AUTOMATED MULTISPECTRAL TEXTURE MAPPING OF 3D MODELS

A. Pelagotti[a]*, A. Del Mastio*[b]*, F. Uccheddu*[b] *and F. Remondino*[c]

[a] CNR - Istituto Nazionale di Ottica Applicata (INOA), Florence, Italy – anna.pelagotti@inoa.it, www.inoa.it
[b] Department of Electronics and Telecommunications, University of Florence, Italy – <delmastio, uccheddu>@lci.det.unifi.it
[c] 3D Optical Metrology Group, Bruno Kessler Foundation (FBK), Trento, Italy – remondino@fbk.eu, www.fbk.eu

ABSTRACT

Recently, thanks to the diffusion of scanning devices and the availability of powerful 3D modelling software, as well as to the improvements in the automation of the image-based modelling approach, 3D models are more and more considered in several research fields, such as inspection, navigation, object identification, visualisation and animation, and recently also in Cultural Heritage. Documentation in case of loss or damage, virtual tourism and museum, education resources, interaction (without risk of damage), and so forth are only few examples of applications where they has become a fundamental step.

However, 3D models obtained with these technologies are often lacking a suitable photorealistic appearance, due to low quality acquired texture, or to the complete absence of it.

Moreover, it is often of specific interest to texture with images different from photos, like multispetral/multimodal images (IR, Xrays, UV fluorescence). In such cases, a post-processing texture mapping is needed, and it is often achieved through manual alignment of the model and the related texture. In this work, we propose a fully automatic approach for multispectral texture mapping. The method relies on the extraction from the model geometry of a depth map, in form of an image, whose pixels maintain an exact correspondence with vertices of the 3D model; the subsequent step is registratio , between such image and the chosen texture, with a very robust registration algorithm, based on Maximixation of Mutual Information.

The results presented show the effectiveness of the proposed method.

1. INTRODUCTION

3D models are currently being used in a variety of applications. In order to increase their visual appearance and realism, a texture is generally applied to them. This texture can be a uniform or a faded color, or, more frequently, one or more images of the object. In 3D models created with photogrammetric methods as well as with some of the more recent range sensors, the shape (x,y,z) data and the colour (R,G,B) data are already available in perfect registration [19]. However the quality of the recorded colour is not always optimal, and, moreover, it is not always the desired texture.

One of the possible applications of textured 3D models is a quantitative evaluation or study of the materials lying on the object surface. In this case multispectral/multimodal images, such as IR or UV fluorescence images may be used to texture the 3D geometry and infer important information about it. For example, in case of paintings, the combination of 3D data with an UV fluorescence image helps to quantify the extension of retouches, which is an information of crucial importance for conservators.

Several methods for texturing a 3D model have been proposed in literature. However, very frequently, they present the need to manually selecting a number of common points between the 3D model and the texture, applying the well known DLT (Direct Linear Transformation) registration method [1]. When the texture has a visual appearance which retains only few common features with the 3D model, as in case of multispectral/multimodal images, the task is complex and might produce inaccurate results. In this paper, we propose a method for automatic texturing 3D models. The algorithm is based on the computation, on a regular grid, of the depth map of the 3D geometry. The depth map is then transformed into a grey-scale image, where the grey level is related to the z coordinate of model.

Computing the depth map of the model from its 3D coordinates preserves the relation between its vertices and map pixels, since for each vertex a single pixel of the map can be found. A custom developed robust registration algorithm, based on Maximization of Mutual Information (MMI) [2], is then used to automatically register a multispectral view of the object with the map. The MMI algorithm shows good performances even when the two images retain scarce similarities with each other. Automatic texturing of the 3D model of the object with multispectral/multimodal images is achieved as a result of a 2D to 2D image registration.

2. 3D MODELING AND TEXTURE MAPPING

There are several methods to produce digital 3D models and usually they are selected according to project requirements, users experience, object's location and project's budget. The actual reality-based 3D modeling technologies involve mainly optical range-based active sensors [3], image-based passive sensors [4] or an integration of them [5], trying to exploit the intrinsic potentialities of each technique. Once the measurements are performed, the 3D geometry is generally textured for photo-realistic results. In fact, the visualization of a 3D model is often the only product of interest for the external world and remains the only possible contact with the model. Therefore a photo-realistic and accurate product is often required and extremely important.

