Revisiting District Six: A Case Study of Digital Heritage Reconstruction from Archival Photographs

Christopher de Kadt*
University of Cape Town

James Gain†
University of Cape Town

Patrick Marais‡
University of Cape Town

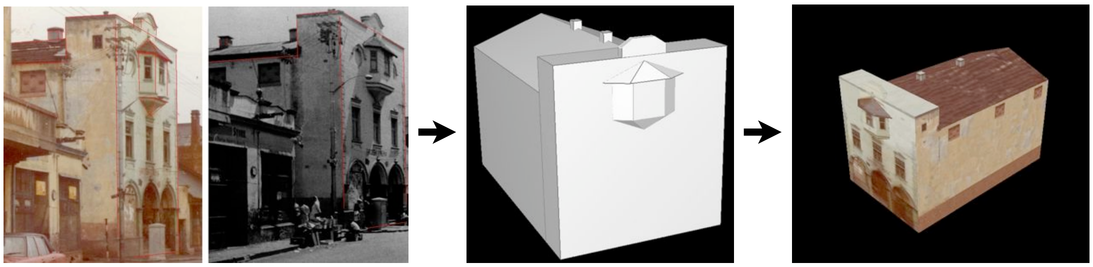

Figure 1: *Reconstruction workflow: building edges are marked in archival photographs, from which a geometric model is derived. Finally, texture synthesis and editing is used to reconstruct the final appearance.*

Abstract

This paper investigates the digital reconstruction of destroyed buildings from small sets of old, uncalibrated photographs. The application domain is the heritage preservation of District Six – a mixed race area in Cape Town that was leveled during the South African Apartheid regime and whose residents were forcibly removed.

Our framework uses a combination of semi-automatic camera calibration, model-based architecture-specific photogrammetry, and texture synthesis to reconstruct the geometry and texture of a building so that it can be incorporated into a heritage-based virtual environment, such as a museum display. These techniques are well established in isolation; the purpose here is to discover if they can be adapted to damaged and uncalibrated photographs, where the time periods and chromatic schemes differ or where, in the worst case, only a single photograph is available.

To test the effectiveness of the reconstruction framework we consider three representative cases of District Six architecture. All three cases were reconstructed successfully with some provisos concerning uneven ground, intricate building features, and unfavourable camera angles.

CR Categories: I.3.8 [Computing Methodologies]: Computer Graphics—Applications; I.4.1 [Computing Methodologies]: Image Processing and Computer Vision—Digitization and Image Capture

Keywords: heritage preservation, photogrammetric reconstruction, texture synthesis

1 Introduction

The use of digital reconstruction techniques in the field of cultural heritage is widespread and increasing, in part due to the recent maturity of supporting technologies such as laser ranging scanning and photogrammetry [Beraldin et al. 2005]. The intent is to generate three-dimensional models of heritage artifacts, ranging from single objects, such as coins and urns, to entire sites or even cities [Müller et al. 2006]. The ultimate purpose is to make these digital models available for archiving, visualisation and detailed study.

District Six, the heritage site under consideration here, is particularly challenging from a reconstruction perspective (as discussed further in Section 3). The area was almost completely destroyed by 1982 (bar a few churches) and the surviving photographs in the District Six museum archive have age-related damage, a variety of chromatic schemes, and no contextual or calibration metadata. Worse, a given building is often only represented by a single photograph.

To handle this, our reconstruction strategy (explained fully in Section 4) has two major components. First, we implement a geometry reconstruction framework, based on the architecture-specific photogrammetric technique of Debevec et al. [1996], that is used to recreate the geometry of a building from photographs. This requires the user to design a basic model of the building, using a set of geometric primitives, and then define correspondences between the edges of this model and the edges of the building visible in the input photographs. This approach enables the exploitation of constraints inherent in the geometry of architectural scenes.

Given the problems associated with material in the District Six photographic archive, this modelling technique offers specific advantages: (a) the ability to reconstruct objects from a single image by exploiting implicit architectural constraints, (b) the ability to model hidden or obscured architecture using symmetry and user knowledge, and (c) tolerance of image error, provided by limiting the number of free parameters.

The second component involves texturing the reconstructed models. To accomplish this, we use a combination of the original textures, extracted from the photographs, and artificial textures, generated from samples of the original using texture synthesis [Efros and Leung 1999]. We also allow features, such as windows and doors to be copied and repositioned on hidden or obscured surfaces.

To validate this reconstruction framework three test cases are con-

*e-mail:chrisdk@gmail.com
†e-mail:jgain@cs.uct.ac.za
‡e-mail:patrick@cs.uct.ac.za

sidered: a large residence (Section 5.1), the landmark British Bioscope (Section 5.2) and a typical small residence (Section 5.3). These represent a range of typical situations: complex architecture (the British Bioscope), a single input photograph (the small residence), and extensive occlusion (the large residence). Most importantly, in Section 6, we discuss the strengths and weaknesses of the framework, including aspects such as highly-detailed geometry, indented features, uneven ground and low quality input photographs.

