SEMI-AUTOMATED DETECTION OF SURFACE DEGRADATION ON BRIDGES BASED ON A LEVEL SET METHOD

KEY WORDS: Deterioration detection, Monitoring, Large infrastructures, Level sets, Image segmentation

ABSTRACT:

Due to the effect of climate factors, natural phenomena and human usage, buildings and infrastructures are subject of progressive degradation. The deterioration of these structures has to be monitored in order to avoid hazards for human beings and for the natural environment in their neighborhood. Hence, on the one hand, monitoring such infrastructures is of primarily importance. On the other hand, unfortunately, nowadays this monitoring effort is mostly done by expert and skilled personnel, which follow the overall data acquisition, analysis and result reporting process, making the whole monitoring procedure quite expensive for the public (and private, as well) agencies.

This paper proposes the use of a partially user–assisted procedure in order to reduce the monitoring cost and to make the obtained result less subjective as well. The developed method relies on the use of images acquired with standard cameras by even inexperienced personnel. The deterioration on the infrastructure surface is detected by image segmentation based on a level sets method. The results of the semi-automated analysis procedure are remapped on a 3D model of the infrastructure obtained by means of a terrestrial laser scanning acquisition.

The proposed method has been successfully tested on a portion of a road bridge in Perarolo di Cadore (BL), Italy.

1. INTRODUCTION

Surfaces of human infrastructures are subject to degradation, mostly due to the effect of (natural) climate elements (sun, rain, wind), to human dependent factors (e.g. increase in the volume of traffic) and to aging.

Nowadays, the monitoring activity of the state of these infrastructures is mostly done by visual inspection performed by expert staff: these highly skilled personnel is involved in periodical recognitions, recording and reporting of the updated infrastructure conditions. It is worth to notice that this procedure typically provides subjective results (conditions are typically subjectively reported on questionnaires) at a quite high cost, related to the relatively frequent use of expert staff.

This paper proposes a different approach in order to reduce the cost of infrastructure monitoring.

The goal of surface infrastructure monitoring is that of periodically updating the detected conditions on a digital representation of the infrastructure, in order to ease the control of the degradation temporal evolution by specific human operators (Costantino and Angelini, 2013a, Costantino and Angelini, 2015, Camarda et al., 2010, Guarnieri et al., 2013).

In this work, the digital 3D representation of the infrastructure is obtained by means of terrestrial laser scanning (TLS). Notice that the use of TLS devices is usually reserved to skilled personnel, hence this is usually a quite expensive operation. For this reason, the use of TLS in this work is limited to una tantum acquisition typically done at the beginning of the monitoring. Nevertheless, other TLS acquisitions can be considered in the following, when required by the occurrence of specific conditions.

Instead, the degradation detections and condition updates are obtained by the analysis of photos acquired by standard digital cameras (not necessarily professional, e.g. cameras embedded in smartphones). Then, the degradation results are mapped onto their correspondent positions on the TLS-based 3D model of the infrastructure.

The main advantages of this integrated strategies approach are as follows:

- The use of digital cameras is commonly considered a much simpler operation than TLS acquisition, hence it can be performed by not highly skilled personnel. This allows a significant monitoring cost reduction, and/or the possibility of more frequent degradation observations.

- Despite some work has been done in order to exploit the luminance of laser scanner radiation in order to classify materials (Costantino and Angelini, 2013b), actually the analysis of the surface degradation status is usually simpler on digital images than on laser scanning data.

- The reprojection of the detected deterioration results on the 3D model of the infrastructure help the operator to realize the overall conditions of the infrastructure (e.g. the areas where deterioration is more remarkable) and to follow the temporal evolution of the degradation.

A quite common choice is to exploit level set methods as an efficient image segmentation tool (Sethian, 1999, Sethian, 2001, Osher and Fedkiw, 2003): accordingly to other previous works (Cerimele and Cossu, 2007, Cerimele and Cossu, 2009), this paper considers the use of level set methods applied to the L^*, a^*, b^* color space. As shown in Section 2., the use of L^*, a^*, b^* instead of the RGB color space allows to successfully reduce the correlation between different color channels.

