3D POINT CLOUD MODEL COLORIZATION BY DENSE REGISTRATION OF DIGITAL IMAGES

N. Crombez, G. Caron, E. Mouaddib

University of Picardie Jules Verne
MIS Laboratory
33 rue Saint Leu
80039 Amiens CEDEX 1
{nathan.crombez, guillaume.caron, mouaddib}@u-picardie.fr

Commission V, WG V/4

KEY WORDS: Point clouds, Virtual and Visual Servoing, Images Registration, Colorization

ABSTRACT:

Architectural heritage is a historic and artistic property which has to be protected, preserved, restored and must be shown to the public. Modern tools like 3D laser scanners are more and more used in heritage documentation. Most of the time, the 3D laser scanner is completed by a digital camera which is used to enrich the accurate geometric informations with the scanned objects colors. However, the photometric quality of the acquired point clouds is generally rather low because of several problems presented below.
We propose an accurate method for registering digital images acquired from any viewpoints on point clouds which is a crucial step for a good colorization by colors projection. We express this image-to-geometry registration as a pose estimation problem. The camera pose is computed using the entire images intensities under a photometric visual and virtual servoing (VVS) framework. The camera extrinsic and intrinsic parameters are automatically estimated. Because we estimates the intrinsic parameters we do not need any informations about the camera which took the used digital image. Finally, when the point cloud model and the digital image are correctly registered, we project the 3D model in the digital image frame and assign new colors to the visible points.
The performance of the approach is proven in simulation and real experiments on indoor and outdoor datasets of the cathedral of Amiens, which highlight the success of our method, leading to point clouds with better photometric quality and resolution.

1. INTRODUCTION

Cultural heritage contains important remnants of our past and has to be carefully safeguarded. That is why patrimony documentation represents an important research issue. Cultural heritage includes monuments, works of art and archaeological landscapes. These sites are frequently faced with environmental conditions, human deterioration and can be victim of their age. The E-Cathedrale[1] research program was born in this context of cultural heritage preservation. It is dedicated to the digitization of the cathedral of Amiens, in France, to get a complete and precise model never obtained. This cathedral reaches an internal height of 42.30 meters for 200,000 cubic meters interior volume. Among French cathedrals of which the construction is over, the one of Amiens is the tallest and has the greatest interior volume. The cathedral was built between 1220 and 1270 and has been listed as a UNESCO World Heritage Site since 1981.

Recent technological advances have enabled the development of tools and methods, like 3D laser scanners which allow to quickly and accurately scan the spatial structure of monuments. The scan result is a dense and panoramic 3D point cloud of the scene. The first three scanning campaigns of the cathedral have allowed to obtain billions of points (Fig. 1) with approximately 2 millimeters sampling resolution on the details and less than 5 millimeters on the other parts. A digital camera is included into the laser scanner to assign color informations to each acquired 3D point. However, the photometric quality of the acquired point clouds model is generally rather low because of several problems. First of all, the scanner camera has a low resolution compared to the 3D scanner acquisition itself, because of that a 3D point and its nearest neighbors are colorized with the same digital image pixel. This
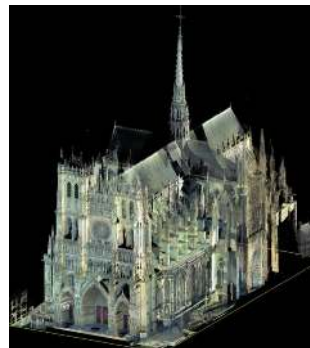
---
[1]https://mis.u-picardie.fr/E-Cathedrale/

Figure 1: View of the Amiens cathedral point clouds model

phenomenon creates a blur effect on the point cloud model which are not representative of the real visual object aspect. Secondly, we do not recover the real-world luminance nuances because of the low photometric dynamic of the scanner camera. Finally, to entirely cover a large scene, the laser scanner has to be placed at different geographical positions. The obtained point clouds are acquired at different time of the day, so at different sun exposures. The merging of these point clouds leads always to visual inconsistencies (Fig. 2). Even if the geometric precision is really accurate the visual aspect of the virtual model is clearly unrealistic. This paper proposes a method to colorize a 3D point cloud model using digital images acquired from any viewpoints.
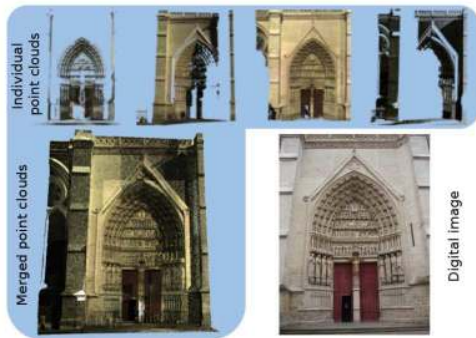
Figure 2: Color problems resulting of the point clouds merging and the scanner camera low quality

2. RELATED WORK

In recent years, the use of 3D technologies has become the major method to accurately measure architectural or historical objects because of their high performance. These measured "objects" can be small as the Sidamara grave (Altuntas et al., 2007) and huge as a desert palace in Jordan (Al-kheder et al., 2009) or as the Todai-ji in Japan, the world biggest wooden edifice (Ikeuchi et al., 2007). In all cases, the aim is to obtain an accurate documentation with a high level of realism for protection, conservation and for possible future restoration or archaeologists studies. Several measurement acquisition methods exist: with terrestrial laser scanners (Haddad, 2011), terrestrial or aerial photogrammetric methods or by close range photogrammetry (Moussa et al., 2012).