The texture mapping process is generally intended as the mapping of colour information onto the 3D data, which are in form of points or triangles (mesh). The texturing of 3D point clouds (namely point-based techniques [6]) allows a faster visualization but for detailed and complex 3D models it is not an appropriate method. In case of meshed data, homologues points between the 3D mesh and the 2D image to-be-mapped should be identified. This is the bottleneck of the texturing phase as it is still an interactive procedure and no automated and reliable approaches were proposed yet. Indeed the identification of homologues points between 2D and 3D data is a hard task, much more complex than image to image or geometry to geometry registration. Furthermore, in applications involving infrared or multispectral images, it is generally quite challenging to identify common features between 2D and 3D data. In the general case, the classical DLT approach [1] (often referred ad Tsai method [7]) is used to retrieve the intrinsic and extrinsic unknown camera parameters and then map the color information on the surface polygons using a color-vertex encoding or a mesh parameterization.

It is clear that in order to have a satisfying texture mapping process it is often not enough to project one or more static images over the 3D geometry. Even when an acceptable image-to-geometry registration is achieved, problems rise because of variations in lighting, surface specularity, camera settings and occlusions. Generally the images are exposed with the illumination at imaging time but it may need to be replaced by illumination consistent with the rendering point of view and the reflectance properties (BRDF) of the object [8]. High dynamic range (HDR) images might also be acquired to recover all scene details [9]. Methods to reduce color discontinuities, aliasing effects and render seamless textured surfaces have been presented in [10, 11]. A methodology to create high-resolution image texture was presented in [12]. An algorithm to generate occlusion-free images for more pleasant and photo-realistic texture mapping applications was presented in [13].

3. THE PROPOSED METHOD

The developed method for automated texture mapping of different typologies of 3D models consists of the following steps:

1. automated generation of a depth map image of the 3D model to produce an intrinsic mapping between depth map pixels and the corresponding vertices of the model

2. registration of the depth map with the image to-be-mapped using a registration method based on the mutual information [2] (see Figure 1). However other robust automated registration methods may also be successfully applied.

3.1 The depth map

Starting from the geometric 3D model, a depth map is built as a two-dimensional array such that the x and y vertices coordinates of the model correspond to the rows and columns of the array (as in an ordinary image). The corresponding depth readings (z values) are referred to a cutting plane and stored in the array's element value (see Figure 2). Such a map is thus a grey scale image where the z information is stored in the intensity information, being therefore a projection of the 3D coordinates into the x y plane. The so generated map intrinsically keeps an exact correspondence between 2D pixels and 3D vertices of the model.

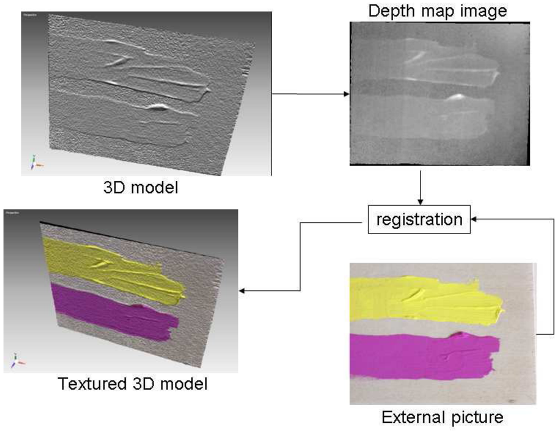

Figure 1 – Sketch of the proposed method. From the 3D model a depth map image is generated. This is automatically registered with an external image and the 3D model textured.

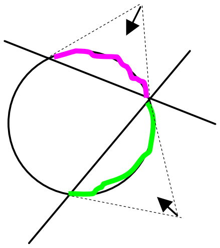

Figure 2 – In case of a solid of revolution the z-data considered is the off-set between the analytical approximation surface of revolution and the actual 3D model. In case of a generic model, we need to consider cuts of the model and the z-data is to be referred to the cut plane

As the vertices of a 3D model are not usually placed along a regular grid, the projected depth map could have sparse values or holes. While the ratio between rows and columns of the map depends on the ratio between the x and y dimensions of the 3D model bounding box, its specific dimensions vary according to the chosen area of each entry/pixel. By associating at most one vertex per pixel, the pixel area is small and several pixels are “empty”, thus generating holes in the map where vertices are far away from their neighbours. On the other hand, by considering a bigger pixel area, more than one vertex could fall in the same pixel, thus leading the system to make a choice. The chosen strategy has been that of selecting the vertex with greater depth value, since generally more representative of the details “emerging” from the surface. The 3D shape is highlighted by the depth map image, and is often related with some “visual” details depicted in the to-be-registered image: for example, a brushstroke offers both a 2D information, due to its colour or shading, and a 3D structure, due to its thickness. This relation is exploited for the following registration step, thus conveying the correspondence between the model vertices and the map pixels to the to-be-mapped image.