Details on the whole segmentation procedure are provided in the next section, while the experimental results on a portion of a road bridge in Perarolo di Cadore (BL), Italy, are shown in Section 3. Finally, some conclusions are drawn in Section 4.

2. ASSESSMENT OF DEGRADED SURFACE AREAS

The data analysis procedure works as follows: once the images are acquired they are separately processed by an ad hoc algorithm. Notice that if the number of images is sufficiently large

and with a proper overlapping level, then a 3D photogrammetric reconstruction would be possible as well. Despite this case would be more informative, actually it typically requires an higher level of accuracy during the image acquirement process. Therefore, in order to keep the image capture task as simple as possible, in the proposed method the data analysis is performed on each single image.

The image analysis procedure is user-assisted in its first step (step (a) of the following procedure), whereas the following ones are performed automatically:

(a) The user pinpoints the matches between a set of points in the image and in the 3D infrastructure representation: these correspondences are used for determining the infrastructure surfaces to be controlled and to assess the local map between the image points and their 3D coordinates.

(b) Automatic modal decomposition of the image values (according to a convenient L^*, a^*, b^* representation) on the region of interest.

(c) Region segmentation (i.e. determination of deteriorated areas) based on a level sets method.

(d) Remapping of the segmentation curves on the 3D representation.

The above steps will be detailed in the following.

2.1 Identification of the regions of interest and of the camera pose

The user manually select a set of matching points on the image and on the 3D model. Let $\{M_i\}_{i=1,...,n}$ be the set of (non-coplanar) 3D points and $\{m_i\}_{i=1,...,n}$ the corresponding $2D$ coordinates on the image plane. The distortion caused by the lens of the acquisition camera is assumed to be negligible or already estimated and properly corrected. Hence, without loss of generality, hereafter the coordinates $\{m_i\}_{i=1,...,n}$ on the image plane are assumed to be distortion free.

A subset of the $\{m_i\}$ points is used by the user to outline polygonal curves for delimiting the areas to be analyzed, e.g. areas to be analyzed are delimited by the user by using closed polygonal curves. Each of these areas is assumed to be well approximated by a planar surface. The j-th area is delimited by the points $\{m_{ij}\}_{i=1,...,n_j}$, with $n_j \geq 3$, for each j.

Camera position and orientation with respect to the reference system used to express the $\{M_i\}$ points can be computed in closed-form as described in the following (Ma et al., 2003).

According to the pinhole camera model, undistorted camera measurements can be modeled as follows:

$$m_i = \frac{P_{12}[M_i^\top\ 1]^\top}{P_3[M_i^\top\ 1]^\top}\,, \qquad (1)$$

where P_{12} and P_3 are the first two rows and the third row, respectively, of the camera projection matrix P. The projection matrix P can be expressed in terms of the matrix of inner parameters K and of the camera position t and orientation matrix R with respect to the global reference system: $P = K[R\ -Rt]$.

By simple matrix manipulations of (1), it immediately follows that the value of the matrix P can be estimated (in closed–form) by solving a simple linear system:

$$\begin{bmatrix} -M_i^\top & 1 & 0 & 0 & u_i M_i^\top & u_i \\ 0 & 0 & -M_i^\top & 1 & v_i M_i^\top & v_i \end{bmatrix} p = 0\,, \qquad (2)$$

for $i = 1, \ldots, n$, where p is a unit vector containing the normalized values of P, and $m_i = [u_i\ v_i]^\top$. A scaled version of P can be obtained by simply rearranging the terms in p.

Once P has been computed (up to a scale factor), the matrices K and $[R|-Rt]$ can be computed as the results of the QR factorization, (Golub and Loan, 1989), of the first three columns of P. The scale factor of P can be obtained by imposing that the term on the last column and last row of K is equal to one. Finally, t can be computed by pre-multiplying the forth column of P with $-R^{-1}K^{-1}$.

The above estimation can be improved by using bundle adjustment–like procedures in order to obtain more accurate estimations (Ma et al., 2003), if needed.

Points $\{m_i\}_{i=1,...,n}$ and $\{M_i\}_{i=1,...,n}$ will be used also in subsection 2.5 to map the 2D segmentation results onto the 3D bridge points.