It seems interesting to use the laser scanner and the digital camera separately. Indeed, terrestrial laser scanner measurements can be combined with Close Range Photogrammetry method or with digital images to improve the texture quality and resolution with a photometric realism (Al-kheder et al., 2009) (Henry et al., 2012). These methods require the registration of the data sets which are collected by the laser scanner and the digital images acquired by the camera. This registration is a crucial step for a good colorization, this is why it has long been studied and many approaches has been proposed. These approaches have been categorized by M. Corsini and al. (Corsini et al., 2012).

The first category contains the simplest methods, when the two device poses are known relatively to each other. This method is used in (Abmayr et al., 2004) where 3D point clouds and 360 degrees high resolution images are acquired from a same geographical position. Therefore, transforming both coordinate systems into each-other is easily done. However, to cover a large environment the scanner has to be placed far from the scene. The camera is far from the scene too, therefore the acquired digital images have a lack of visual details.

Registration can be also done using corresponding features from both data sets. These corresponding features can be points, lines or else. These can be manually chosen or automatically detected. In (Adan et al., 2012), points are manually chosen. Point clouds are used to generate ortho-images and corresponding points are manually selected in the virtual ortho-images and the digital one. A transformation matrix is computed using these sparse sets of corresponding points to register images and the 3D point cloud. Points can be automatically detected using SIFT, ASIFT or other point features extraction methods. In (Moussa et al., 2012), features are extracted using SIFT in the digital image and in a point-based environment model. In some cases (Alshawabkeh and Haala, 2004), manually or automatically selected, lines may be more appropriate given the straight geometry of the scanned environment. Laser scanners assign to the acquired points a reflectance value.

This reflectance is sometimes used to make the data of the two systems more similar (Mastin et al., 2009) and facilitate the features detection.

Registration can be done comparing 3D model shapes and objects contours detected in the digital image (Belkhouche et al., 2012). Some works are using several images to obtain a sparse 3D point cloud, for example by Structure from Motion (SfM). Then, this point cloud is registered with the 3D target model. This 3D/3D registration performed, the images used for the SfM can be aligned to the model and used to colorize it. However, these methods require a large number of digital images.

Finally, other methods try to minimize mutual informations between digital images and 3D model geometric features. Different features like normals, reflection, ambient occlusions are compared in (Corsini et al., 2009). The results of their study show better image-to-geometry registrations using normals maps and a combination of normals and ambient occlusion.

The huge size of the cathedral leads us to use the automatic features extraction methods. The color problems presented above make difficult the features detection and registration. Thanks to a point clouds colors homogenization, we are able to automatically detect and match point features between a virtual image of the 3D scene and a digital one. A camera pose is computed using these sparse sets of 2D/3D corresponding points to register the digital image and the 3D point cloud model. However, these kind of registrations are not enough accurate, the digital and the virtual images are not correctly aligned. We propose to correct the obtained camera pose using the entire images intensities as features. This dense photometric feature has the advantage of avoiding geometric features detection and matching and has the potential to reach a higher registration precision. The photometric feature has been used for visual tracking of planes (Benhimane and Malis, 2004). In the latter work, image intensity variations are related to the plane motion in the image, expressed as the so called homography, a 2D motion model. In our case, we propose a method to directly use image pixel intensities to estimate the extrinsic and intrinsics camera parameters in a virtual 3D space. For this reason, taking inspiration of (Benhimane and Malis, 2004), we need to relate an image intensity variation to the camera pose variation. This theoretical framework was first tackled for the photometric vision based robotic system control, that is the photometric visual servoing (Collewet and Marchand, 2011). For over twenty years, visual servoing allows to control a robot using various visual features (points, lines, moments etc...). It has been previously extended to Virtual and Visual Servoing (VVS) (Marchand and Chaumette, 2002), aiming to compute a camera pose using image measurements, by non-linear optimization, and exploiting visual servoing knowledge. Our approach extends the photometric visual servoing (Collewet and Marchand, 2011) to the VVS involving the photometric feature.

This paper is organized as follows. First, Section 3.1 compares different similarity measures to determine the most suitable cost function for our data. Section 3.2 describes the photometric virtual and visual servoing. Then Section 4. presents some results of simulations to validate our method. Section 5. shows results of point clouds colorizations using our method. Finally, conclusions and future works are presented in Section 6..

3. PROPOSED APPROACH

The formation of a two-dimensional representation of a three-dimensional world has been extensively studied. A perspective camera transforms a 3D point $\mathbf{X} = (X\,Y\,Z\,1)^T$ in the image plane at homogeneous coordinates $\mathbf{x} = (x\,y\,1)^T$ by the perspective pro-

jection :

$$\mathbf{x} = (x\,y\,1)^T = pr(\mathbf{X}) \quad \text{with} \quad x = f\frac{X}{Z} \quad \text{and} \quad y = f\frac{Y}{Z} \quad (1)$$

where f is the camera focal length. To obtain a visual representation of the image, homogeneous coordinates image points are converted in pixel coordinates. A pixel $\mathbf{u} = (u\,v\,1)^T$ is obtained by $\mathbf{u} = \mathbf{Kx}$ where the matrix $\mathbf{K}$ contains the camera intrinsic parameters :

$$\mathbf{K} = \begin{pmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \quad (2)$$

α_u and α_v are scale factors which relate pixels to distance and u_0 and v_0 represent the principal point. The focal length f is integrated in α_u and α_v, thus the complete 3D point $\mathbf{X}$ transformation to pixel point $\mathbf{u}$ is expressed as :

$$\mathbf{u} = \mathbf{K}pr(\mathbf{X}) \quad (3)$$

A digital image taken in the real world depends on the extrinsic camera parameters (position and orientation in the world) and the intrinsic camera parameters $\mathbf{K}$.