3.2 Registration step

Since the synthetic depth map image is intrinsically mapped to the model, a texture mapping with a different image can be accomplished, by registering such an image with the depth image itself, and then conveying the obtained correspondence to the 3D model.

Image registration is a fundamental step in many image processing and analysis tasks, such as in remote sensing, in medicine, in computer vision, and more recently in Cultural Heritage applications (multispectral analysis of pigments) [14, 15, 16, 18]. It aims to determine the correct displacement (that is a geometrical transformation) to align two views of the same subject.

In our work, we exploited a powerful automatic registration technique [2]; it is based on the computation of the *Mutual Information (MI)* between the two images, which is a similarity measure coming from the Information Theory, and it is a measure of how much information one image shares with another one. The Maximization of Mutual Information (MMI) criterion states that when the two images are correctly aligned, MI assumes its maximum value [17]. MMI is a very general and powerful criterion, since no assumptions are made over the nature of this dependence and no constraints are posed on the image contents, thus enforcing MMI criterion to be particularly effective when a low amount of information is shared between the two images (such as in multimodal images registration).

Given two grey-scale images X and Y, related by a geometric transformation T_α such that the pixel p of X (whose intensity is x) corresponds to the pixel $T_\alpha(p)$ of Y (whose intensity is y), the MI of the two images is given by:

$$MI(X;Y)=\sum_{x,y} p_{XY}(x,y)\cdot\log_2\frac{p_{XY}(x,y)}{p_X(x)\cdot p_Y(y)}$$

where $p_{XY}(x,y)$ is the joint probability distribution of the two images, and $p_X(x)$ and $p_Y(y)$ are their marginal probability distribution. The goal of the registration step is to find the transformation maximizing the MI among a suitable set of transformations.

Marginal and joint probability distributions can be estimated by normalization of the *joint histogram* (say $h_\alpha(x,y)$) of the two images, obtained in turn by binning the intensity value pairs $x=X(p)$ and $y=Y(T_\alpha(p))$, for all the pixels and depending on each accounted transform [16]:

$$p_{XY,\alpha}(x,y)=\frac{h_\alpha(x,y)}{\sum_{x,y}h_\alpha(x,y)} \quad p_{X,\alpha}(x)=\sum_{y}p_{XY,\alpha}(x,y) \quad p_{Y,\alpha}(y)=\sum_{x}p_{XY,\alpha}(x,y)$$

The maximum MI will be reached when (i) the cut plane will be perfectly perpendicular to the camera position at the moment of the texture acquisition, (ii) the scale between model and texture will be the same and (iii) the relative translation and rotation agree. Therefore, the algorithm should try every possible position and for each position evaluate the MI. In order to obtain optimal results, the texture image should be undistorted, otherwise the registration algorithm should also evaluate the MI trying to compensate for the deformations of the texture image, at the price of a higher computational time of the algorithm. In this work, we took into account an affine transformation as deformation model for the to-be-registered image, that is a combination of translation (along the rows and columns direction separately), a rotation with respect to the center of the image, and a scaling (the same factor both for height and width). We also chose an initial cut plane which was a good approximation of the correct one.

Since the applied transformation can lead to non-integer values for the transformed pixel coordinates, the bilinear interpolation technique was used. The exhaustive search was also used as maximizing scheme; the elapsed time thus depends on the range and on the step chosen for each one of the parameters. In order to provide a rough indication of the processing time, for a target 500 px wide by 583 px tall image (Figure 6), the elapsed time for each iteration (Intel Core 2 – 2.00 GHz – 2.00 GB RAM) was of 7.1118 sec.

4. RESULTS AND DISCUSSION

To demonstrate the effectiveness of the proposed method, we generated several 3D models and textured them with different kind of views of the objects. The models were produced both with photogrammetry and range sensors. The 3D models used were not always perfect. This choice was made on purpose in order to test the effectiveness of the proposed method. Our first goal was to substitute the texture of a 3D model which was not corresponding with the actual visual appearance, with a colour photograph.

