RECONSTRUCTION OF URBAN SITES FROM PHOTOS

F. Agnello

University of Palermo, Department of Architecture, Palermo, Italy, fabrizio.agnello@unipa.it

Commission II

KEY WORDS: Perspective restitution, Single image 3D restitution, Camera pose estimation, Virtual reconstruction, Panoramic images

ABSTRACT:

This paper reports the results of an experiment that aims at the virtual reconstruction of a urban site that has been partially reshaped during the 20th century; the reconstruction process is based on period photos. The chosen case study is the eldest harbour of Palermo, named ‘Cala’; the site was bombed during World War II and new buildings took the place of the ruined ones. Two period photos, taken from an aircraft, document the buildings that were destroyed by bombs.
The 3D restitution process used the ‘inverse’ projection from period photos (2D) to the virtual space (3D) ; the first step was therefore addressed to the calculation of the inner parameters (focal length, principal point coordinates) and of the cameras’ pose, i.e. their position and orientation. The data needed for these calculations were extracted from the 3D laser scanning survey of those buildings that appear in photos and that are still on site. The calculation of inner parameters and poses has been computed with a motion tracking commercial package.
The second part of the paper focuses a process for the fruition of virtual reconstructions, based on the alignment of real panoramic images, generated by photos taken on site, and of virtual panoramic images, extracted from the 3D reconstruction model; the proposed method uses SfM photogrammetric tools.

1. INTRODUCTION

All along the 20th century the biggest towns in Europe were deeply transformed, by the reconstruction plans that followed war damages, or simply by urban renewal plans (larger roads, more services, etc.).
War destructions often ruined monuments and relevant urban sites that deserve today a posthumous documentation and a chance of new life by means of virtual or augmented reality.
A huge amount of period photos, stored in private or public repositories, display the destroyed monuments and urban sites (squares, roads) as they appeared in the second half of the XIX century and in the first half of the XX century.
The paper reports the results of an experiment aiming at the setup of a process for the 3D virtual reconstruction of no extant sites and buildings that are documented in period photos.
The 3D reconstruction from photos is not simple nor accurate; period photos come to us printed or published in books. The original film and the features of camera and lenses are very rarely available.
When images are extracted from books, the size and the ratio (e.g. 4:3, 2:3) of the original image are unknown, since images were often cut to fit the layout of the page.
The monocular 3D restitution demands the availability of geometric or dimensional data referred to the buildings that appear in the photo.
Perspective restitution is the traditional approach to 3D monocular restitution; if some geometric features are known (e.g. the angles formed by couples of intersecting straight lines), the perspective restitution leads to the detection of the inner orientation parameters of the image and to the restitution of the size of segments in an unknown unit. If a length is known, than the reconstructed models can be properly scaled and measured.
Perspective restitution has been a neglected subject for decades; the lack of accuracy, due to the limits of traditional drawing tools, made it the subject of collateral lessons in the courses of descriptive geometry.
Digital tools for drawing offer a new chance to perspective restitution, since they allow the detection of the furthest vanishing points and the real time evaluation of the correspondence between image and 3D reconstruction (Agnello 2018).
Nonetheless, the most impressive improvements to the 3D restitution from photos do not come from descriptive geometry, but from the researches addressed to the development of automated solutions for the computation of the position and orientation of a monocular camera in a given space.
The problem, usually known as camera pose estimation, has a peculiar relevance in robot navigation and in augmented reality.
The solution assumes that 3D data (Point clouds, textures, reflectance) of the environment that is displayed in the image are available. Camera pose estimation can be solved by matching textures and intensities, or geometric features. Scholars consider the latest approach more robust, because detection is not conditioned by illumination, shadows, etc. (Li-Chee Ming et al. 2017).
The geometric-based camera pose estimation demands 3D data of the environment that is displayed in the image, e.g. the coordinates of some relevant points (corners and the like); that is why the process is usually referred to as model-based tracking.
If the inner orientation is known, one of the simplest solutions of the model-based tracking is provided by the Perspective-3-Point, usually reported as P3P. The pose of a camera is detected when the positions of 3 points displayed in the photo are known.
The P3P system provides four possible solutions, which can be retrieved both by analysis (Gao, 2003) or graphic representation (Fallavolita et al., 2017).
In the graphic approach, the theoretic solutions are easily discarded, since they suppose unreal locations of the camera; in the analytical approach, the best solution is detected if more than three points are considered; the solution becomes therefore P*n*P.
The pre-requisite of model-based tracking is a strict correspondence between the images and the 3D model of the site.