2.2 L^*, a^*, b^* representation

Images acquired by standard cameras are usually represented in the RGB (red, green, blue) color space. Despite being convenient for the visualization on standard displays, this space is not the most suitable when dealing with segmentation purpose. Indeed in the RGB representation each color channel is typically highly correlated with the others, making the analysis of the three channels mostly redundant.

In order to tackle this issue, the L^*, a^*, b^* representation for the image colors is considered as it allows to obtain much less correlated color channels.

The conversion between RGB and L^*, a^*, b^* color space is usually described by passing through the XYZ representation:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.4125 & 0.3576 & 0.1804 \\ 0.2127 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9502 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \qquad (3)$$

and

$$\begin{aligned} L^* &= 116 f(Y/Y_n) - 16 & (4) \\ a^* &= 500[f(X/X_n) - f(Y/Y_n)] & (5) \\ b^* &= 200[f(Y/Y_n) - f(Z/Z_n)] & (6) \end{aligned}$$

where X_n, Y_n, Z_n represent the (normalized) values representing the white color in the XYZ representation, and the function $f(\cdot)$ is defined as follows:

$$f(t) = \begin{cases} t^{1/3} & \text{if } t > (6/29)^3 \\ \left(\frac{29}{6}\right)^2 \frac{t}{3} + \frac{4}{29} & \text{otherwise.} \end{cases} \qquad (7)$$

Figure 1: Example of region to be segmented.

In order to validate the above considerations on correlations among RGB and L^*, a^*, b^* channels, the Pearson correlation coefficient has been computed as well. In the case of the image shown in Fig. 1, this coefficient results to be 0.97 for the red and green color channels (Fig. 2), whereas it is 0.03 for the L^* and a^* channels (Fig. 3). Correlation between the red and green channel is also apparent in Fig. 2, while L^* and a^* are weakly correlated as shown in Fig. 3. Similar considerations can be repeated for the the other possible couples of channels.

Figure 2: Red (top) and green (bottom) color channels for the image in Fig. 1. Brighter pixels are associated to higher values of the considered variables.
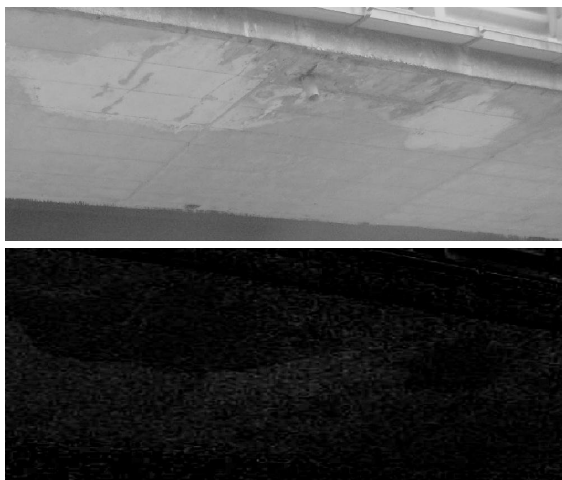

Figure 3: L^* (top) and a^* (bottom) channels for the image in Fig. 1. Brighter pixels are associated to higher values of the considered variables.

To simplify the notation hereafter a one dimensional signal will be considered (i.e. the lightness L^*). Nevertheless the approach can be extended to the three dimensional signal case.

2.3 Lightness multimodal density

The goal of this step is to estimate the probability density of the lightness L^* values in the region of interest: such density is usually a multimodal density, where each mode is typically associated to a region with different characteristics. Hence, the rational is that of obtaining a rough segmentation of the region in different areas according to the distribution of the lightness values in different modes.

The histogram of the lightness values will be considered as a sample approximation of the probability density. Fig. 4 shows the histogram corresponding to the multimodal lightness density of image in Fig. 1.
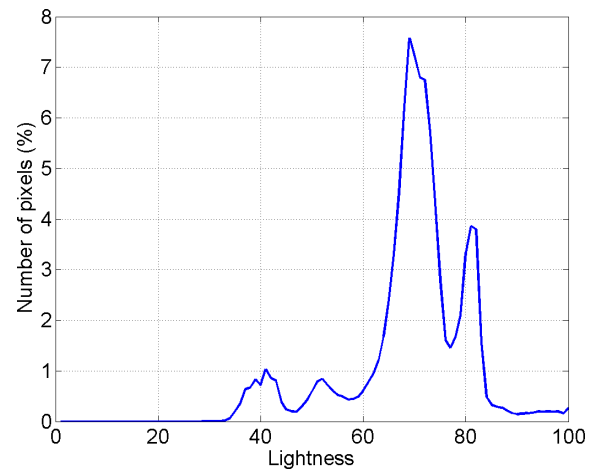

Figure 4: Example of histogram of lightness values, corresponding to the image in Fig. 1.