We note $P_{i,j}(\mathbf{X},I)$ the jth 3D point from the point cloud acquired by the station i where $\mathbf{X} = (X\,Y\,Z\,1)^T$ is its coordinates and I its intensity. The point cloud acquired by the station i is noted $\mathbf{PC}_i = \{P_{i,j}(\mathbf{X},I)/\forall j \in [\![0, N_i - 1]\!]\}$ with N_i the number of points acquired by the ith station. The complete model $\mathbf{PCm}$ contains the N registered point clouds: $\mathbf{PCm} = \{\mathbf{PC}_i/\forall i \in [\![0, N-1]\!]\}$.
A virtual image is generated by the projection of 3D points $P_{i,j}(\mathbf{X},I)$ to 2D image points by the complete perspective projection (3). We expressed the formation of a virtual image of a 3D point cloud by $\mathbf{I_v}(\mathbf{PCm}, \mathbf{K}, {}^c\mathbf{M}_o)$ where $\mathbf{K}$ contains the intrinsic virtual camera parameters and ${}^c\mathbf{M}_o$ is the virtual camera pose in the 3D point cloud. The matrix ${}^c\mathbf{M}_o$ represents the transformation from the $\mathbf{PCm}$ (object (o)) frame to a camera (c) frame :

$$^c\mathbf{M}_o = \begin{pmatrix} {}^c\mathbf{R}_{o(3\times3)} & {}^c\mathbf{t}_{o(1\times3)} \\ \mathbf{0}_{(3\times1)} & 1 \end{pmatrix}_{(4\times4)} \quad (4)$$

This matrix ${}^c\mathbf{M}_{o(4\times4)}$ is formed by a rotation matrix ${}^c\mathbf{R}_{o(3\times3)}$ and a translation vector ${}^c\mathbf{t}_{o(1\times3)}$.

To register a digital image $\mathbf{I}$ on a point cloud $\mathbf{PC}_m$, we need the extrinsic parameters of the camera which took $\mathbf{I}$ expressed in the $\mathbf{PC}_m$ frame (the transformation matrix ${}^c\mathbf{M}_o$) and the intrinsic camera parameters $\mathbf{K}$.

3.1 Photometric cost function comparisons

This section aims to determine the most suitable cost function in our study case for a dense visual servoing. We are looking for a robust measure of similarity to handle the color problems previously mentioned (Figure 2). The different similarity measures compared are the Zero-mean Normalized Sum of Squared Differences (ZNSSD) and the Zero-mean Normalized Cross Correlation (ZNCC) between digital images and virtual one in grayscale. We also compare these results with the Mutual Information (MI) between digital images and normal maps and reflexion maps. Normal and reflexion maps have been employed in (Corsini et al., 2009), we use the same methods to create our renderings and compute Mutual Information.
We compute these different measures around a virtual camera pose close to a digital image pose varying two degrees of freedom of the virtual camera. Camera motions are made along its x and y axes. We use the intrinsic camera parameters of the camera

which took the digital image.
This study compares cost function shapes obtained with these correlation criteria (Figure 3), to facilitate the comparison we plot the opposite ZNCC and MI. We can observe that the ZNSSD and the ZNCC cost functions have a more convex shape than the others and a more pronounced minimum. The MI cost functions are more noisy than the others. These observations indicate that it is preferable to use the ZNSSD or ZNCC or the mutual information based on the intensities in order to converge the pose of the virtual camera. Finally, the low complexity of the ZNSSD calculation whether for the criteria minimization/maximization or for the interaction matrix computation, confirms the relevance of the choice of the ZNSSD.
Centering and normalizing the intensities of both virtual and digital images permits to handle the color problems related to the merged point clouds. The error to minimize is computed on these intensities. That is why the cost function shape is so clean, convex and has a clearly distinguishable minimum. Moreover, the ZNSSD is faster to compute than de MI.
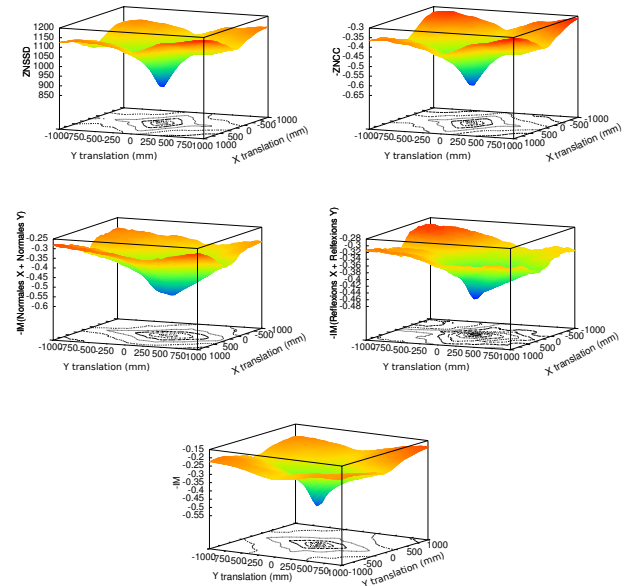