Nonetheless, the P*n*P approach leads to a good camera pose estimation even when the depicted scene matches *only partially* the 3D model of the extant site.
This property becomes relevant when the camera pose estimation process is addressed to period photos that show both extant and disappeared buildings. Survey or dimensional data of the buildings that survived the transformations is therefore enough to support the pose estimation, especially if the survived buildings, and therefore the surveyed points, are widely spread in the image.

historic photos and the 3D reconstruction of the buildings that no longer exist.

In this paper a commercial motion-tracking package supported the calculation of both the inner orientation and the pose of period photos. The laser scanning survey of those buildings that survived the transformations and the destructions of the site, provided the coordinates of the points needed to solve the camera pose.
The second part of the paper discuss a method for the mutual orientation of real panoramic images, generated by photos taken on site, and of virtual panoramic images, extracted from the 3D reconstruction model; the proposed method uses SfM photogrammetric tools. The matching of real and virtual panoramic images allows transitions between the two and thus supports the comprehension of the links between the present and the past features of the site.

2. RELATED WORKS

Fallavolita et al discuss the graphic solution for monocular camera pose estimation by means of known features, such as the position of points (Fallavolita et al., 2013). The authors recall that the P3P solution is nothing but the problem of the ‘vertex of the pyramid’ (inverse resection) quoted by Gaspard Monge in his writings. The authors propose a solution apart, based on perspective restitution, for the estimation of the inner orientation of the studied photo.
Algorithms and tools for the automatic detection of vanishing points and for the semi-automatic restitution of geometric features have been widely discussed (Aguilera, 2004 and Grammatikopoulos, 2004).
The estimation of the initial pose of the camera without knowledge of the previous camera pose, a critical issue in camera tracking, is solved by monocular camera pose estimation (Lahdenoja, 2015).
Finally, the alignment of real and virtual panoramic images is discussed in Li-Chee-Ming et al. The authors provide a detailed review of the researches on monocular camera pose estimation and propose a geometric-based pose detection that uses panoramic images (Li-Chee-Ming et al., 2017).

3. DEVELOPED METHODOLOGY

3.1 The case study and the reference data

The paper focuses the reconstruction of the ‘Cala’, the old harbour of Palermo. The ‘Cala’ and the surrounding area were hardly bombed during World War II and modern buildings took the place of ruined houses, palaces and monuments.
Two photos, taken in 1924 from an aircraft, show that, at that time, the buildings facing the ‘Cala’ formed an uninterrupted curtain; many of those buildings were destroyed by bombs, but many other survive in their original location (fig. 1).
Another historic datum that supported the 3D reconstruction process is a 1:5000 map of the area, dated 1935; the map does not report elevation data, nor the limits of the buildings.
The lines on the map represent the boundary of the blocks, delimited by roads or open spaces.

Figure 1: The southern front of the ‘Cala’ in 1924 and today.

The site has been surveyed with a laser scanner, in order to provide the coordinates of points (mainly corners on the roofs and on the vertical edges of buildings) to be used in the camera pose estimation process.
The point cloud and the map have been loaded in a digital drawing package and have been referred to a sole coordinate system (fig. 2).

Figure 2: Laser scanning survey and the 1935’ map.

3.2 Camera pose estimation and 3D reconstruction

The pose estimation of historic photos has been solved with the motion tracking commercial software Syntheyes. The software provides a simple workflow to upload points’ coordinates, position the points on the image and finally estimation the inner orientation and the pose of the photo.

The coordinates of the point of view, the focal length of the camera and the Euler angles that allow the reconstruction of the direction of the vector perpendicular to the image plane, can be exported in various formats (fig. 3).