2 Related Work

Heritage preservation is one of the major application areas for three-dimensional reconstruction techniques. Heritage organisations, such as museums, have an interest in digital reconstruction of artifacts, as virtual exhibits augment physical displays in ways that make them far more attractive to visitors. Historians, archaeologists and artists also use digital reconstructions, both for preservation and when answering questions about objects under examination. For instance, digital reconstructions offer researchers insight into the creation and proportions of statues that they would otherwise not be able to answer [Berndt and Carlos 2000].

In line with this demand, researchers have developed many techniques to facilitate the reconstruction of heritage sites and artifacts. However, no one technique is suitable for all situations. For instance, in reconstructing smaller objects, it is important to capture detail. On the other hand, in reconstructing larger scenes, the problems of high resolution are eclipsed by issues of physical scale. The material available for reconstruction (Does the object still exist? Do we have access to it? Are there accurate measurements available?) and the requirements of the reconstruction (level of detail, texturing, lighting independence, etc.) further emphasise the need for an array of different reconstruction methods.

As a result, a wide variety of reconstruction techniques have been applied to the field of heritage preservation. For instance, Lehner [1992] uses CAD-based methods to assist in building a three-dimensional model of the Giza Plateau in Egypt, including both architecture and terrain. The project made extensive use of maps, survey data and excavation reports and took approximately nine months to complete.

A further example is the well known Digital Michelangelo project [Levoy et al. 2000]. The aim of this project was to develop a framework for building digital models of real-world statues, using laser range scanning. The results of the project are highly-detailed, three-dimensional scans of numerous statues, each up to two billion polygons in size. These models have since been used to render thousands of images of the statues in various synthetic conditions. Despite the excellent results, the scope is limited since it requires expensive equipment and can only be used where the objects still exist.

The virtual recreation of the Buddha statues that once stood at Bamiyan, Afghanistan [Gruen et al. 2002] is closer to our situation, since the statues were destroyed by Taleban militia. A photogrametric approach was taken, using three different sets of photographs – a set found on the Internet, a set of tourist images and a set of three metric images. Although results from this project were good, with the metric image set being used to create a point cloud of over 170,000 points, it is not appropriate in all cases. Where the structure being reconstructed is very regular, a point cloud is not an ideal representation, as the geometry can be modelled more accurately as a set of flat surfaces and regular curves, and a more constrained optimization technique is appropriate.

For extensive reconstruction, such as entire cities, where general architectural style is known but not the specifics of every individual building, a procedural approach is effective. Architectural rules and variations are encoded using procedural shape grammars [Müller et al. 2006], which allow for extensive building variation. For instance, Pompei and Rome have been reconstructed with landmarks modelled in detail but procedural architecture placed elsewhere according to the street plan.

In some cases a hybrid approach is employed for optimal results. For instance, El-Hakim et al. [2004], utilize both photogrammetry and range scanning. Their project is aimed at the reconstruction of highly-constrained objects, and uses a photogrammetry-based solution to determine basic shape. In a few select areas, where a high level of geometric detail is required, this is augmented by range scanning. This technique is extremely effective when used to reconstruct instances of architectural heritage, as show-cased by their model of the Abbey of Pomposa.

Finally, examples of heritage reconstruction also include work by Streilein and Nierderost [1998], Buehrer et al. [1999] and Beraldin et al. [Beraldin et al. 2002]. The first describes the reconstruction of the Disentis monastery, performed through the analysis of still video imagery, the second outlines the use of photogrammetry to assist in the reconstruction of a rock-hewn church in Ethiopia, and the third documents the reconstruction of a Byzantine crypt that was achieved through the use of laser scanning and high-resolution texture information.

3 District Six

Figure 2: *Some examples of the District Six photographic material. [Top] three photographs of the Hanover Building, a local landmark. One of the photographs is in greyscale, while another was taken during the demolition of District Six. [Middle] two photographs of the British Bioscope — a building we go on to reconstruct in Section 5. Again, one image is in greyscale, while the second is in colour. [Bottom] two images of different District Six buildings, for which we were unable to find additional, matching photographs*

District Six, located in Cape Town, South Africa, was one of very few mixed-race areas in the country during the Apartheid regime. However, in 1966, it was declared a white area under the Group areas Act of 1950. By 1982, the entire community had been uprooted;

sixty thousand people were forcibly removed and all of their homes destroyed.

Of what was once a diverse and vibrant community, all that now remains are a select few buildings (primarily churches) and a museum housing a number of artifacts and an archive of several hundred photographs taken before and during the demolition of the area.