It is assumed that the mode with the largest number of associated pixels corresponds to the non-deteriorated region, this way, it is possible to easily distinguish between non-deteriorated and deteriorated regions (those corresponding to the larger mode and to the others, respectively). Similar considerations allow to draw analogous conclusions if the algorithm can associate the most significant mode to deteriorated regions.

Otherwise, when neither of the above cases can be considered, associating a different region to each mode (i.e. to each local maximum in the probability density), it is possible to segment regions characterized by different statistics, but without automatically distinguishing which correspond to the deteriorated ones

Once the number of modes m has been determined (for instance by using local maxima detection), the separation between different modes is done on the histogram based on the Otsu's m-thresholding method (Otsu, 1975, Otsu, 1979). Otsu's method provides the threshold τ_i to separate the histogram bins associated to the i-th detected mode from those associated to the $(i+1)$-th, for $i = 1, \ldots, m-1$.

In order to simplify the presentation, without loss of generalization hereafter only two regions are considered, i.e. $m = 2$, separable by means of the threshold τ. The generalization to a generic value of m is immediate.

2.4 Region segmentation

Region segmentation considered in this subsection is the final step to be done in order to separate deteriorated areas from the non-deteriorated ones. The approach considered here is level set segmentation-like (Sethian, 1999, Sethian, 2001, Osher and Fedkiw, 2003).

First, a rough segmentation is obtained by means of the model decomposition presented in the previous subsection: the number of segmented areas to be considered is set equal to the number of detected modes m. Then, a pixel is assigned to the i-th segmented area if its lightness value has been assigned in the previous subsection to the i-th mode, for $i = 1, \ldots, m$.

Unfortunately, the above rough segmentation is usually very noisy, hence the following level set method is applied:

- The closed curve Γ separating the two regions of interest on the image is implicitly described as

$$\Gamma = \{(x, y) \mid \phi(x, y) = 0\} , \qquad (8)$$

where (x, y) are the coordinates on the image domain and the level set function $\phi(x, y)$ will be described in the following.

- The level set function is initialized as $L^*(x, y) - \tau$, then it is evolved by considering both the positions where edges are more probable in the images (i.e. where it is more likely to have the border between the two regions) and the regularity (i.e. smoothness) of the curve:

$$\frac{\partial \phi}{\partial t} = -S \cdot \nabla \phi + k|\nabla \phi| , \qquad (9)$$

where the first term on the right side of the above equation, $S \cdot \nabla \phi$, corresponds to the introduction of an external vector field related to the image edges, whereas the second term, $k|\nabla \phi|$, aims at increasing the curve smoothness.

- In this work the external vector field S corresponds to an edge detection function. Let I_G be the original image filtered by a Gaussian and LoG be the Laplacian of Gaussian filter:

$$I_G = G * I , \qquad (10)$$

and

$$LoG(x, y) = -\frac{1}{\pi\sigma^4}\left(1 - \frac{x^2 + y^2}{2\sigma^2}\right)e^{-\frac{x^2+y^2}{2\sigma^2}} , \quad (11)$$

where $*$ stands for the 2-dimensional convolution and the Gaussian filter G is defined as follows

$$G(x, y) = \frac{1}{2\pi\sigma^2}e^{-\frac{x^2+y^2}{2\sigma^2}} . \qquad (12)$$

Let $I_L = LoG * I$, then, S on the point (x, y) is defined as follows:

$$S(x, y) = I_L(x, y)\frac{\nabla I_G(x, y)}{|\nabla I_G(x, y)|} , \qquad (13)$$

that is, the magnitude of the external field in (x, y) is determined by $I_L(x, y)$ (i.e. the value of the original image I filtered with the LoG operator), whereas the direction of the the external field is determined by the filtered gradient of the image, $\nabla I_G(x, y)$, normalized by its length. Where the gradient is equal to 0, $S(x, y)$ is set to be a zero vector as well.