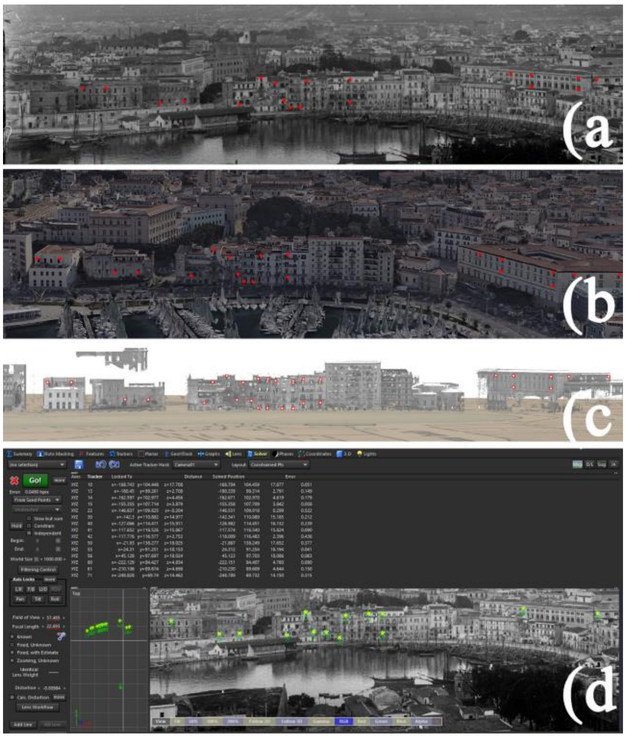

Figure 3: Camera pose estimation. The points used for camera pose estimation in period photo (a), in a image from google earth (b), in the point cloud (c) and in Syntheyes (d).

The availability of two historic images taken in a sequence (same photographer, same camera), allowed evaluating by comparison the results of inner orientation. Calculated errors range from few centimetres up to 30cm.
Both period images are in 2/3 proportion and they could reasonably be the prints of a 6*9 film. If we scale the size of the photo according to this hypothesis, the focal length result 111.25mm in one photo and 107.02mm in the second one; both values fit the focal length of lenses used at the time and the residual between the values can be accepted.

Figure 4: Cameras poses in the 3D scene.

After the images were positioned in the 3D scene, together with the 1935 map and the point cloud (fig. 4), the projective correspondence between the photos and the point cloud has been validated by visual inspection, setting a virtual camera on the point of view and assigning the plane of the image an increasing transparency value (fig. 5).

Figure 5: Visual validation of external orientation.

The reconstruction process proceeded by steps, literally performing an inverse projection (fig. 6):
a) in the first step a vertical plane from the point of view is traced; in the photo, a corner belonging to a vertical edge of a building is marked by a point; the vertical line and the marked point detect a plane that intersects the 1935 map;
b) if the intersection line cuts the lines of the map in the expected point, a new vertical line is drawn from such point;
c) the ray through the point of view finally detects the 3D position of the point marked in the image.

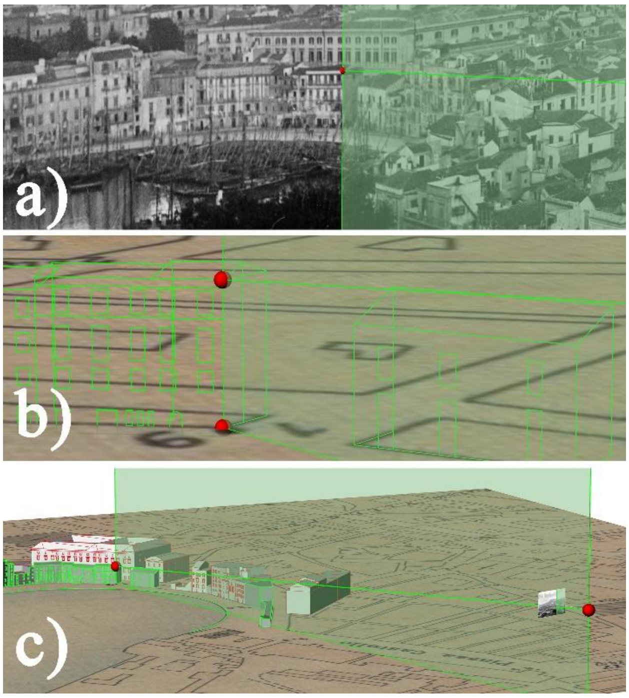

Figure 6: The Reconstruction process.

