INSERTING THE PAST IN VIDEO SEQUENCES

Elli Petsa, Stefanos Kouroupis
Department of Surveying, Technological Educational Institute of Athens
GR-12210 Athens, Greece (e-mail: petsa@teiath.gr)

George Karras
Department of Surveying, National Technical University of Athens
GR-15780 Athens, Greece (e-mail: karras@central.ntua.gr)

KEY WORDS: single image techniques, 3D reconstruction, panoramic image, video sequence

ABSTRACT

There is currently a growing demand for photo-textured renderings in fields such as virtual tourism, architectural walk-throughs or archaeological reconstruction. Although this usually takes the form of "artificial" VRML views of modeled 3D scenes draped with image content, video sequences can also be augmented with the insertion of non-existing structures. While these are mostly structures planned to be built, this contribution addresses the question of demolished buildings. In a video sequence showing part of a central square of Athens one building, torn down long ago and replaced by a high hotel, is inserted. Its 3D model has been produced from an old single photograph using its vanishing points. The video itself has been produced in a simple way. First, a panoramic image has been generated by matching cocentric overlapping images. A photo-textured perspective view of the reconstructed building was then inserted into the panoramic image, parts of which were subsequently resampled to produce a frame sequence. This video sequence, combining real scene views with a virtually inserted building, imitates the movement of a stationary rotating camera. Our next step will be to produce a full sequence incorporating further demolished buildings of the square, for which old photographs are available.

1 INTRODUCTION

Perhaps one of the most challenging tasks of 3D modeling and visualisation techniques is "reconstruction" of the past. Indeed, with the exception of the (usually highly controversial) instances where an actual rebuilding is proposed, this is apparently the only means available for visually "regenerating" past scenes. Such approaches may have many important uses, e.g. as tools for historical and architectural research or for education purposes and visual representations for the public. The present contribution focuses in particular on the visualisation of demolished buildings. Such objects may be approached only though historic records (pictorial or written) and, in many cases, reconstruction has to rely on single images. The use of such single-image techniques for the reconstruction of, mainly polyhedral, structures is well documented (CIPA Task Group 2, 2001) and, indeed, alternative techniques have been reported at the last CIPA Symposia on this subject (Karras and Petsa, 1999; Bräuer-Burchardt, 2001; van den Heuvel, 2001).

Generally, however, a reconstructed building needs to be depicted within its environment, be it an artificially generated or even a real scene. The first alternative requires 3D modeling of the surroundings, while in the second instance augmented video sequences could be used. A simple approach, adopted here, is to imitate the movement of a stationary rotating camera by resampling parts of a panoramic image to produce a succession of frames. Indeed, it is much easier to integrate one perspective view of the reconstructed photo-textured building into a panoramic image (rather than insert the suitable perspective view into each frame of a sequence). Obviously, sequences generated in this mode are "chained" to a fixed perspective centre (that of the panoramic image). This is a further use of panoramic images, as these are mainly used for visualisation but also for mapping purposes. Such images can be produced either by mutually rectifying single cocentric images (Haggrén et al., 1999), an approach adopted also here, or directly with special rotating cameras (Pöntinen, 2001; Reulke et al., 2001). Thus, all methods employed in the present contribution are single-image techniques.

2 OBJECT AND AVAILABLE DATA

Our site of interest is the Constitution Square in Athens, a most central and historic square facing the Greek Parliament. This square has been considerably reshaped during the last decades, as most of its 19th century neo-classical buildings have been demolished and replaced by "modern" new structures. A today's view of the square is shown in Fig. 1. The main intention of the authors is to reconstruct all buildings for which photographs are available, and subsequently insert them into a full video sequence of the square.