Figure 3: Example of cost functions for five correlation criteria

3.2 Photometric virtual and visual servoing

We express the extrinsic and intrinsic camera parameters estimation problem as an optimization problem under the visual servoing framework. However, unlike the photometric visual servoing (Collewet and Marchand, 2011), our method does not try to minimize the difference between two real images from the same vision sensor. Our method minimizes the difference between a desired digital image $\mathbf{I}^*$ and a virtual image $\mathbf{I_v}(\mathbf{PCm}, \mathbf{K}, {}^c\mathbf{M}_o)$. For this, the criterion to minimize is:

$$\mathbf{e}(\mathbf{r}) = ||\mathbf{I}^* - \mathbf{I_v}(\mathbf{PCm}, \mathbf{K}, {}^c\mathbf{M}_o)|| \quad (5)$$

where $\mathbf{r} = (t_X, t_Y, t_Z, \theta u_X, \theta u_Y, \theta u_Z, \alpha_u, \alpha_v, u_0, v_0)$. The first six elements of $\mathbf{r}$ represent the current virtual camera pose ${}^c\mathbf{M}_o$ by 3 translation and 3 rotation elements (considering the rotation is represented as an angle and a unit vector) and the last four are the intrinsic camera parameters.
The virtual camera motion control law is computed thanks to a

Levenberg-Marquardt based one:

$$\dot{\mathbf{r}} = \lambda \left(\mathbf{H} + \mu \, diag(\mathbf{H}) \right)^{-1} \mathbf{L}_I^{\mathrm{T}} (\mathbf{I}^* - \mathbf{I_v}(\mathbf{PCm}, \mathbf{K_v}, {}^c\mathbf{M}_o)) \quad (6)$$

with $\mathbf{H} = \mathbf{L}_I^{\mathrm{T}} \mathbf{L}_I$, considering $\mathbf{L}_I$ is the interaction matrix related to luminance of image $\mathbf{I}$. This interaction matrix is also named the image Jacobian, i.e. the matrix linking the image feature motion to the camera pose variation. The interaction matrix is derived under the constant illumination consistency (see (Collewet and Marchand, 2011) for details). In the iterative process, the pose ${}^c\mathbf{M}_o$ is updated with the pose increment $\mathbf{r}_6$ composed by the first six elements of $\dot{\mathbf{r}}$ using the exponential map of the special euclidean group SE(3).

$$
\begin{aligned}
{}^{c^{[t+1]}}\mathbf{M}_o &= (e^{(\mathbf{r_6})})^{-1} \times {}^{c^{[t]}}\mathbf{M}_o = \\
&\left({}^{c^{[t]}}\mathbf{M}_{c^{[t+1]}} \right)^{-1} \times {}^{c^{[t]}}\mathbf{M}_o = \\
&\quad {}^{c^{[t+1]}}\mathbf{M}_{c^{[t]}} \times {}^{c^{[t]}}\mathbf{M}_o \quad (7)
\end{aligned}
$$

The intrinsic camera parameters are incremented with the last four elements of $\mathbf{r}$.

The optimization stops when the residual error becomes constant. The optimization is made more robust using an M-Estimator (Comport et al., 2003) which is based on robust statistics (Huber, 1981).

4. SIMULATION

In this Section, simulation results are presented to validate our approach. We use a 3D scene composed by four simple geometric meshed objects as test environment (figure 4).
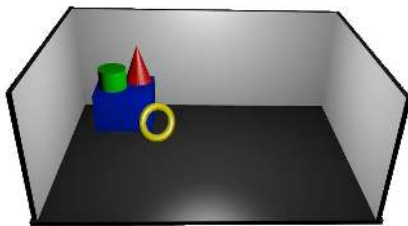
Figure 4: Virtual 3D test environment

A virtual camera is placed inside the 3D environment and a virtual image is generated from this pose. The intrinsic virtual camera parameters are known. We consider this virtual image as the desired image $\mathbf{I}^*$ (figure 5a).

The virtual camera is randomly moved inside a sphere of 0.2 meters around the desired pose. The camera orientations are randomly incremented by an angle between -5 and 5 degrees around its three axes. The intrinsic camera parameters are randomly incremented by a number between -20 and 20 pixels. A virtual image is generated with these new camera extrinsic and intrinsic parameters. This image is considered as the initial image. We use the photometric virtual and visual servoing from this initial pose. Figure 5b shows the image differences between the desired and the initial image. Figure 5c shows the image differences between the desired and the image generated at the end of the servoing. We can see that the virtual camera has converged to the desired camera pose (figure 5d) by minimizing the residual error between the desired image and the images generated during the servoing. The intrinsic camera parameters have also been correctly estimated as shown by the figure 5e.