The photos used for the reconstruction have a slight overlapping; the edges and lines displayed in both images supported the mutual assessment of the reconstruction process and allowed the estimation of the accuracy, that has been evaluated in decimetres. The projective correspondence between photos and 3D model and the accuracy of digital drawing allowed the restitution of doors and windows in the buildings' fronts (fig. 7).

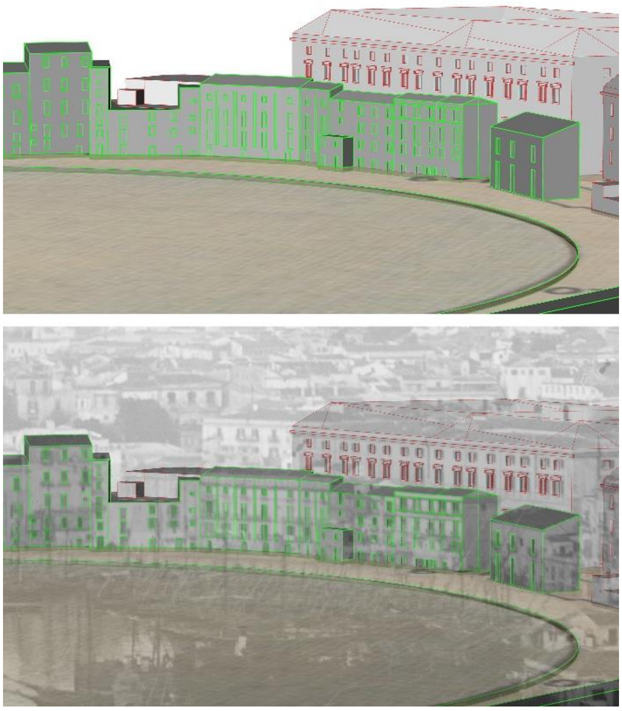

Figure 7: The reconstruction process.

At the end of the reconstruction process, a complete 3D model of the southern front of the 'Cala' has been built (fig. 8). In the 3D model buildings that survived the destruction are red coloured, whereas the colour green features the buildings that are definitely lost.

Figure 8: The 3D reconstruction model of the 'Cala'.

The last step of the reconstruction process aimed at the photogrammetric texturing of the 3D model; both the reconstructed 3D model and the oriented historic cameras have been loaded in the SfM photogrammetric package Photoscan. The black and white pixels of historic images were used to calculate the texture of the 3D model (fig. 9).

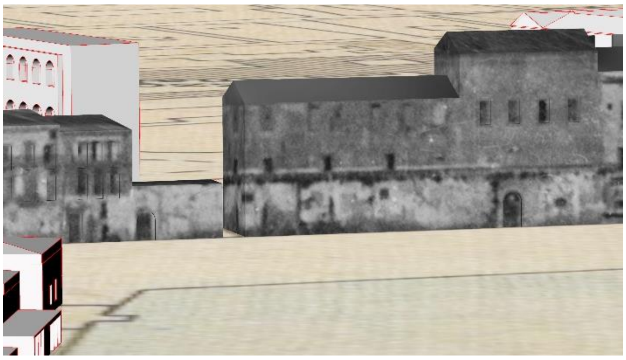

Figure 9: The textured reconstruction model.

3.3 Aligning Real and Virtual panoramic images

In order to set up an effective support for the fruition of the virtual reconstruction of the 'Cala', a workflow for the alignment of real and virtual panoramic images, bases on SfM photogrammetric tools, has been tested.

Aligned panoramic images add to the usual panoramic exploration the chance of a gradual transition between the extant site and its virtual reconstruction, thus making easier the understanding of where and how the modifications changed the site.

The transitions between 'real' panoramic images of the site and 'virtual' panoramic images of a 3D model requires the knowledge of the coordinates of the centre of the sphere generated after fish-eye images' stitching; if such coordinates are known, then the positioning of the virtual camera in the virtual 3D space becomes easy.

The alignment of panoramic images becomes complete when the mutual orientation is fixed, that is when the barycentre points of equirectangular images, extracted from the panoramic images, refer to the same direction in 3D space.

In order to setup a reference system for the orientation of fish-eye images, a photographic documentation of the site has been taken with a reflex camera equipped with a 35mm lens.

The photogrammetric model has been scaled and referred to the coordinate system of point clouds (and of the 3D reconstructed model) via the creation of markers and the extraction of their coordinates from the point clouds. This way the external orientation of photos was calculated (fig. 10).