Figure 1: Panoramic image of the Constitution Square in Athens, today

The data available to this end are twofold. First, several existing aerial images dating between the 1920 and the 1970, which could be used for extracting scaling data and rough planimetric as well a heighting information. Besides, historic photographs of buildings or scenes, found in photographic archives (of museums, banks, the National Gallery) but also in books or as postcards, will be used for reconstruction and texturing purposes. In Fig. 2 images of several demolished buildings are shown over an old aerial image. They include public buildings, hotels and private residences, yet at times most of them have been actually used for more than one purposes.
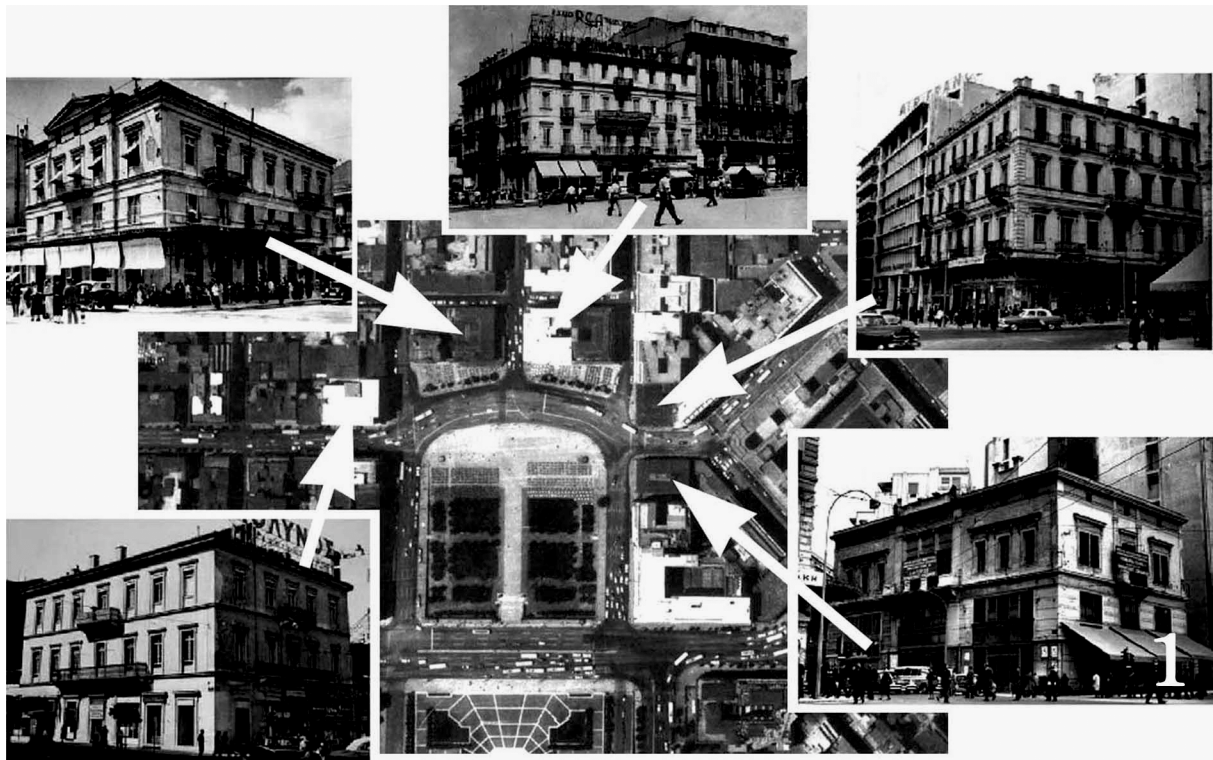

Figure 2: Historic images of demolished buildings (Athens, Constitution Square). Background: aerial image dating from 1962

For the purposes of this presentation only one of these objects (building 1 in Fig. 2) has been treated. This house, the Vouros private residence, had been designed in 1872 by a well-known architect of the time (E. Troumpe) and was torn down in 1960 giving its place to a hotel. The house is characterised as 1 in Fig. 3 (let it also be noted that building 2 in Fig. 3, previously the "Grand Hotel", was house to the largest bookstore in Athens at the time of demolition). The three images of the Vouros residence which have been used here are seen in Fig. 4. The photograph on the left served for single-image reconstruction. From the other two images textures were extracted, the image on the right being the most recent one (the announcement of demolition is already hanging on the balcony).

Figure 3: Historic image of the Vouros residence (1, left) and today's photograph of its successor (1, right)

3   THE PHOTO-TEXTURED MODEL

Although for this building there existed more than one images, a 'monoscopic' method was applied since for other buildings only one image is available. Reference to different methods regarding single-image 3D reconstruction has already been made. As the intention here was simply to visualise rather than accurately reconstruct, we adopted a direct approach by first finding the image coordinates of the three vanishing points with a simple graphical construction (Gracie, 1968). From the six equations thus obtained the six required parameters of orientation (*interior*: image constant c, principal point x_o, y_o; *exterior*: image rotations ω, φ, κ) are computed. With this information, only a known distance on the object façade $Z = 0$ is required to determine the imaging distance Z_o which is sufficient for reconstruction (Karras et al., 1993). This distance was the width of the existing structure. The following values were obtained:

Results for the 5616 x 4214 digital image A					
image constant	image principal point		rotations		imaging distance Z_o
c = 7337 pixels	x_o = 2682 pixels	ω (°)	φ (°)	κ (°)	19 m
	y_o = 2224 pixels	16	-45	-12	

Figure 4: The old photographs of the Vouros residence used. Only image A on the left was used for 3D reconstruction

Figure 5: General views and detail of the photo-textured model of the Vouros residence

These data allow metric modeling if the object's polyhedral nature is exploited. Thus, 3D reconstruction was based on transferring X, Y or Z point coordinates through intersecting planes, assuming their orthogonality or parallelism. The model was completed with all required additional details, e.g. the balcony lattices. For texturing the model other images were also used, while artificial texture was added wherever needed (texturing and animation was performed in 3DStudio Max). Frames of the textured model are seen in Fig. 5.
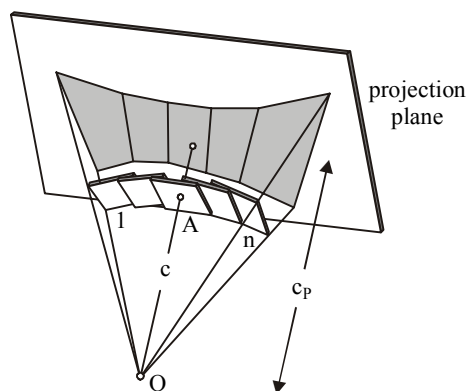

Figure 6: Creation of a panoramic image by rectifying cocentric images (1 to n) with respect to image A

4 THE PANORAMIC IMAGE

As already mentioned, panoramic images can be generated either with specially designed rotating cameras or simply through mutual rectification of images taken from the same standpoint. If full view (360°) panoramic images or images with very large angular openings (for instance > 150°) are required, then a cylindrical projection is needed. For smaller angles, projection on a plane suffices.

For this latter alternative, which has been adopted here, the basic process is as follows according to Fig. 6. Images 1 to n have been acquired from the same perspective centre O with considerable overlap (> 50%) and, for the sake of simplicity, with the same camera constant c. Thus, the n images have been produced by exactly the same bundle of rays through O and, consequently, are projectively related to each other. Accordingly, they can be rectified onto a plane with 2D projective transformations and then mosaicked, to form a much "wider" central projection. The plane of the central image A usually serves as the projection plane to avoid extremely strong perspective distortions and very large panoramic images. The projection may be performed either through the collinearity equations (whereby, for instance, a different panoramic image constant c_P can be used) or through the 2D projective equations, conventionally applied in photogrammetric rectification. In the last case, not only the exterior but also the interior orientation of the resulting panoramic image is that of the reference image A.

Figure 7: The images used for the creation of the panoramic image (five images were rectified with respect to image A)

The six images shown in Fig. 7 were rectified and matched to create the panoramic image. They had been taken using a small format amateur camera with a wide-angle lens (f = 24 mm). All images were projectively fit to the reference image A, whose interior and exterior orientation was again found via its vanishing points, and then merged into a panoramic mosaic of dimensions 11656 x 6054. The resulting colour panoramic image is shown in Fig. 8. As the orientation of reference image A also describes the orientation of the panoramic image, the principal point of A was identified and transferred onto the panoramic image (x_o = 6060, y_o = 3900).

Figure 8. The initial panoramic image

The last step for producing the final panoramic image consisted in back-projecting onto it the textured (grayscale) 3D model of the building using the orientation data of the panoramic image. Finally, the resulting "augmented" image was locally processed to allow a smooth fit of the reconstructed building into the image. A part of the final image can be seen in Fig. 9.

Figure 9. Part of the final panoramic image

5 GENERATION OF THE VIDEO SEQUENCE

Finally, the sequence frames were generated from the augmented panoramic image. Following Fig. 10, the panoramic image has a known rotation matrix $\mathbf{R}_P(\omega_P, \varphi_P, \kappa_P)$ and an image constant c_P which are those of the reference image. It was decided to produce