Unfortunately, the photographs in this collection (a selection of which appear in Figure 2) are all at least 30 years old, often in poor condition, and with little or no contextual information. For the purposes of photogrammetric reconstruction, this presents a number of challenges:

- **Finding sets of matching photographs.** Because of the large number of images, the fact that the archive is not indexed by street address, and the similarity of much of the Cape architecture of the period, the task of finding two photographs of any one building is both difficult and labour intensive. As a result, it is sometimes necessary to produce a reconstruction from only a single picture.
- **The period over which the photographs were taken.** First, as a consequence of changes in camera technology and age degradation, the quality of photographs varies greatly and this has a direct effect on the reconstruction. Second, if two photographs of a building were taken many years apart, changes may have occurred to the building in the intervening years.
- **Insufficient meta-data.** Most of the photographs in the archive have little or no accompanying information on camera parameters, photograph location, or date.
- **Different chromatic schemes.** Many of the photographs were taken in greyscale, but a number were also captured in colour or sepia. This presents problems when texturing the reconstructed buildings, as it is difficult to combine these schemes in a visually appealing way, without resorting to converting all images to greyscale.
- **The digitization process.** In order to use images from the archive for reconstruction, they must first be digitized. Minor rotations of the image while scanning, or slightly warped images will adversely affect the quality of the digital copy and increase reconstruction error.
- **The problem of calibration.** For most photographs, the camera and lens properties are unknown and this leads to problems with calibration. This is further exacerbated by the practice of cropping photographs, making it difficult to approximate the principal point.

4 Reconstruction Framework

The Reconstruction Framework discussed in this paper comprises two discrete components. The first of these handles the reconstruction of a geometric model from the photographs, while the second deals with the texturing of that model.

This section presents a brief technical discussion of each of these components, followed by a descriptive overview of the entire framework from the perspective of the user.

4.1 Geometry Reconstruction

To facilitate the reconstruction of geometry from photographs a technique, based on the work of Debevec[1996] was implemented.

Essentially, this method involves creating a parameterized model of the scene (both of the geometry and of the cameras), and then optimizing this model by minimizing the error between user-defined observations on the photographs and the equivalent components of the model.

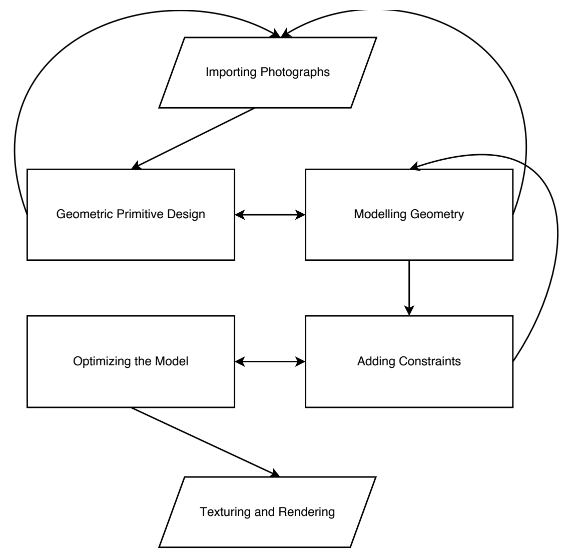

Figure 3: *A user's view of the overall reconstruction framework.*

To this end, the geometry is modelled as set of parameterized, user-defined geometric primitives, where each primitive is defined by at most 9 parameters — 3 for width, depth and height, 3 for its position vector, and 3 for its rotational offset.

To reduce the total number of free parameters introduced by the geometry, constraints can be imposed whereby a parameter of one primitive can be defined as a function of any number of constants or parameters of other primitives.

Similarly, a camera (equivalent to a photograph) is defined by 6 parameters — 3 for its position vector and 3 for its rotational offset. While constraints can be imposed on these parameters, they typically remain unconstrained.

Finally, constraints on the model are added when the user marks out edges on the photographs and associates these edges with the equivalent edges of the model. Indeed, it is through these associations that the principle of stereopsis is applied in the reconstruction.

Once the scene model has been defined, a two-phase Newton's method optimization (as discussed in Taylor and Kriegmann [1995]) is used to minimize an objective function that quantifies the error between the model and the user-defined observations, thus fitting the model to the photographs. This process first optimizes a rough, linear objective function to obtain a good initial estimate, after which a non-linear function is used to fine-tune the model parameters.

4.2 Texturing

Once a scene has been reconstructed, all that remains is to texture the resulting geometry. Many texturing schemes [Debevec et al. 1996] rely on multiple photographs of a scene taken under the same lighting conditions. To overcome the inherent deficiencies of the District Six data in this respect, our framework offers the user a combination of texture extraction, texture synthesis and image manipulation tools.

First, wherever possible, texture information is extracted from the original photographs. This is an automated process that reconstructs textures for each non-occluded face of the model (for each photograph). In cases where these textures are obstructed, incomplete or damaged, a texture synthesis process [Efros and Leung 1999] is used to either complete, repair or entirely reconstruct the texture.

This method of texture synthesis, which can generate a large texture from a small seed, works by modelling a texture as a Markov Random Field [Efros and Leung 1999]. What this means, is that the probability distribution for the value of a particular, unknown, pixel is assumed to be dependant only on the values of the set of other pixels within a certain neighbourhood of that pixel, rather than being dependant on the entire image. The result of this model is that globally a certain amount of noise is permitted, while the local texture structure remains generally well preserved.