2.5 Remapping of the segmented regions

Once the regions have been segmented the resulting boundaries have to be remapped on the 3D space in order to spatially positioning them on the infrastructure.

By assumption each of the areas to be segmented can be well approximated as a planar surface, hence a local two-dimensional coordinate system can be defined on such surface.

Consider the j-th analyzed area, delimited by the points $\{m_{ij}\}_i$, $i = 1, \ldots, n_j$, with $n_j \geq 3$, for each j. Since the considered area can be approximated as a planar surface, each of its point can be expressed as a linear combination of 3 non–collinear points in $\{m_{ij}\}_i$, $i = 1, \ldots, n_j$. Without loss of generalization, let m_{i1}, m_{i2}, m_{i3} be non-collinear.

Furthermore, let $u = m_{i1}$ and let U be the result of the Gram-Schmidt orthogonalization of the matrix $[m_{i2} - m_{i1} \quad m_{i3} - m_{i1}]$.

Then, each point M^* of the planar surface of interest can be expressed as follows:

$$M^* = u + Ux , \qquad (14)$$

$x \in \mathbb{R}^2$, where, with a slight abuse of notation, hereafter x corresponds to the point coordinates according to the new reference system.

Let M^* be the 3D point associated to a 2D point m^* on the segmentation boundary. Accordingly to (1), the 3D position M^* can be obtained from the following:

$$(m^* P_3 - P_{12})\begin{bmatrix} M^* \\ 1 \end{bmatrix} = 0 . \qquad (15)$$

Substituting (14) in the above equation,

$$(m^* P_3 - P_{12})\left(\begin{bmatrix} u \\ 1 \end{bmatrix} + \begin{bmatrix} U \\ 0 \end{bmatrix}x\right) = 0 , \qquad (16)$$

and, after simple matrix manipulations of the above equation:

$$x = -\left((m^* P_3 - P_{12})\begin{bmatrix} U \\ 0 \end{bmatrix}\right)^\dagger (m^* P_3 - P_{12})\begin{bmatrix} u \\ 1 \end{bmatrix} . \qquad (17)$$

Finally, M^* can be obtained from (14).

3. RESULTS

The proposed method has been tested on a road bridge located in Perarolo di Cadore, a small village in the Italian alps. The considered bridge is shown in Fig. 5, whereas the cloud points of a portion of the bridge (acquired by means of TLS) is shown in Fig. 6.

The image segmentation results obtained (by means of the level set method described in the previous section) for the image in Fig. 1 are shown in Fig. 7. The region considered in the analysis of this image is a portion of the planar surface under the bridge, delimited by green lines in the figure. Instead, segmented regions are delimited by red lines: the regions corresponds to degraded areas, where degradation has been probably caused by percolation issues. It is worth to notice that segmentation shown in the figure corresponds to the separation between the two largest modes in Fig. 4: this is sufficient to distinguish the areas where degradation is more clear, whereas, if needed, segmentation based

Figure 5: Road bridge located in Perarolo di Cadore considered in as test case in this work.
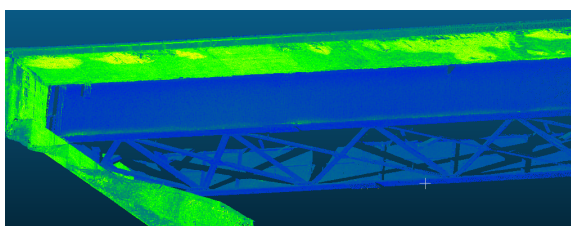

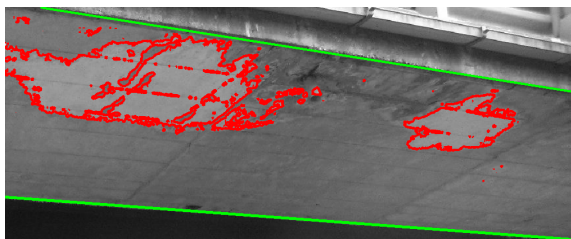Figure 6: Cloud points of a portion of the road bridge of Fig. 5.