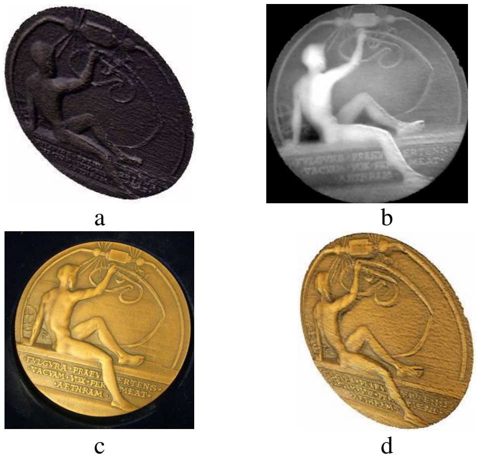

Figure.3 - In (a) the textured model automatically generated is shown. In (b) the computed depth map generated from the scanned 3D model, where the relieves are clearly visible, varying from the dark (lower areas) to the white (upper areas), is displayed. Once the photo (c) of the coin has been correctly registered with the map, it is possible to accomplish the texture mapping, leading to the model depicted in (d).

In Fig. 3 it is shown a comparison between a 3D model with a low quality texture generated by the laser scanner

(Fig. 3a), and the model with an higher quality texturing generated with a photograph and our texture mapping system (Fig. 3d). It is clearly visible the more realistic appearance.

We then wanted to apply a texture which could not be acquired with a colour 3D laser scanner. In Fig. 4 we show the results obtained by texturing the 3D geometry with a UV Fluorescence image, which is not directly achievable with any colour 3D laser scanner, since it needs a special illuminator. Visible UV Induced Fluorescence is an investigation technique, which is capable of highlighting some features of paintings, such as restored areas which often are not to be distinguished to the naked eye. Texturing the 3D model of this object with the fluorescence enabled restorers to have a much more accurate map of restored areas.

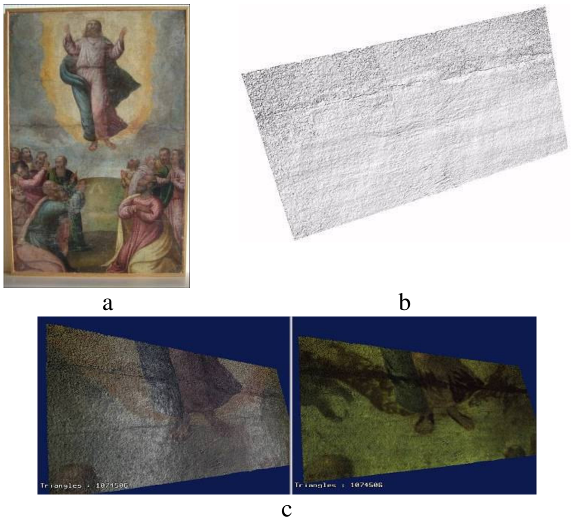

Figure 4 – (a) a colour image of a XVIII century painting, and its 3D model (b); (c) the same detail of the 3D model textured with a colour image, on the left, and a UV fluorescence image, on the right. Retouchings are clearly visible as darker areas.

In Fig. 5 three different textures, i.e. standard photograph for the appearance under natural light (Fig. 5a), the UV fluorescence image (Fig. 5b) and a IR thermal map (Fig 5c), are used to texture the same 3D model (Fig. 5d-e-f) of a XIX century vase, which presents restoration interventions which are invisible to the naked eye, but are documented using these other imaging techniques. In fig. 6 the results on challenging test. We borrowed a low resolution model of Donatello's Maddalena and downloaded some free photos from the web. Despite the relatively little amount of information shared between the produced depth map and the image to-be-mapped, a proper registration was achieved.

5. CONCLUSIONS

In this paper, an automatic texture mapping technique has been presented. The method is based on the computation of the depth map image of a 3D model as a two-dimensional array where the x and y-coordinates of the model corresponds to rows and columns of the array, as in an ordinary image, and the z-coordinates (depth) are stored in the corresponding array values. The depth map image maintains intrinsically the connection with the vertices of the 3D model; thus, simply registering a view of the object with the depth map image, it is possible to accomplish an accurate texture mapping of the model itself.