frames 1 to n which would have no rotations with respect to the object system except for a rotation φ about the vertical Y-axis. Thus, let frame i have an image constant c_i and a rotation matrix $\mathbf{R}_i(\varphi_i)$. Besides, as the panoramic image had a considerable upwards rotation ω_P about the horizontal X-axis, a constant downward shift y_o of the principal point of the "video camera" was also introduced in order to compensate for "lowering" the view of the frames due to neglecting the upward rotation of the panoramic image.
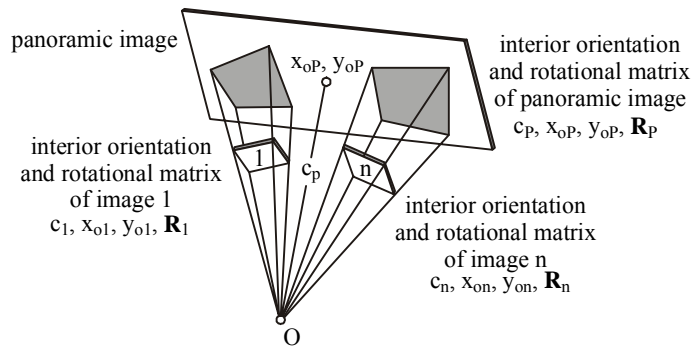

Figure 10. Basic scheme of frame generation from the final panoramic image

If the image coordinates of the panoramic image, referred to its principal point, are denoted by x_P, y_P and the corresponding image coordinates of frame i by x_i, y_i then they are related to each other through the following transformation:

$$\begin{bmatrix} x_P \\ y_P \\ -c_P \end{bmatrix} = k\mathbf{R}_P\mathbf{R}_i^T \begin{bmatrix} x_i \\ y_i - y_o \\ -c_i \end{bmatrix}$$

Scale factor k can be eliminated by dividing the two first equations with the third. From the resulting two equations we can find the projection of the four corners of each frame onto the panoramic image. The area thus defined on the latter is subsequently rectified to produce the final image content of the frame. The first and last frames of the final sequence, along with the corresponding areas of the panoramic image from which they have been resampled, are seen in Fig. 11.
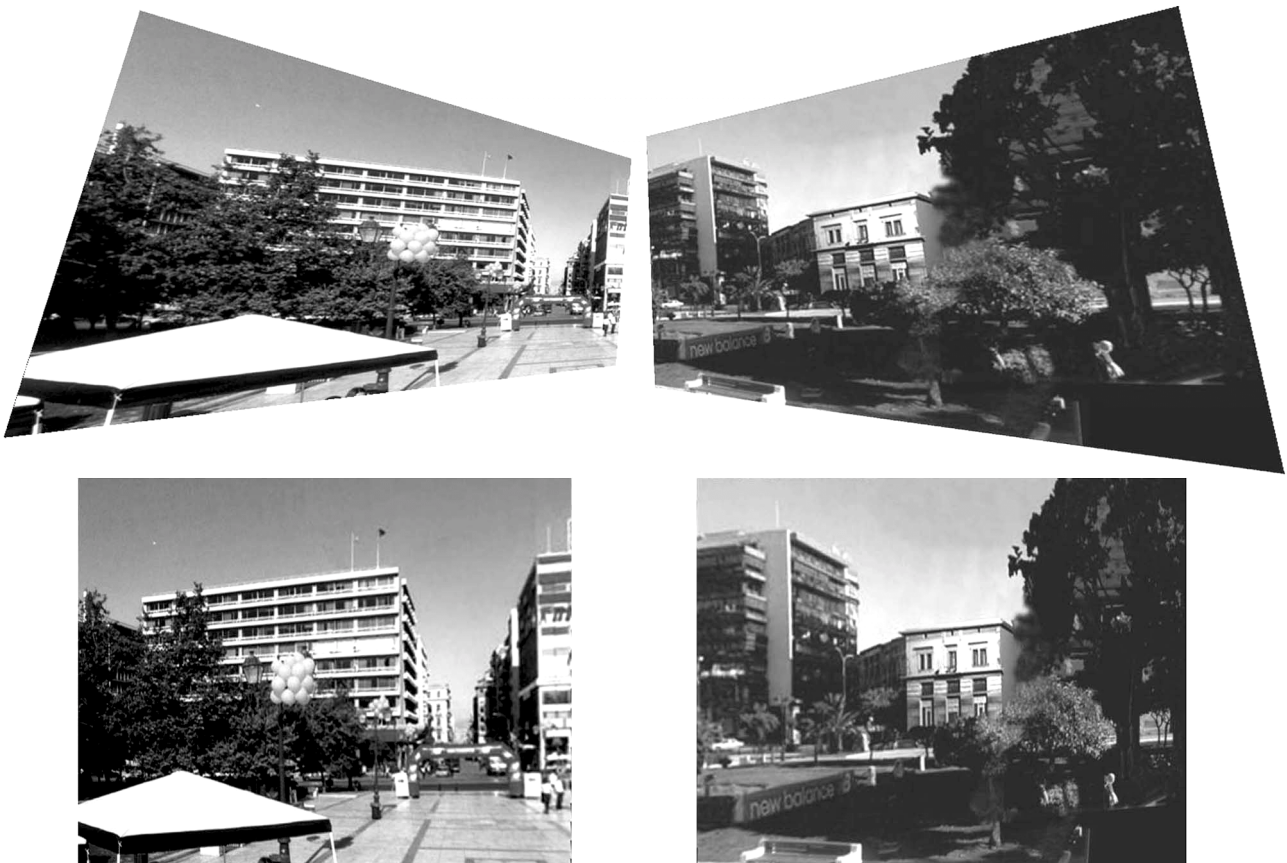