Finally, to complete the reconstruction of these damaged or missing textures, features (such as windows, or doors or drain pipes), either imported from external images or extracted from other textures, can be inserted into the newly synthesized texture.

4.3 The User's View

From the user's perspective, the functionality of the framework can be partitioned into six actions (as illustrated by the flowchart in Figure 3):

- **Calibrating and Importing Photographs.** Before any uncalibrated photograph can be used in the reconstruction framework, its calibration parameters must be determined. The absolute minimum requirement is a good estimate of focal length. Usually, it is also necessary to have an estimate of the principal point (the center of camera projection) but, by paying careful attention to cropping issues, it is possible to assume that the principal point is at the image centre. Computing an estimate of the focal length of an image then becomes simple, provided at least one cuboid is visible in the image. Wilczkowiak et al. [2001] show the duality between the intrinsic parameters of a camera and a parallelepiped. In fact, they argue that only when the calibration parameters of a camera are correct, will a parallelepiped in space match one that has been observed in the image. Using this property, by repeatedly modelling (and optimizing) a single cuboid in an image, while varying the assumed focal length, a good estimate of the correct focal length can be obtained. Essentially, the focal length, for which the measured error is minimal after optimization of the model, is the most accurate estimate.
- **Geometric Primitive Design.** A tool has been created for the user to design geometric primitives for later use in the modelling process. These primitives represent canonical architectural structures such as different styles of roof, building, and window. It is important to note that once such a primitive has been created, it is placed in a library for access in subsequent reconstructions.
- **Modelling.** The user selects from the set of pre-designed primitives and arranges them into a rough model of the architectural scene. It is not essential that the model and actual scene have the same proportions, only that they have equivalent geometric primitives in a similar configuration. The image-model correspondence and optimization steps will take care of scale parameters. While modelling the user is able to add constraints, such as a roof resting on top of the walls, to the system. These constraints are essential in reducing the number of free parameters that must be solved during optimization.
- **Defining Image-Model Correspondences.** In order to measure the accuracy of model geometry and camera parameters, correspondences between the images and the user model must be defined. These correspondences are line-based; user-marked edges (or observations) of the real world object in the photographs are matched to edges of the model. When observed edges from different photographs are linked to the same model edge, the result is an implicit feature match. The task of defining these correspondences involves (a) marking out edge features, and (b) linking observed edges to geometry edges. These correspondences form the basis for the objective function that is used to optimize the model.
- **Optimizing the model.** At any point during the modelling phase, provided there is some geometry, at least one image, and one correspondence, the user may elect to automatically optimize the model. This invokes the aforementioned optimization process, manipulating the free parameters of the model, until the geometry and cameras in the system conform to the observations.
- **Texturing and Rendering.** The final operation that the user performs to complete a reconstruction is the texturing of the model. This involves texture extraction, synthesis and manipulation as discussed above, and Figure 4 illustrates the typical workflow for this process.

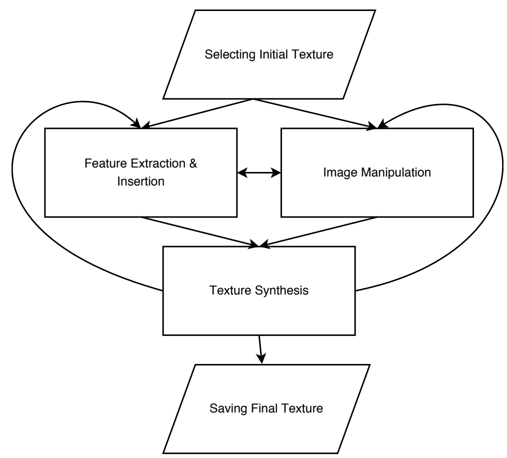

Figure 4: *A user's view of texture reconstruction.*

5 Case Studies

In order to analyse the success of the reconstruction and texturing techniques it is necessary to critically consider a number of real-world situations. We examine three District Six reconstructions in detail. The first two – a large residence and the British Bioscope – are reconstructed from two photographs, and the last is a small, simple, residential building, of which only a single image exists.

5.1 A large District Six residence

The first reconstruction is of a large residence that existed prior to the demolition of the District Six region. The input for this reconstruction comprised two photographs of the residence, shown in Figure 5. The first photograph is from nearby, encompassing only part of the building, and the second from a significant distance, showing the entire building, but at a very low resolution. The field of view for both cameras was computed to be approximately 45 degrees.

Figure 5: *Photo input for a Large District Six residence. The relevant edges of these photographs have been marked out in red by the user.*

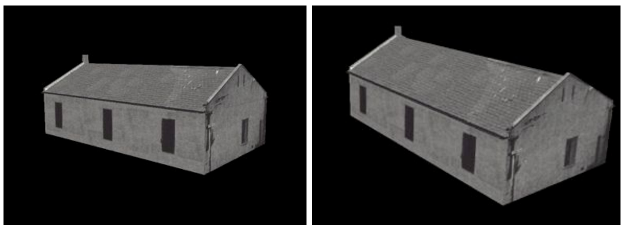

Figure 6: *Final reconstruction results for a large District Six residence.*

The major problem with the reconstruction of this building, is the difficulty of marking edges where the building joins the ground. This is partially as a result of the sloping ground, and partially as a result of a lack of landmarks on the photograph. Aside from this, however, the model, when projected onto the image-planes, matches the marked edges and visible geometry within the images with a high degree of accuracy.