This experiment has been repeated one hundred times with randomly extrinsic an intrinsic parameters. The virtual camera parameters have successfully converged to the desired parameters
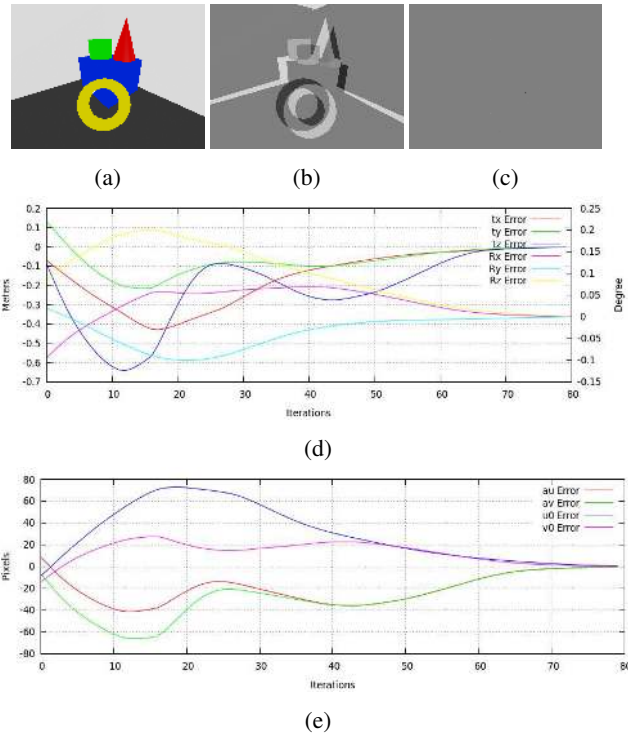
Figure 5: Simulated experiment : (a) desired image, (b) differences between desired and inital image, (c) differences between desired and final image, (d) translations and rotations error, (e) intrinsic camera parameters error

for 95 experiments. The 5 others experiments have failed because the overlap of the geometric objects projection in the desired image and in the initial image was not enough.

5. COLORIZATION RESULTS

5.1 Methodology

The estimated pose allows us to transform each point of the cloud into the digital image frame. The 3D point cloud data are initially expressed in object reference frame. We use the ${}^c\mathbf{M}_o$ from our pose estimation to express the point cloud in the camera reference. Then, using the perspective projection (3) with the estimated intrinsic camera parameters $\mathbf{K}$, we express the points into the 2D digital image plane $\mathbf{I}^*$. Our points have now real coordinates included in the digital image dimensions. We use a bilinear interpolation to assign a new color value to each 3D point. This RGB color value is computed with the weighted average of the RGB values of the four neighboring pixels in the digital image.

In most cases, the colorization or the texturing is done when the point cloud is meshed. In our case, we want to colorize directly the original point cloud. Because of that, we can not expect good colorization results without handling the occlusion issues. Indeed some points coming from parts of the scene which are invisible from the camera position may be considered as visible and colored. This can occur in two situations: when two (or several) points from different depths are on the same projected ray and so are projected into a same pixel or when every points, projected in a pixel, come from invisible parts of the scene. In these two cases, points, which are supposed to be invisible, receive a false new color. To fix these issues, before the colorization we apply the Hidden Point Removal (HPR) (Katz et al., 2007) on the point cloud to remove the 3D points which should be invisible as viewed from the estimated pose ${}^c\mathbf{M}_o$.

5.2 Experimental results

First Experiment : we use our approach to colorize the point cloud of the south portal shown in figure 2 with the digital image shown in figure 6a. The camera is calibrated, so its intrinsic parameters are known.
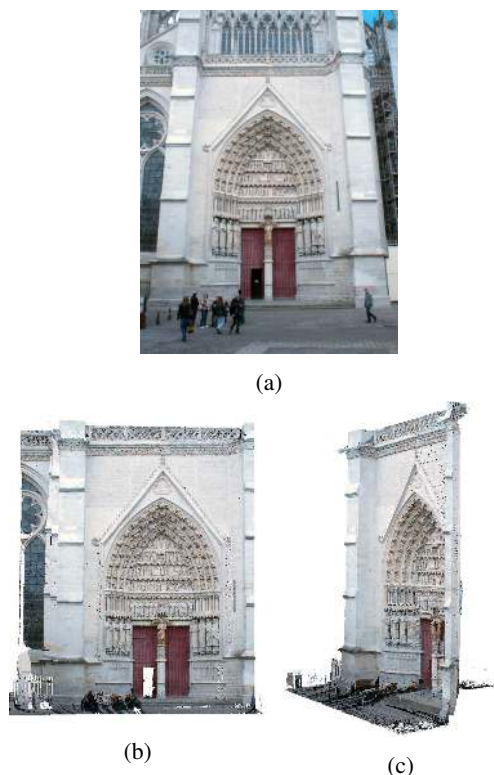

(a)

(b)

(c)

Figure 6: Colorization of the south portal : (a) desired digital image, (b-c) colorized point cloud, (e) evolution of the residual error during the servoing