In a second step, further the photos were taken with a camera mounting a fish-eye lens; in order to use these photos for the creation of panoramic images, the camera has been mounted on a Nodal Ninja mechanical arm, fixed to a photographic tripod.

From each stationing point seven photos have been taken; six photos at 30° intervals with the axis of the camera approximately horizontal and the camera rotated in portrait mode; the seventh photo with the camera slightly oriented upwards (fig. 11).

Figure 10: Marker coordinates from Recap and Marker creation in Photoscan.

Figure 11: Fish-eye photos taken from a station point.

In order to calculate the coordinates of the point of view of each 7-photos fish-eye images set, these photos have been added to the SfM photogrammetric model and a new alignment process has been launched.

It is well known that in SfM photogrammetry photos should not be taken from the same point; nonetheless, Photoscan allows grouping photos into a specific folder that can be classified as ‘station’ folder. Thanks to this classification the software assumes that photos have been shot ‘exactly’ from the same point and processes their orientation (fig. 12).

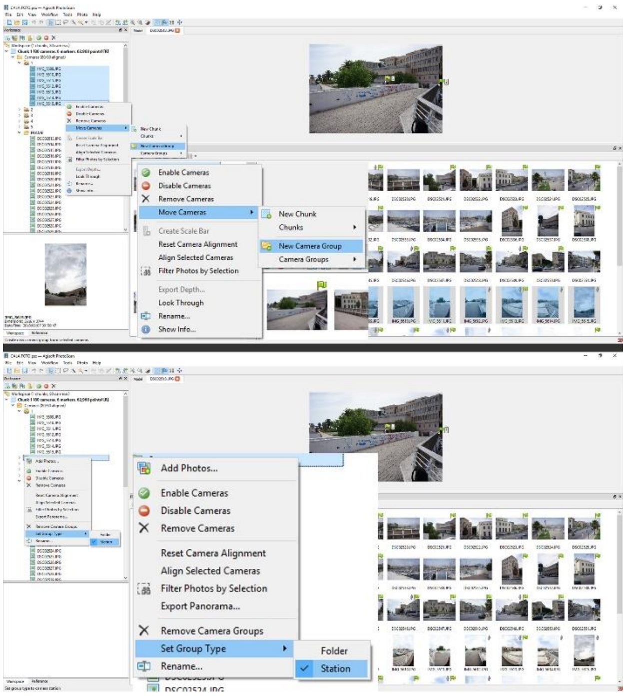

Figure 12: Uploading Fish-eye photos in Photoscan.

The tested workflow succeeded and the photogrammetric model provided the coordinates of the nodal point of each 7-photos group (fig. 13) and the Euler angles reporting the orientation of the normal to the plane of each image (fig. 14).

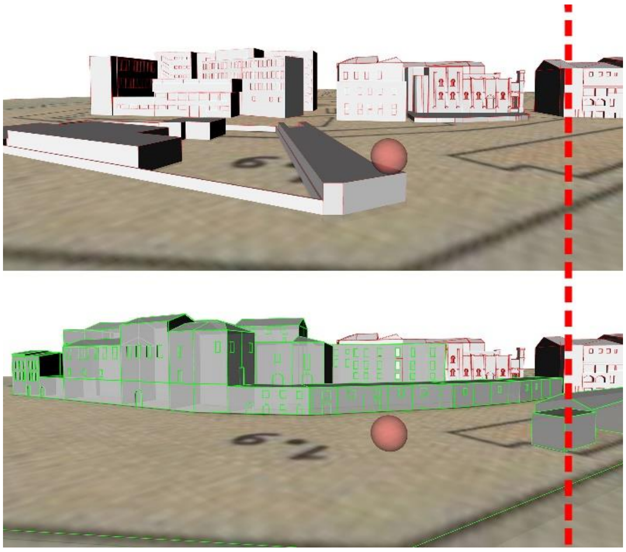

Figure 13: Position of the point of view of a 7-fish-eye photos group in the real site (above) and in the reconstructed model.