Texturing of the building was performed using the close-up photograph almost exclusively. The second photograph was so far away that the extracted texture was almost entirely without detail. The side wall of the house, visible in both images, was textured using the original information from the first image. The roof and front wall were badly obscured in both images, and thus new textures had to be synthesized. To complete the front wall texture, a window was extracted from the first photograph and inserted in a number of places. Finally, the chimney was textured using information obtained from the second image.

Texture synthesis was particularly successful on the roof, as the texture had a very defined, repeating pattern. The synthesis of paint texture of the front wall was less convincing, as the paintwork texture sample was not particularly well structured. To produce more authentic results in such situations the synthesis procedure is randomly seeded at a number of points.

The final phase in the texturing of the obscured walls involved inserting a number of window features into the base texture. To do this, one copy of a window was extracted from the close-up input image, and then pasted in three places on the newly synthesized texture. For the remaining surfaces, (i.e., the back wall, and other side of the roof) the corresponding front-facing textures, with minor changes, were replicated. The final reconstruction results are shown in Figure 6.

5.2 The British Bioscope

The second sample reconstruction is of the British Bioscope that was located in District Six prior to the re-zoning. This case is substantially more complex than the previous one for a number of reasons. First, the building has far greater geometric complexity. This can be seen in the two input photographs of the building appearing in Figure 7. The building has a bay window, two chimneys and decorative geometry atop the façade. A second complication, is that the two photographs have different chromatic schemes. Furthermore, the second image is substantially older than the first, and of very poor quality. As a result, it was only useful in reconstructing the geometry, and not for texturing purposes. Finally, both photographs are taken from a very similar vantage point, offering a poor view of the front wall. This is exacerbated by large portions of the building being obscured by external geometry, including other buildings, a lamp post, and a pillar box.

Figure 7: *Photo input for the British Bioscope. Note the poor viewing angle with respect to the front wall, and the obstructions present in both photographs.*

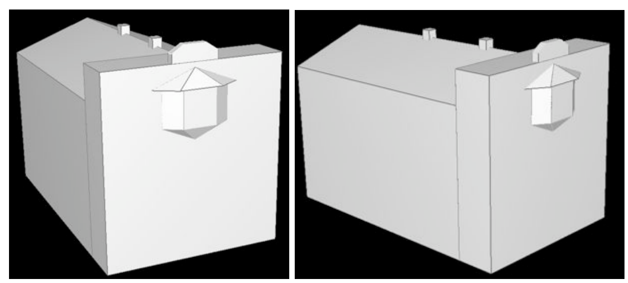

Figure 8: *The reconstructed geometry of the British Bioscope.*

The geometry of the Bioscope was reconstructed successfully as shown in Figure 8. The building was modelled using two cuboids for the base structure (since two different heights were needed), a prism for the roof of the second cuboid, two cuboids for the two chimneys, three "bay window" primitives for the bay window, and two further primitives for the decorative architecture on the crest of the façade. This resulted in a total of 31 free model parameters, including those describing the two cameras.

Figure 9 illustrates the reconstructed model projected onto the original photographs. In general, the projected lines conformed very closely to the user observations as well as the visible geometry. The bay window and decorative geometry did, however, introduce a small amount of additional error into the system (discussed more fully in Section 6.1).

As with the previous example, it was difficult to determine where the building joined the sloping ground but this time, it was possible to use a set of stairs visible in both photographs to establish where the foundations ended and the building began.

The texturing process required extensive synthesis, since most of the roof, large portions of the side wall, and smaller portions of the

Figure 9: *Perceptual test of the British Bioscope reconstruction. Edges in the reconstructed geometry (in green) are overlaid onto the original photograph and closely conform to user drawn edges (still partly visible in red).*

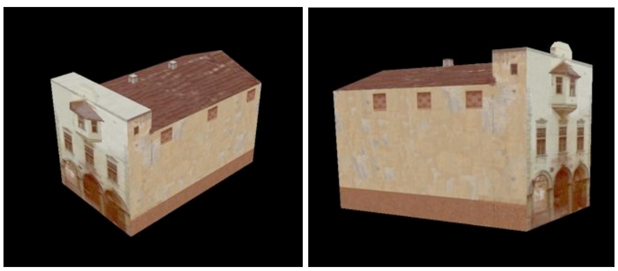

Figure 10: *The final reconstruction results for the British Bioscope.*

front wall are obscured. The side wall was textured using multiple seeds taken from visible portions, to achieve a convincing "old paint" texture. A window, extracted from the visible portion of the side wall, was inserted at regular intervals.

Synthesis of the roof was less successful, largely due to the very low quality of the samples. The visible portion of roof is seen from a grazing angle, and after perspective correction this results in a very low-resolution texture source.