The proposed method overcomes the need of a manual detection of common points in the model and in the texture image. Since the process is completely automated, the accuracy depends on the registration algorithm, while manual detection is employed correctness of mapping is very subjective. The method has been initially developed for flat-like models, thus is applicable when dealing with objects such as paintings, coins, or in general objects whose depth is extremely lower with respect to its height and width.

The main challenges in case of a generic models are linked to the fact that the results of the registration process depend on the degrees of freedom available in the chosen implementation of the transformation algorithm, and, moreover, to the actual overlap between the texture and the range data map created. In case of revolution solids, the depth map will be computed as a difference between the actual z-coordinate and the z-coordinate of an ideal solid similar to it, and the texture can be a panoramic view of the object. In case of generic solids , we need to consider cuts of the model and the z-data is to be referred to the cut plane. The method proved to be robust also for generic model.

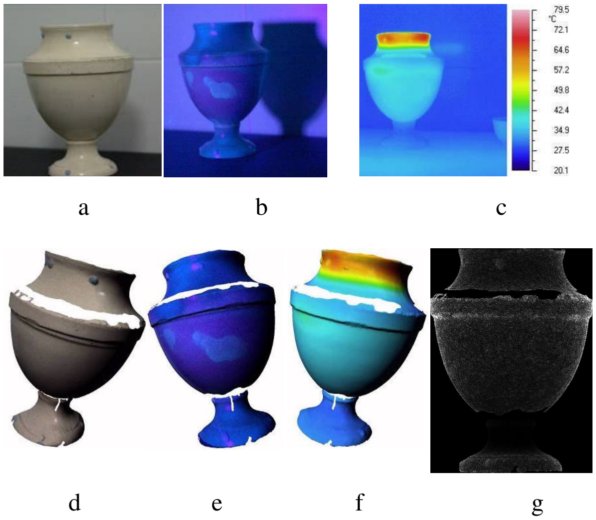

Figure 5 – (a) a RGB photograph of a XVIII century pot, (b) a UV light photograph, (c) a thermal IR image, (g) the depth map image of the pot. Range-based 3D models of the pot textured respectively with a visible (d), UV fluorescence (e), near-IR image (f). The texturing was correctly performed despite missing data in the 3D model

ACKNOWLEDGMENTS

The authors would like to thank R. Vannetti for his contribution to the presented work. We are also thankful to Gabriele Guidi (Polythecnic of Milan, Italy), for providing us the Donatello's Maddalena 3D model. The sculpture is kept in Museo dell'Opera del Duomo, Florence, Italy. The 3D model was acquired during the campaign reported in [20].

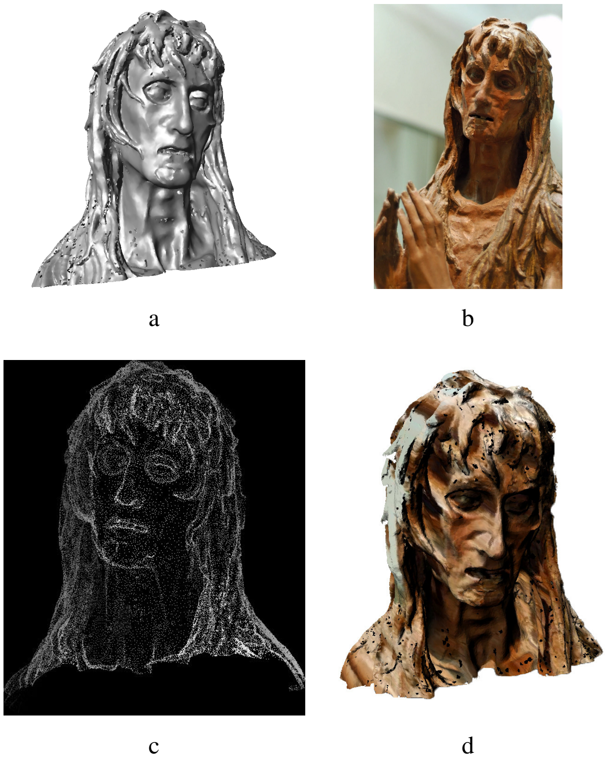

Figure 6 – 3D geometric model of Donatello's Maddalena head (a) textured with an Internet image (b) aligned onto the depth map (c) with the proposed method. The final textured 3D model is shown in (d)