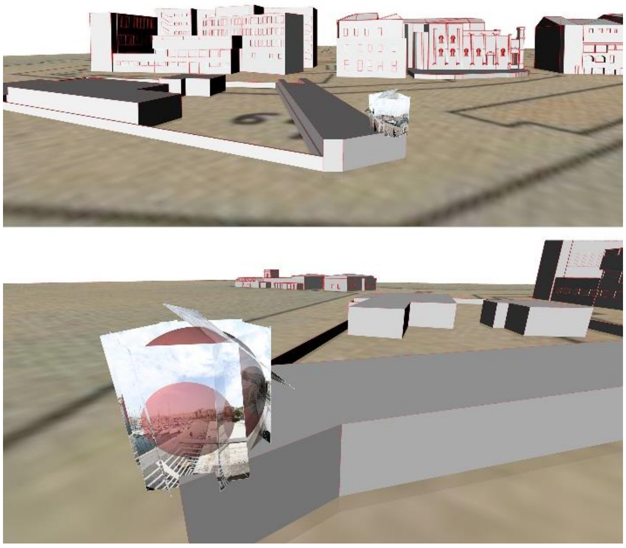

Figure 14: Orientation of a 7-fish-eye photos group.

It is worth remarking that the combination of fish-eye and undistorted photographs is particularly effective when panoramic photos are taken in a wide site, since the detection of points for the orientation of the photogrammetric model becomes difficult, with fish-eye images, if the points are far from the camera.
The textured reconstruction model has been exported in the *.fbx format file and has been uploaded in Blender. Here, a virtual panoramic camera has been positioned on a point whose spatial position matches the calculated position of the point of view of the chosen 7-fish-eye photos group.
The equirectangular image generated by the virtual panoramic camera was oriented so that its starting point is aligned to the vertical edge of a window that has maintained its position and size, even if the façade of the building has been hardly transformed (fig. 15).

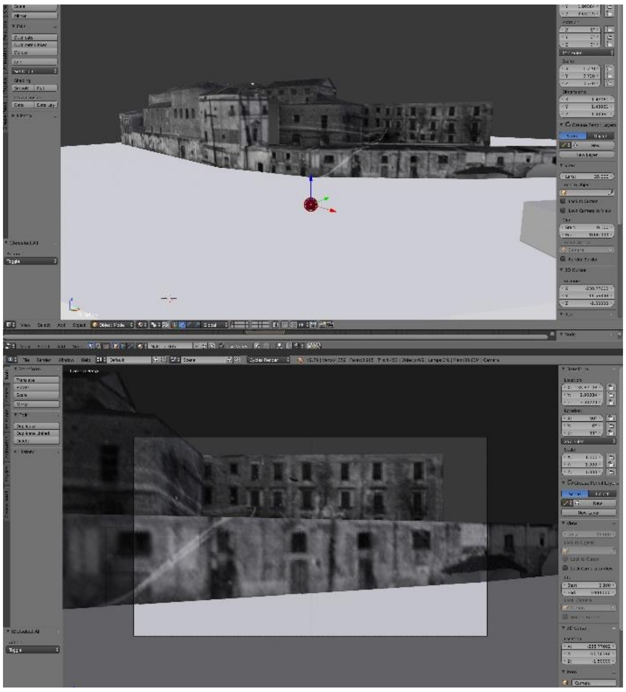

Figure 15: Orientation of the virtual panoramic image.

In order to get an effective alignment of the two panoramic images, the last step was addressed to match the orientation of the real and of the virtual panoramic images.
It is known that the starting direction of a panoramic image becomes the vertical axe of symmetry of the corresponding equirectangular image. The final step was therefore addressed to adjust the starting direction of the real panoramic image to match the vertical axe of symmetry of the virtual equirectangular image. Equirectangular images of each 7-fish-eye photos group have been created with the stitching software PtGUI. When the image is created, the software allows shifting its axe of symmetry. The new barycentre has been aligned to the chosen vertical edge of the window (fig. 16).
In the last step of the visualization experiment, the virtual image has been overlaid to the real one (fig. 17). The resulting ‘mixed’ image is particularly effective when transitions are executed, since it makes the reconstructed model appear inside the real panoramic image.
At the end of the process real and virtual, mixed, historic panoramic image result almost exactly aligned (fig. 18). A slide bar in the panoramic viewer allow a gradual and effective transition between present and past (fig. 19).