The Bioscope façade was a challenge: it is detailed, obscured in part by external features, and has a poor viewing angle, with the result that features with an offset depth, such as windows, appear distorted. Obscured texture was repaired by simply erasing the invalidated portions, and then refilling with a combination of texture synthesis and feature insertion. Specifically, the left hand side of the front texture was obscured by a lamp post. This was removed, and replaced with a decorative brickwork coining feature, that had been extracted from the right hand side of the façade. The pillar box, obscuring part of the lower texture, was replaced using texture synthesis, with minor features inserted for realism.

The problem of distorted windows was solved by extracting a window, saving it to file, and repairing it using image editing to mirroring the left side onto the right. The window was then loaded as a new feature, and inserted over each of the distorted windows.

The texture for the bay window was largely in a usable condition, with invisible faces reconstructed by copying from visible regions. Finally, the remaining faces (chimneys and decorative roof-piece) used either the original or synthesized textures. These faces are very small, and required no feature insertion.

The overall result of this reconstruction is illustrated by Figure 10. This demonstrates the ability of our reconstruction framework to handle complex cases.

5.3 A small District Six residence from a single photograph

Figure 11: *Photo input for a small District Six residence. [Left] The original input photograph, after the relevant edges have been marked out. [Right] The photograph with the reconstructed model projected onto it from the perspective of the reconstructed camera. Note the large error at the top right-hand corner of the roof.*

Figure 12: *The final reconstruction results for a small District Six residence.*

The final case study is of a small residence. This building is architecturally simple, but reconstruction remains problematic as there is only one sample image (shown in Figure 11). Furthermore, the photograph, although providing an excellent view of the front wall, has a correspondingly limited view of the side (and hence its depth). Additional difficulties are caused by the degree of radial distortion evident in the image, and by the roof of the building, which sags along the depth axis, both of which introduce non-linearities.

The geometric reconstruction of the house is correct with two caveats. First, due to sagging and radial distortion, the top right-hand corner of roof in the reconstructed model, fails to conform closely to the user's observations (as shown in Figure 5.3). Second, the side wall appears to be longer than expected. However, due to the poor camera angle, it is impossible to visually determine the actual wall length, and the projected model does match up accurately with that portion of the image. Although this building sits on a slope, this presented no problems since the border of the foundation is clearly visible. The building was modelled, very simply, with a cuboid for the base and a prism for the roof. Including the camera parameters, the entire model comprised only 10 free parameters.

The texturing of the house was relatively simple. For the front wall, the original high fidelity texture from the photograph could be used directly. Similarly, texture for the visible side of the roof was automatically extracted without any difficulties and then mirrored onto the hidden side. Although one side wall is entirely visible in the photograph, the view is so grazing that the extracted texture was unusable. To solve this problem, a new texture was synthesized from a paint sample taken off the front wall, after which a window was inserted at the correct position in the texture. This procedure was repeated for hidden wall surfaces.

This example represents a worst case scenario for reconstruction. However, the results remain acceptable. An additional photograph would ensure that the depth of the building is perceived correctly, but that was not possible in this case. The final reconstruction results are shown in Figure 12.

6 Discussion

In order to analyse the success of the reconstruction strategy, the accuracy of the geometry reconstruction, and quality of the texturing strategy must be independently assessed.

6.1 Geometry Reconstruction

When considering the accuracy of the reconstruction of geometry from images, it is essential to understand that this technique simply fits a constrained model to a set of user defined observations. As a result, any inaccuracies caused by lens distortion, poor image quality, or user input error will result in some degree of model error. Furthermore, where geometry cannot be exactly modelled by a set of primitives, additional approximation errors may be introduced.

To measure reconstruction error we use the area between the marked edge and the projected edge. This metric is then normalised by the overall length of the marked edge. Effectively, this measure describes the distance (on average), in pixels, between the marked line and the projected model edge.

One example of where error is introduced into the system through a combination of poor image quality and geometric approximations is the roof of the small residence. In this case, the model edges are, on average, 1.14 pixels out, with a standard deviation of 1.23 pixels. However, much of this error can be attributed to a single edge at the top right of the roof of the building, which deviates from its observation by 4.32 pixels on average. Discounting this edge, the model edges are, on average, 0.86 pixels out, with a standard deviation of 0.53 pixels.

In the case of the British Bioscope, the model generally conforms very closely to the user observations. The edges are, on average, only 0.71 pixels out, with a standard deviation of 0.46 pixels. Although this reconstruction has a low degree of error, there are still problems with the bay window and decorative geometry above the façade. For these edges, the average error rises to 0.96, with a standard deviation of 0.43. In this case, the additional error is clearly a result of the complex, curved geometry in the scene.

Considering these problems, the three reconstructions have an acceptable level of error. The quality of photographs is generally poor, and it is often difficult to see exactly where the edges of the building are, which makes marking observations imprecise. In addition, there is always some degree of lens distortion, and as a result edges are often very slightly curved in the photograph.