We use our method to determine from where the digital image has been taken (in the 3D scene which contains the point cloud). The virtual and visual servoing moved the virtual camera by computing iteratively its six degrees of freedom and reached a stable state after 17 iterations. Figure 7 shows the evolution of the residual error during the servoing. Figure 8 represents the difference between the digital image that we want to reach and the virtual one at the first (a) and the last (b) iterations. The distance between the initial virtual camera pose and the optimal one is 674.50mm.
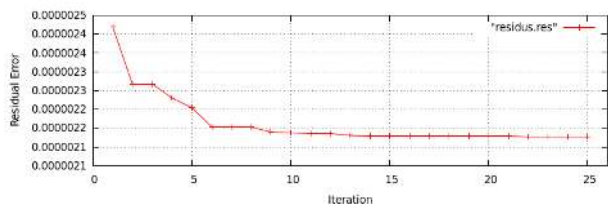

Figure 7: Evolution of the residual error during the servoing

With the estimated camera pose $^c\mathbf{M}_o$, the visible points of $\mathbf{PC_m}$ are projected in the digital image $\mathbf{I}^*$ (figure 6a) and are colorized. Figure 6b and figure 6c show the point cloud $\mathbf{PC_m}$ after its colorization. We obtain a point cloud which seems perfectly colored, with much more uniform colors but if we closely look at the cloud (figure 9), we can see that we need to have several digital images for covering the portal entirely. We can also see the lack of accuracy on the statue face because of a bad digital image point of
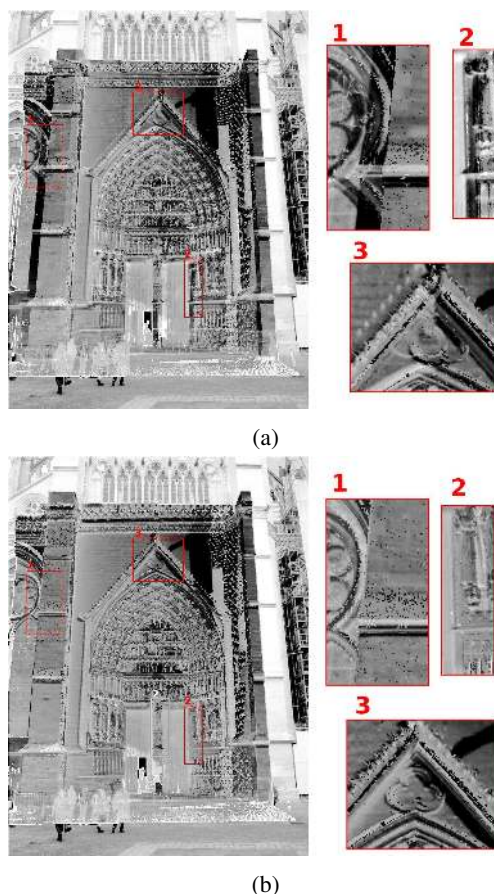

(a)

(b)

Figure 8: Differences images between the desired digital image and (a) the intial image (b) the final image

view choice. To improve the detail visibilities we need a better image resolution for the colorization. Therefore, we have to take images from nearest positions from the statue.
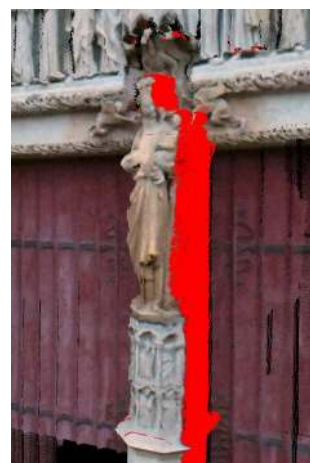

Figure 9: The south portal point cloud colorized. The points recolorized in red are invisible from the digital image pose

Second experiment : we use our pose optimization method to entirely colorize a smaller part of the south portal : the Virgin Mary statue. Our goal is to obtain an accurately colorized point cloud with photo-realistic quality. To obtain this, ideally, the result would be to have each 3D point visible from the camera pose

projected in one different pixel. Knowing the intrinsic camera parameters and the point cloud resolution, we computed the distance needed between the object and the camera to ensure that a pixel receives a unique projected 3D point. Empirically, we calculated a four meters distance needed between the statue and our camera to project each visible 3D point in one different pixel in the 2D image. Three digital images have been taken I_1^*, I_2^*, I_3^* around the statue for a maximum covering (figure 10). The three
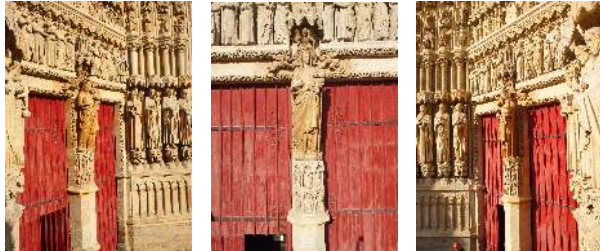
Figure 10: Images of the Virgin Mary used to the colorization (These digital images have a 4032x3024 pixels resolution)