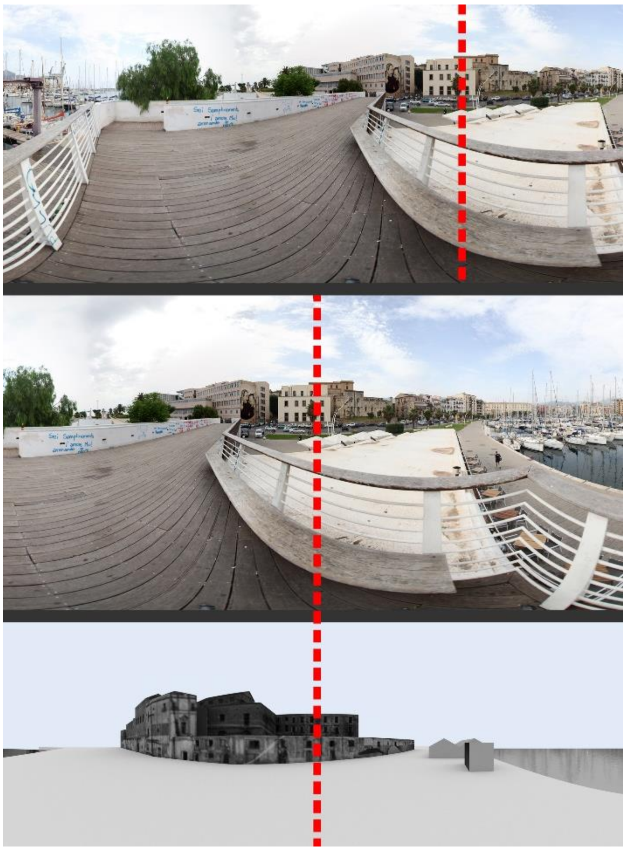

Figure 16: The orientation of the real panoramic image is shifted to match the orientation of the virtual panoramic image.

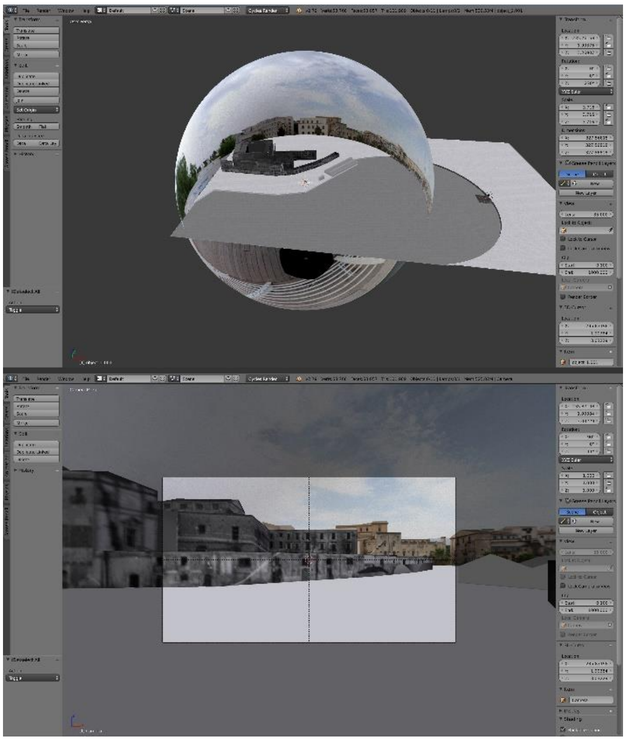

Figure 17: Real and virtual image are mixed in Blender.

Figure 18: Alignment of Real and mixed panoramic images.