Figure 11. Below: first (left) and last (right) frame of the sequence. Above: corresponding areas on the panoramic image.

For creating the sequence following parameters were used. The size of the frame had a typical proportion (1440 x 1152). For the "rotating video camera" a camera constant $c = 1875$ pixels was used. The sequence concludes with a zooming of the "stationary camera" on the inserted building, during which the camera constant was finally doubled. The constant shift of the principal point, mentioned above, was $\Delta y_o = 200$ pixels. Finally, the camera rotation about the vertical axis had a range of $\varphi_{max} - \varphi_{min} = 82°$, whereby successive frames had a step of $\Delta\varphi = 0.25°$ during the rotational movement. The final video consists of about 800 frames in total and runs at a speed of 25 frames/sec.

6   FUTURE TASKS

In the present contribution simple monoscopic techniques have been combined to create an "artificial" video sequence showing a part of the central square in Athens, which has been largely rebuilt in the last decades. Cocentric images were acquired, mutually rectified and mosaicked into a panoramic image. The 3D model of a demolished building was reconstructed, photo-textured and then introduced into the panoramic image. The resulting image was used to generate frames which were finally combined into a sequence. This experience indicates the possibility to employ simple means for generating a product which, although not highly accurate, is visually powerful and hence very useful in architectural and archaeological representation.

A question encountered has to do with the somewhat non-uniform perspective of the panoramic image, which might be attributed to the fact that the overlapping images were only approximately cocentric. If no panoramic camera is used, image acquisition may cause such difficulties. Besides, imprecision may be also explained by errors of manual measurements during rectification (image matching is an alternative). Furthermore, the wide-angle lens used here has introduced significant lens distortion. This may be avoided by using normal lenses or by correcting distortion with one of the simple methods available. As a consequence, the reconstructed building did not fit perfectly into the panoramic perspective. Such questions will be taken into account in the next steps of this work. The intention is actually to reconstruct all buildings of this historic square for which old photographs are available. This reconstruction will be used both to produce more complete video sequences derived from images taken today but also to fully model the site and produce VRML models in a fully virtual environment.

REFERENCES

Bräuer-Burchardt, C., 2001. Façade reconstruction of destroyed buildings using historical photographs (in this volume).

CIPA Task Group 2 on "Single images in Conservation", 2001. Web site: http://info.uibk.ac.at/sci-org/cipa/tg2_1.html.

Gracie, G., 1968. Analytical photogrammetry applied to single terrestrial photograph mensuration. Presented at the Xth I International ISP Congress, Lausanne, Switzerland, 8-20 July, pp. 24.

Haggrén, H., Pöntinen, P., Moronen, J., 1999. Cocentric image capture for photogrammetric triangulation and mapping and for panoramic visualisation. In: Videometrics VI (eds. S. F. El-Hakim, A. Grün), Proc. SPIE, vol. 3641, pp. 17-21.

Karras, G.E., Patias, P., Petsa, E., 1993. Experiences with rectification of non-metric digital images when ground control is not available. Paper presented at the XVth CIPA International Symposium, Bucharest, 22-26 September.

Karras, G. E., Petsa, E., 1999. Metric information from single uncalibrated images. XVIIth CIPA International Symposium, Olinda, Brazil (proceedings available in CD).

Pöntinen, P., 2001. Image mosaics with rotating camera (in this volume).

Reulke, R., Börne, A., Scheele, M., Klette, R., Huang, F., Wei, S.-K., 2001. Stereo-recording and stereo-processing of high resolution digital panoramic cameras (in this volume).

van den Heuvel, F. A., 2001. Reconstruction from a single architectural image from the Meydenbauer archive (in this volume).