Figure 7: Example of segmentation results, corresponding to the image in Fig. 1. The region considered for the analysis is delimited by the green lines, whereas boundaries of the segmented areas are indicated by red lines.

on the other modes with lower intensity in Fig. 4 can be used to distinguish other areas in the figure with other (i.e. lower) levels of degradation.

Similarly, the results for another portion of the bridge are shown in Fig. 8.
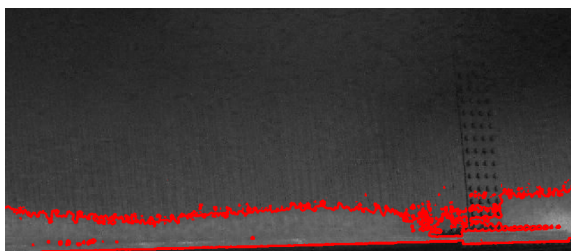

Figure 8: Example of segmentation results.

Once the 2D regions have been segmented, their boundaries can be remapped on the 3D space accordingly to the procedure described in subsection 2.5. For instance, the segmentation bound-

aries in Fig. 7 can be remapped as shown in Fig. 9 (in order to improve the readability of the figure, x points (computed as in (17)) are shown instead of M^*).
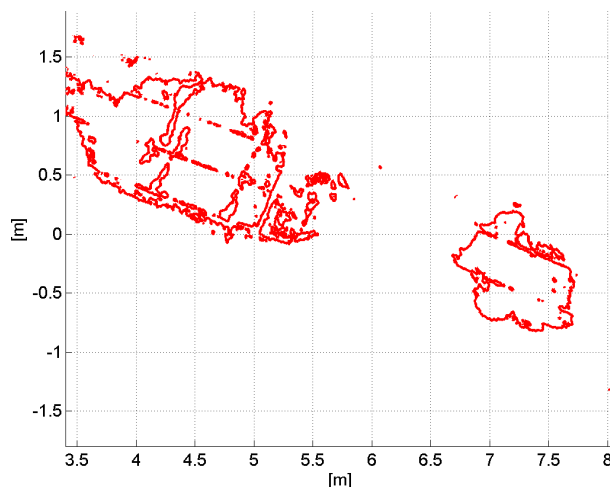

Figure 9: Example of segmentation results: boundaries of detected degradation areas of Fig. 7 remapped in the 3D space (in order to improve the readability of the figure, x points are shown instead of M^*).

Actually, the segmentation boundaries can be remapped on the 3D-model as well: Fig. 10 shows the curve formed by the segmentation boundary points $\{M^*\}$ (obtained from (14)) superimposed on the cloud points data. In order to ease the visualization of the results the figure shows a zoom on a specific (and significant) portion of the original image, viewed from an observation direction approximately orthogonal to the analyzed surface.
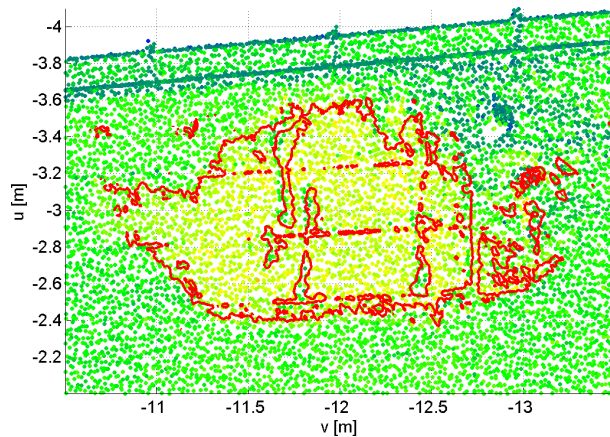

Figure 10: Example of segmentation results: a portion of the segmentation boundaries (red lines) remapped in the 3D space. Boundary lines are superimposed to the cloud points acquired by means of TLS. Observation direction is approximately orthogonal to the analyzed surface.

4. DISCUSSION AND CONCLUSIONS

This paper proposed a semi-automated method that exploits level set image segmentation in order to estimate the boundaries of degraded regions on complex infrastructures, e.g. the bridge considered in this work.