camera poses are estimated with the pose estimation presented before and we use our complete colorization process. Right now, the points colorization is done one image after the other. Points are projected in the first digital image, points covered by this one are colored and the others are labelled as not yet colored. Then, points are projected in a second digital image and only the covered and labelled points are colorized. This operation is repeated with all the digital images or when all the points are colorized. A comparison between the Virgin Mary point cloud before and after our colorization is shown in Figure 11. We took another
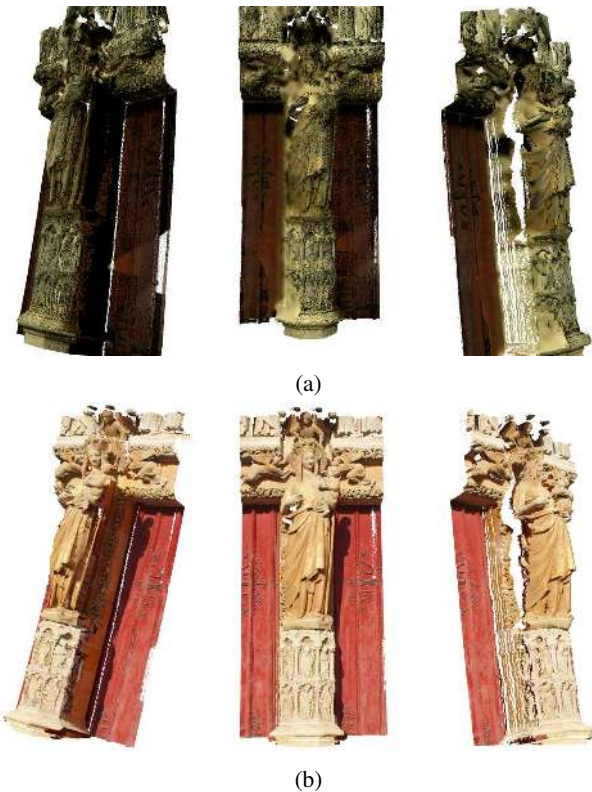
Figure 11: The Virgin Mary point cloud (a) before and (b) after our colorization.

digital image I_4 of the Virgin Mary from a different view angle for a visual validation. This image has not been used for the point

cloud colorization. Figure 12 shows the Virgin Mary point cloud from the same camera pose than the new image after and before our colorization process. As we can see, the good colorization results are clearly undeniable. The two images focused on the Virgin's face (Figure 12) show us the really accurate color resolution of the point cloud after its colorization. The digital image I_4 has been taken 5 months (different season) after we took the three images I_1^*, I_2^*, I_3^* used for the colorization. This explains the colorimetry slight differences between I_4 and the colored point cloud.
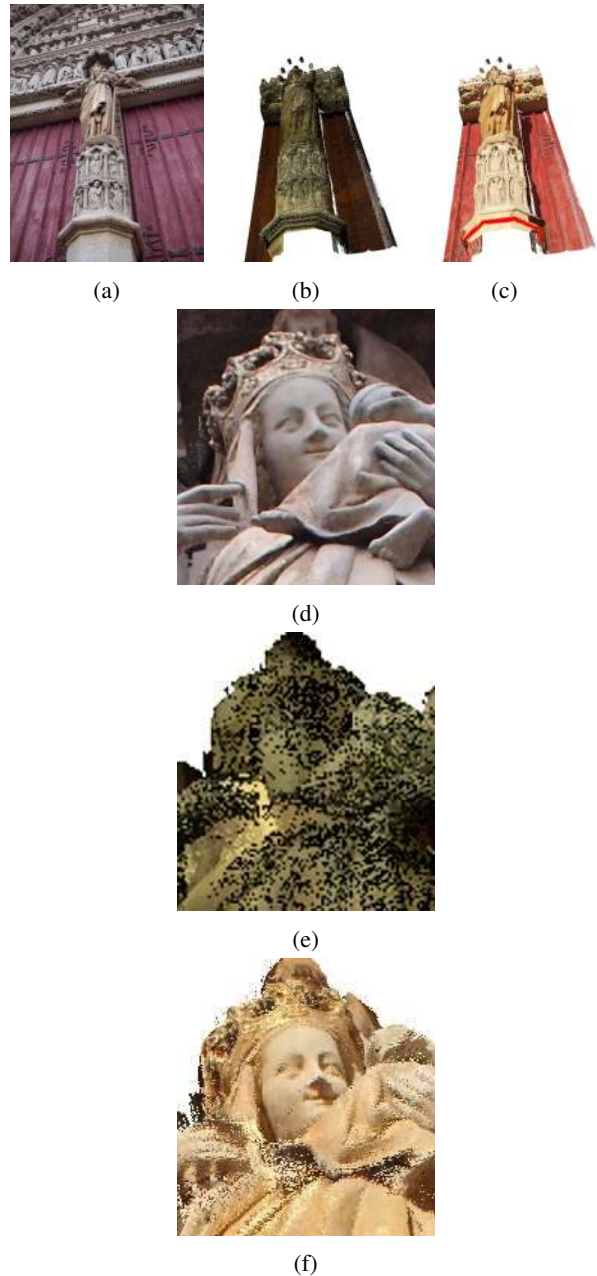
Figure 12: (a) A Virgin Mary statue image from a low-angle shot. (b) The original point cloud of the Virgin Mary from the same camera pose. (c) The colored point cloud of the Virgin Mary from the same camera pose. (d-f)A focus on the color resolution of the Virgin's face.

Third Experiment : we also used the colorization process on a point cloud of the painting in the chapel of Saint-Sebastian.

Figure 13 shows a comparison between the point cloud colors acquired by the scanner (bottom area) and after its colorization (top area).

Figure 13: Painting in the chapel of the Saint-Sebastian before (bottom) and after (top) its colorization.

Fourth Experiment : thanks to the intrinsic camera parameters estimation, we can also use digital images without any information about the camera which took them. We colorize a part of the point cloud of the western faade portal using a digital image from the internet (figure 14a).

