Relightable Buildings from Images

F. Melendez[1][†], M. Glencross[2][‡], G. J. Ward[3][§] and R. J. Hubbold[1][¶]

[1] The University of Manchester
[2] Loughborough University [3] Dolby Canada

Abstract
We propose a complete image-based process that facilitates recovery of both gross-scale geometry and local surface structure to create highly detailed 3D models of building façades from photographs. We approximate both albedo and sufficient local geometric structure to compute complex self-shadowing effects, and fuse this with a gross-scale 3D model. Our approach yields a perceptually high-quality model, imparting the illusion of measured reflectance. The requirements of our approach are that image capture must be performed under diffuse lighting and surfaces in the images must be predominantly Lambertian. Exemplars of materials are obtained through surface depth hallucination, and our novel method matches these with multi-view image sequences that are also used to automatically recover 3D geometry. In this paper we illustrate the effectiveness of our approach through visually realistic reconstructions of historic buildings in two test cases, together with a table showing the breakdown of effort required to reconstruct each.

Categories and Subject Descriptors (according to ACM CCS): I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Color, shading, shadowing, and texture;

1. Introduction

Existing techniques for building complete 3D models involve a tradeoff between the quality of the recovered models and the amount of effort and cost to create them. The accepted way to create a high-quality model is to use a laser scanner and reflectance measurement equipment, but this requires both expensive equipment and significant effort for both data capture and final model assembly. In contrast, image-based approaches offer the promise of low-cost equipment, simple capture requirements and automated model assembly. The challenge for image-based methods is to build complete models with sufficient local detail that compare favourably with more labour intensive approaches. Few current methods exist to easily capture and incorporate local detailed surface *meso-structure*, such as cracks and protrusions in stone walls, necessary to create realistic appearance under novel lighting. Often, surface detail is con-

veyed by textures recovered from images, *whose appearance is only valid under the originally photographed viewing and lighting conditions*.
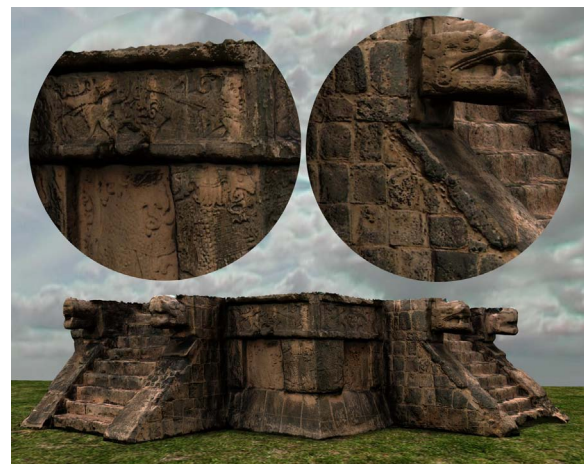

Figure 1: *Eagles and Jaguars model reconstructed with our system and rendered under novel lighting conditions.*

[†] e-mail: francho.melendez@manchester.ac.uk
[‡] e-mail: M.Glencross@lboro.ac.uk
[§] e-mail: gward@dolby.com
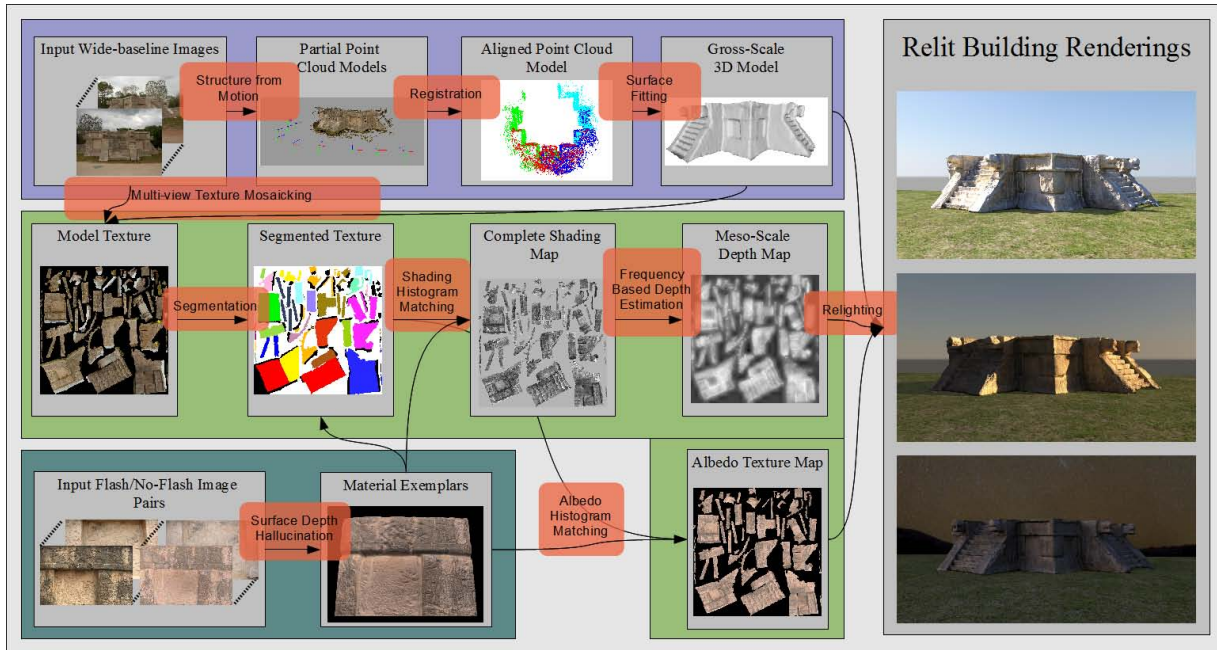[¶] e-mail: roger.hubbold@manchester.ac.uk