REFERENCES

[1] Y.I,Abdel-Aziz and H.M. Karara, "Direct linear transformation from comparator coordinates into object space coordinates in close-range photogrammetry", in Proceedings of the Symposium on Close-Range Photogrammetry (pp. 1-18). Falls Church, VA: American Society of Photogrammetry, 1971

[2] A. Pelagotti, A. Del Mastio and A. Razionale "Active and passive sensors for artworks analysis and investigation", Videometrics IX Proc. SPIE-IS&T, Electronic Imaging 2007, Vol. 6491

[3] F. Blais, "A review of 20 years of Range Sensors Development". *Journal of Electronic Imaging*, Vol. 13(1), pp. 231-240, 2004

[4] F. Remondino., S. El-Hakim, "Image-based 3D modelling: a review", Photogrammetric Record, Vol. 21(115), pp. 269-291, Sept. 2006

[5] G. Guidi, F. Remondino, M. Russo, F. Menna, A. Rizzi, "3D modeling of large and complex site using multi-sensor integration and multi-resolution data", in Proc. of 9th Int. Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST'08), pp. 85-92, Braga, Portugal, Dec. 2008

[6] L. Kobbelt and M. Botsch, "A survey of point-based techniques in computer graphics", *Computers and Graphics*, Vol. 28(6), 801–814, 2004

[7] R.Y. Tsai, "A versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses", *IEEE Journal of Robotics and Automation*, Vol. RA-3(4), August 1987, pp. 323-344

[8] H. P. A. Lensch, J. Kautz, M. Goesele, W. Heidrich, and H.-P. Seidel, "Image-based reconstruction of spatial appearance and geometric detail", ACM Transaction on Graphics, 22(2), 234–257, 2003.

[9] E. Reinhard, G. Ward, S. Pattanaik, and P. Debevec, "High dynamic range imaging: acquisition, display and image-based lighting", Morgan Kaufmann Publishers, USA, 2005.

[10] K. Umeda, M. Shinozaki, G. Godin, and M. Rioux, "Correction of color information of a 3d model using a range intensity image", in Proc. of 5th International Conference on 3-D Digital Imaging and Modeling (3DIM'05), pp. 229–236, 2005.

[11] M. Callieri, P. Cignoni, M. Corsini, and R. Scopigno, "Masked photo blending: Mapping dense photographic dataset on high-resolution sampled 3d models", Computer and Graphics, 32(4), 464-473, 2008.

[12] F. Remondino and J. Niederoest, "Generation of high resolution mosaic for photo-realistic texture-mapping of cultural heritage 3D models", in Proceedings of the 5th International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST04), 7-10 December 2004, Brussels, Belgium. 279 pages: 85-92

[13] D. Ortin and F. Remondino, "Generation of occlusion-free images for realistic texture mapping", International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 2005, 36 (5/W17)

[14] B. Zitovà, and J. Flusser, "Image registration methods: A survey", *Image and Vision Computing*, vol. 21, pp. 977–1000, June 2003

[15] J. M. Fitzpatrick, D. L.G. Hill, and C. R. Maurer Jr., *Handbook of Medical Imaging - Volume 2, Medical Image Processing and Analysis - ch. 8, Image Registration*. Address: SPIE PRESS, Bellingham, Washington, 2000.

[16] V. Cappellini, A. Del Mastio, A. Pelagotti and A. Piva, "An automated registration technique for Cultural Heritage images", in *8th International Conference on Non Destructive Investigations and Micronalysis for the Diagnostics and Conservation of the Cultural and Environmental Heritage*, Lecce (Italy), May 2005

[17] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information", *IEEE Trans. Medical Imaging*, vol. 16, pp. 187–198, 1997

[18] A. Pelagotti, A. Del Mastio, A. De Rosa and A. Piva, "Multispectral imaging of paintings", *IEEE Signal Processing Magazine*, Vol. 25 n.4, pp. 27-36, July 2008

[19] F. Remondino, S.F. El-Hakim, and A. Gruen, "Turning Images into 3-D models", *IEEE Signal Processing Magazine*, Vol. 25 n.4, pp. 55-64, July 2008

[20] G. Guidi, J.-A. Beraldin, C. Atzeni, "High accuracy 3D modeling of cultural heritage: the digitizing of Donatello's Maddalena", *IEEE Transaction on Image Processing*, 13, (3), pp 370–380, 2004