As shown in the obtained results, the method allows to obtain

appropriate 2D region boundaries (Fig. 7 and 7), which can be remapped on the 3D-model as well (Fig. 9 and 10).

Despite our results are quite encouraging, some work has still to be done in order to make the whole procedure as simple as possible to the user. In particular, the case of the analysis of periodicals recordings of the same infrastructure will be considered in the future. The rational in such case is that of properly exploiting information already available from previous analyzed data to reduce the interaction with the user, i.e. to make the procedure more autonomous.

Furthermore, this work relies on the use of differences in the lightness values in order to distinguish different areas. Despite this is quite intuitive, it might be subject to errors in certain cases. Future investigation will be dedicated to the use of different image analysis techniques (e.g. (Geladi and Grahn, 1996, Facco et al., 2011, Facco et al., 2013)) in order to provide more robust results from this point of view.

Finally, our future investigations will also consider the evaluation of possible issues related to the use of reconstructed 3D model (Tucci et al., 2012).

REFERENCES

Camarda, M., Guarnieri, A., Milan, N., Vettore, A. et al., 2010. Health monitoring of complex structure using tls and photogrammetry. Proceedings of the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Newcastle upon Tyne, UK pp. 21–24.

Cerimele, M. M. and Cossu, R., 2007. Decay regions segmentation from color images of ancient monuments using fast marching method. Journal of Cultural Heritage 8(2), pp. 170–175.

Cerimele, M. M. and Cossu, R., 2009. A numerical modelling for the extraction of decay regions from color images of monuments. Mathematics and Computers in Simulation 79(8), pp. 2334–2344.

Costantino, D. and Angelini, M., 2013a. Topographic survey for structural monitoring, case: Quadrifoglio condominium (lecce). ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XL-5/W3, pp. 179–187.

Costantino, D. and Angelini, M., 2015. Three dimensional integrated survey for building investigations. Journal of Forensic Sciences.

Costantino, D. and Angelini, M. G., 2013b. Qualitative and quantitative evaluation of the luminance of laser scanner radiation for the classification of materials. ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 1(2), pp. 207–212.

Facco, P., Masiero, A. and Beghi, A., 2013. Advances on multivariate image analysis for product quality monitoring. Journal of Process Control 23(1), pp. 89–98.

Facco, P., Masiero, A., Bezzo, F., Barolo, M. and Beghi, A., 2011. Improved multivariate image analysis for product quality monitoring. Chemometrics and Intelligent Laboratory Systems 109, pp. 42–50.

Geladi, P. and Grahn, H., 1996. Multivariate image analysis. John Wiley & Sons Inc., New York (U.S.A.).

Golub, G. and Loan, C. V., 1989. Matrix Computations. Johns Hopkins University Press: Baltimore, MD.

Guarnieri, A., Milan, N. and Vettore, A., 2013. Monitoring of complex structure for structural control using terrestrial laser scanning (tls) and photogrammetry. International Journal of Architectural Heritage 7(1), pp. 54–67.

Ma, Y., Soatto, S., Košecká, J. and Sastry, S., 2003. An Invitation to 3D Vision. Springer.

Osher, S. and Fedkiw, R., 2003. Level set methods and dynamic implicit surfaces. Vol. 153, Springer Science & Business Media.

Otsu, N., 1975. A threshold selection method from gray-level histograms. Automatica 11(285-296), pp. 23–27.

Otsu, N., 1979. A threshold selection method from gray-level histograms. Systems, Man and Cybernetics, IEEE Transactions on 9(1), pp. 62–66.

Sethian, J., 1999. Level Set Methods and Fast Marching Methods: Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Materials Science. Cambridge Monographs on Applied and Computational Mathematics, Cambridge University Press.

Sethian, J., 2001. Evolution, implementation, and application of level set and fast marching methods for advancing fronts. Journal of Computational Physics 169(2), pp. 503 – 555.

Tucci, G., Cini, D. and Nobile, A., 2012. A defined process to digitally reproduce in 3D a wide set of archaeological artifacts for virtual investigation and display. Journal of Earth Science and Engineering 2(2), pp. 118–131.