6. CONCLUSIONS AND FUTURE WORKS

We proposed an alternative method for automatically registering digital images over 3D point cloud. The existing manual methods used to detect the correspondence features of the two data sets can provide good results. However, these methods can be really tedious to handle an important size of data. Automatic detections of geometric features can be delicate because they require to have similarities between the digital images and the virtual images generated with the 3D point clouds.

In our approach, rather than using geometric image features, we use the entire image intensities which make the registration more robust and precise than by manually selecting points. In the future, it seems interesting to use the RGB information rather than only the intensities.

The optimization of the intrinsic camera parameters allows to use digital images without any informations about the camera which took them. The presented results show that our process provides point clouds with a realistic texture quality. Future works will be focused on the research of the optimal positions needed to obtain digital images to colorize a point cloud.
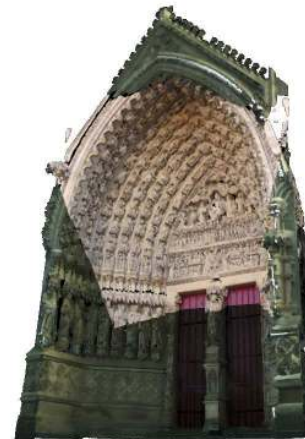

(a)

(b)

Figure 14: (a) Digital image of a part the western façade portal from the internet. (b) Point cloud of the western faade portal with points colored using the digital image

REFERENCES

Abmayr, T., Hartl, F., Mettenleiter, M., Heinz, A., Neumann, B. and Fr02hlich, C., 2004. Realistic 3d reconstruction - combining laserscan data with rgb color information.

Adan, A., Merchan, P. and Salamanca, S., 2012. Creating Realistic 3D Models From Scanners by Decoupling Geometry and Texture. International Conference on Pattern Recognition (ICPR) pp. 457–460.

Al-kheder, S., Al-shawabkeh, Y. and Haala, N., 2009. Developing a documentation system for desert palaces in Jordan using 3D laser scanning and digital photogrammetry. Journal of Archaeological Science 36(2), pp. 537–546.

Alshawabkeh, Y. and Haala, N., 2004. Integration of digital photogrammetry and laser scanning for heritage documentation. IAPRS pp. 12–23.

Altuntas, C., Yildiz, F., Karabork, H., Yakar, M. and Karasaka, L., 2007. Surveying and documentation of detailed historical heritage by laser scanning. XXI International CIPA Symposium pp. 01–06.

Belkhouche, Y., Buckles, B., Duraisamy, P. and Namuduri, K., 2012. Registration of 3d-lidar data with visual imagery using shape matching. Int. Conf. on Image Processing, Computer Vision, and Pattern Recognition pp. 749–754.

Benhimane, S. and Malis, E., 2004. Real-time image-based tracking of planes using efficient second-order minimization. IEEE/RSJ International Conference on Intelligent Robots and Systems pp. 943–948.

Collewet, C. and Marchand, E., 2011. Photometric Visual Servoing. IEEE Transactions on Robotics 27(4), pp. 828–834.

Comport, A., Marchand, E. and Chaumette, F., 2003. A real-time tracker for markerless augmented reality. In: ACM/IEEE Int. Symp. on Mixed and Augmented Reality, ISMAR, Tokyo, Japon, pp. 36–45.

Corsini, M., Dellepiane, M., Ganovelli, F., Gherardi, R., Fusiello, a. and Scopigno, R., 2012. Fully Automatic Registration of Image

Sets on Approximate Geometry. International Journal of Computer Vision 102(1-3), pp. 91–111.

Corsini, M., Dellepiane, M., Ponchio, F. and Scopigno, R., 2009. Image-to-geometry registration: a mutual information method exploiting illumination-related geometric properties. Computer Graphics Forum 28(7), pp. 1755–1764.

Haddad, N. A., 2011. From ground surveying to 3D laser scanner: A review of techniques used for spatial documentation of historic sites. Journal of King Saud University - Engineering Sciences 23(2), pp. 109–118.

Henry, P., Krainin, M., Herbst, E. and Ren, X., 2012. Rgb-d mapping : Using kinect-style depth cameras for dense 3d modeling of indoor environments. The International Journal of Robotics Research 31(5), pp. 647–663.

Huber, P. J., 1981. Robust Statistics. John Wiley and Sons.

Ikeuchi, K., Oishi, T. and Takamatsu, 2007. The great buddha project: Digitally archiving, restoring, and analyzing cultural heritage objects. Int. J. Comput. Vision 75(1), pp. 189–208.

Katz, S., Tal, A. and Basri, R., 2007. Direct visibility of point sets. ACM Trans. Graph. 26(3), pp. 24.

Marchand, E. and Chaumette, F., 2002. Virtual visual servoing: A framework for real-time augmented reality. In: Eurographics Conference Proceeding, Vol. 21(3), pp. 289–298.

Mastin, A., Kepner, J. and Fisher, J., 2009. Automatic registration of lidar and optical images of urban scenes. IEEE Conference on Computer Vision and Pattern Recognition pp. 2639–2646.

Moussa, W., Abdel-Wahab, M. and Fritsch, D., 2012. An automatic procedure for combining digital images and laser scanner data. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences pp. 229–234.