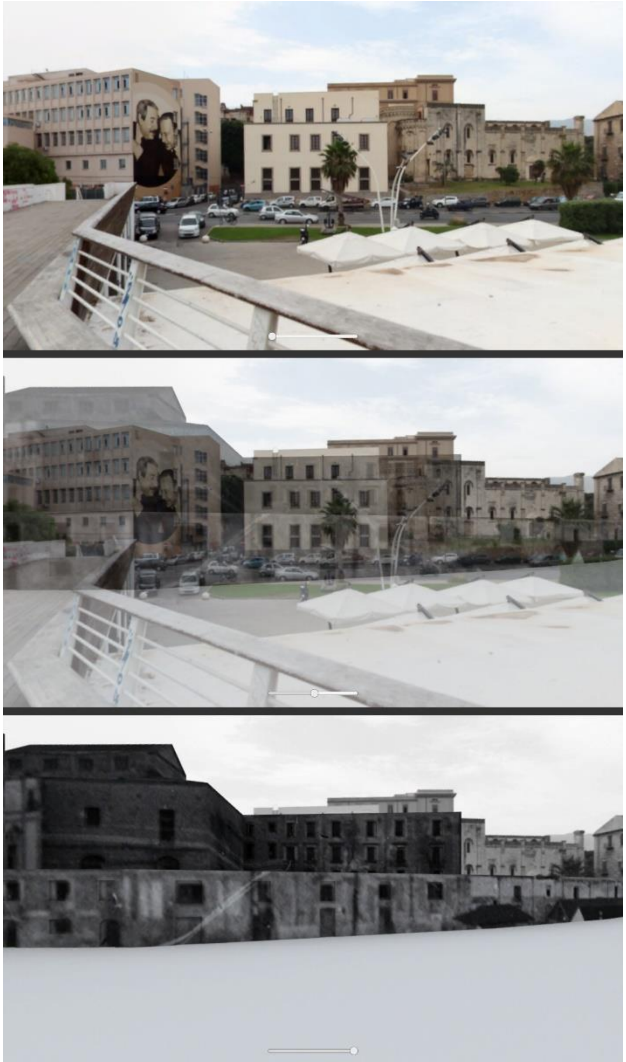

Figure 19: Gradual transitions from real to virtual image

4. CONCLUSIONS

The paper presented a workflow for the reconstruction of a urban site using historic photos and for the visualization of the virtual reconstruction with the transition between real and virtual panoramic images. The proposed method is effective only if some features displayed in the historic photos (edges, points) are still in their original location. The spatial position of these features, detected by a survey, supported both the detection of the pose of period photos and the alignment of real and virtual images.
The development of motion tracking solutions has offered new opportunities to virtual reconstruction from a single image, allowing an almost accurate estimation of camera's inner orientation and pose.
Future works could address the application of the proposed workflow to different sites (e.g. archaeological sites), or the use of augmented reality solutions for the fruition of the virtual reconstructions.

5. ACKNOWLEDGEMENTS

The work on 3D reconstruction from period photos started as a degree thesis project in Architecture; the author wishes to thank the student Giulia Agnello for having developed an interesting research on photography and photographers working in Palermo in the first half of the 20th century and for having modelled the extant buildings of the 'Cala'.
The research on historic photos has been encouraged and supported by Arch. Fabio Militello, official at the 'Centro Regionale per l'Inventario e la Catalogazione e la Documentazione' (CRICD) of the Sicilian Regional Administration.

REFERENCES

Agnello, F., 2018. La ricostruzione da immagini con tecniche di restituzione prospettica. *Proceedings of the 40th UID Conference*, pp. 247-256.

Aguilera, D., Claro-Irisarri, M.A., Gomez Lahoz, J., Finat, J., Gonzalo-Tasis, M., 2004. Development of a simulator for reliable and accurate 3D Reconstruction from a single view. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences,* Vol. 35, pp. 142-147.

Fallavollita, F., Migliari, R., Salvatore, M., 2013. Monge e il problema del vertice di piramide: una applicazione alla restituzione di quote e volumi da una fotografia del 1892, in *DisegnareCon*, Vol. 6, n. 12 (2013), pp. IX/1-9.

Gao, X., Rong Hou, X., Tang, J., Cheng, H., 2003. Complete solution classification for the perspective-three-point problem. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Volume: 25, Issue 8, pp. 930-943.

Grammatikopoulos, L., Karras, G., Petsa, E., 2004. Camera Calibration Combining Images with Two Vanishing Points. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, Vol. 35, No. 5, pp. 99-104.

Lahdenoja, O., Suominen, R., Säntti, T., Lehtonen, T., 2015. Recent advances in monocular model-based tracking. *University of Turku Technical Reports*, n. 8.

Li-Chee-Ming, J. and Armenakis, C., 2017. Matching real and synthetic panoramic images using a variant of geometric hashing, *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Volume IV-1/W1, 2017, pp. 199-206.