The geometry reconstruction process is painstaking and time consuming. Marking and modelling a reasonably complex building will take in the region of two hours. However, the optimizations run in a reasonable time, with the non-linear optimization taking less than 10 minutes for even the most complex cases, on a Pentium IV, 3.0 GHz single core machine.

Although the geometry reconstruction technique was, on the whole, successful, there are a number of issues to consider:

- **Highly detailed geometry.** Where there is very complex geometry that cannot be modelled by geometric primitives (such as decorative architecture), or where the resolution of the image does not permit the user to accurately distinguish subtleties within the building structure, approximate models must be used, with a resulting loss of detail.
- **Indented features.** It is difficult to accurately model walls with slight irregularities and indentations (at windows, for example). Usually, such surfaces are modelled as flat, and the textures create the necessary effect, but, when a photo offers a very poor view of a wall, this complicates the texturing process.
- **Uneven ground.** Many buildings lie on sloping or uneven surfaces, and at times determining where the foundation ends, and the building begins, can be problematic. Fortunately, in real world situations it is often possible to find delimiting features that allow a solution.
- **Low quality input images.** With low quality input images, it is more difficult to achieve accurate reconstructions, because the quality of the reconstruction depends almost entirely on the quality of the observations.

6.2 Image-based Texturing

Determining the success of a texturing approach to reconstruction is difficult, as what constitutes a "good" or "realistic" texture is highly subjective. Furthermore, it is impossible to directly compare the textured model with the original photographs, as so much of each texture is regenerated, and large portions of the buildings are obscured from view in the photographs. This implicitly undermines the validity of any quantitative analysis.

Nevertheless, we can make qualitative observations. The synthesis procedure is effective; damaged textures were repaired, yielding results that closely matched the original sample textures. This was most successful for structured textures, such as the roof tiles in Figure 6, but also yielded satisfactory results for paint textures, as shown in Figure 10. In general, this technique works best if seeded at several points, and with a larger sample texture.

The extraction and insertion of features allows a user to modify the synthesized textures, and ensures that they conform closely to the appearance of the original building. Figures 6 and 10 clearly illustrate how inserted features, such as windows, add greatly to the overall quality of the final texture. In Figure 10, the window features were extracted, modified, and then reinserted to solve the indentation problem.

Figure 13 illustrates the success of our texturing scheme. These are renderings of the British Bioscope, using only the textures that were extracted from the input images. In Figure 13[left], none of the damaged texture material is removed, while in Figure 13[right], all damaged material has been erased. Both of these renderings are clearly not in a usable state, and are, in contrast to those in Figure 10, of very poor quality.

As with the geometry reconstruction, there are some issues associated with the texturing strategy:

- **Synthesis speed.** The texture synthesis procedure is slow: the computation cost scales linearly with the size of the empty texture, the size of the sample texture and the size of the neighbourhood window. Synthesizing a large texture from a large input sample can take up to half an hour. However this is certainly amenable to acceleration [Wei and Levoy 2000; Nealen and Alexa 2003; Zelinka and Garland 2002].
- **Synthesis errors.** Occasionally, the texture synthesis procedure produces visually annoying repeating texels or a loss of structure, a problem also observed by Efros et al. [1999].

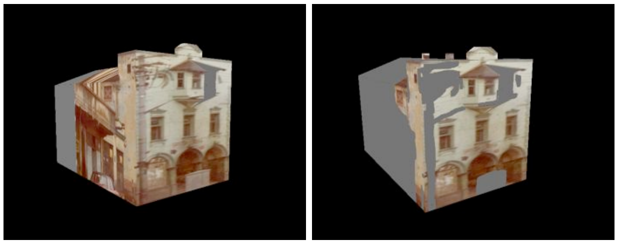

Figure 13: *The original British Bioscope textures. [Left] a synthetic view of the British Bioscope, using all extracted texture material. [Right] The same viewpoint, but with damaged material removed.*

These problems can mostly be prevented, however, by inserting multiple seeds into the new texture.

- **Poor input textures.** Where the input texture is particularly poor, it becomes difficult to produce acceptable results, as the low quality of such textures is propagated by synthesis.

7 Conclusion

Using our framework a user is able to reconstruct with an acceptable degree of geometric accuracy and visual fidelity, models of heritage buildings, from as little input as a single photograph. The framework successfully combines parallelepiped camera calibration, model-based architectural photogrammetry, texture synthesis and feature editing. However, one consideration, clearly evident in all aspects of the reconstruction strategy, is that the quality of the results is highly dependant on the quality of the input material. Where the original photographs are damaged or distorted, it becomes harder to produce accurate models with high quality textures.

While some issues can be overcome: chromatic schemes can be matched [Welsh et al. 2002], radial image distortion corrected [Heikkila and Silven 1997; van den Heuvel 1999] and texture synthesis accelerated [Wei and Levoy 2000; Nealen and Alexa 2003; Zelinka and Garland 2002], perhaps the best investment of effort would be in improving the user interface. Currently, the reconstruction of a reasonably complex scene can take a number of hours to perform. While automation of the reconstruction would be problematic (due to the quality and variety of input data), it would be worth investigating methods of speeding up the modelling process through an efficient user interface.

References

BERALDIN, J.-A., EL-HAKIM, S., GODIN, G., VALZANO, V., BANDIERA, A., AND LATOUCHE, D. 2002. Virtualizing a byzantine crypt by combining high-resolution textures with laser scanner 3d data. In *VSMM 2002*, 3 – 14.

BERALDIN, J.-A., PICARD, M., EL-HAKIM, S., GODIN, G., BORGEAT, L., BLAIS, F., PAQUET, E., RIOUX, M., VALZANO, V., AND BANDIERA, A. 2005. Virtual reconstruction of heritage sites: Opportunities and challenges created by 3d technologies. In *The International Workshop on Recording, Modeling and Visualization of Cultural Heritage*.

BERNDT, E., AND CARLOS, J. 2000. Cultural heritage in the mature era of computer graphics. *Computer Graphics and Applications 20*, 1, 36–37.

BUEHRER, T., LI, Z., GRUEN, A., FRASER, C., AND RUTHER, H., 1999. Photogrammetric reconstruction and 3d visualisation of bet giorgis, a rock-hewn church in ethiopia.

DEBEVEC, P. E., TAYLOR, C. J., AND MALIK, J. 1996. Modeling and rendering architecture from photographs: A hybrid geometry- and image-based approach. In *SIGGRAPH '96*, ACM Press, 11–20.

DEBEVEC, P. E. 1996. *Modeling and Rendering Architecture from Photographs*. PhD thesis, University of California at Berkeley, Computer Science Division, Berkeley CA.

EFROS, A. A., AND LEUNG, T. K. 1999. Texture synthesis by non-parametric sampling. In *IEEE International Conference on Computer Vision*, IEEE Computer Society, Corfu, Greece, 1033–1038.

EL-HAKIM, S. F., BERALDIN, J.-A., PICARD, M., AND GODIN, G. 2004. Detailed 3d reconstruction of large-scale heritage sites with integrated techniques. *IEEE Computer Graphics and Applications 24*, 3, 21–29.

GRUEN, A., REMONDINO, F., AND ZHANG, L. 2002. Image-based reconstruction of the great buddha of bamiyan, afghanistan. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XXXIV*, 5/W12.

HEIKKILA, J., AND SILVEN, O. 1997. A four-step camera calibration procedure with implicit image correction. In *Computer Vision and Pattern Recognition*, IEEE Computer Society, 1106–1112.

LEHNER, M., 1992. The gize plateau mapping project.

LEVOY, M., PULLI, K., CURLESS, B., RUSINKIEWICZ, S., KOLLER, D., PEREIRA, L., GINZTON, M., ANDERSON, S., DAVIS, J., GINSBERG, J., SHADE, J., AND FULK, D. 2000. The digital michelangelo project: 3d scanning of large statues. In *SIGGRAPH 2000*, ACM Press, K. Akeley, Ed., 131–144.

MÜLLER, P., WONKA, P., HAEGLER, S., ULMER, A., AND GOOL, L. V. 2006. Procedural modeling of buildings. In *SIGGRAPH '06*, ACM, New York, NY, USA, 614–623.

NEALEN, A., AND ALEXA, M. 2003. Hybrid texture synthesis. In *EGRW '03: Proceedings of the 14th Eurographics workshop on Rendering*, Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, 97–105.

NOCEDAL, J., AND WRIGHT, S. J. 1999. *Numerical Optimization*. Springer-Verlag.

STREILEIN, A., AND NIERDEROST, M., 1998. Reconstruction of the disentis monastery from high resolution still video imagery with object oriented measurement routines.

TAYLOR, C. J., AND KRIEGMAN, D. J. 1995. Structure and motion from line segments in multiple images. *IEEE Transactions on Pattern Analysis and Machine Intelligence 17*, 11, 1021–1032.

VAN DEN HEUVEL, F. 1999. Estimation of interior orientation parameters from constraints on line measurements in a single image. *International Archives of Photogrammetry and Remote Sensing 32*, 5/W11, 81–88.

WEI, L.-Y., AND LEVOY, M. 2000. Fast texture synthesis using three-structured vector quantization. In *SIGGRAPH 2000*, ACM Press, 479 – 488.

WELSH, T., ASHIKHMIN, M., AND MUELLER, K. 2002. Transferring color to greyscale images. *ACM Trans. Graph. 21*, 3, 277–280.

WILCZKOWIAK, M., BOYER, E., AND STURM, P. 2001. Camera calibration and 3d reconstruction from single images using parallelepipeds. In *Proceedings of the 8th International Conference on Computer Vision, Vancouver, Canada*, IEEE Computer Society Press, vol. 1, 142–148.

ZELINKA, S., AND GARLAND, M. 2002. Towards real-time texture synthesis with the jump map. In *EGRW '02: Proceedings of the 13th Eurographics workshop on Rendering*, Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, 99–104.

ZHANG, Z. 2000. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence 22*, 11, 1330 – 1334.