Figure 2: *Overview of our façade modeling pipeline.*

Furthermore, while texture includes colour information, separating diffuse reflectance *albedo* from lighting is an inherently ill-posed problem. This separation is key to conveying visual realism of textured surfaces under novel lighting. In this paper, we present a largely automatic process for recovering visually faithful, relightable models of building façades from photographs. A key aspect of our process is that it uses only simple digital SLR equipment and has low data capture and labour requirements, requiring only diffuse lighting (cloudy skies) and predominantly Lambertian materials. Our novel contribution is the use of an image-based statistical transfer method for approximating surface detail for entire building façades from material appearance models captured at close range. This transfer technique can be easily incorporated into other image-based 3D reconstruction pipelines to improve the visual richness of reconstructed scenes. The remainder of this paper provides an overview of our implementation, explains our appearance transfer process, and shows the effectiveness of our method through reconstructions of two historic sites: the platform of the Eagles and Jaguars, from the Ancient Mayan site of Chichén Itzá, Mexico; and Clifford's Tower in York, UK.

2. System Overview

An overview of our 3D reconstruction pipeline is shown in Figure 2 and described in detail later in this paper. The pipeline has three main stages: gross-scale geometry capture; material appearance capture, including albedo and meso-scale geometry; and texture recovery and material appearance transfer. For the first two parts, our implementation builds upon existing approaches, the third part (shaded green in Figure 2) details our contribution.

Our 3D reconstruction system obtains a quasi-dense point cloud, camera parameters for each photograph, and a corresponding surface mesh for the photographed structure. A brief overview of this is given in Section 4. We refer to the 3D model recovered from this part of the system as *gross-scale geometry*, and in the context of our pipeline this represents the global 3D shape of the acquired façade (shaded lilac in Figure 2). To complement this global shape information with surface detail, we first begin by capturing accessible samples (at close range) of the different textured materials present in the façade, by applying surface depth hallucination [GWJ*08]. This yields *exemplars* for each material consisting of both albedo and meso-structure (blue-green shading in Figure 2). Using the reconstructed model and camera parameters we create a high-resolution texture mosaic for the complete model from the multi-view image sequence by selecting the best view for each texel. Next we segment the texture mosaic image and assign materials to appropriate exemplars. We then transfer albedo and shading from the exemplars to the segmented texture using histogram matching. Finally we estimate a depth map from the transferred shading image, which can either be fused with the gross-scale geometry or rendered as a bump map. The fine detail in the complete 3D model results in a visually rich and faithful appearance of the original surfaces.

After data capture our entire process, with the exception of segmentation, is fully automatic. To provide context for our contribution, we briefly discuss prior work spanning the three sections shown in our 3D reconstruction pipeline.

3. Previous work

Gross-scale geometry recovery from images uses established methods in the computer vision and graphics literature. These methods fall broadly into semi-automatic [DTM96,GHCH03,DTC04,SSS*08] and fully automatic approaches [FZ98,Nis05,PvV*04,LQ05,Liu07]. In recent years, the focus of research has shifted to automatic methods that take advantage of large photo collections from the Internet to reconstruct buildings [GSC*07, SSS06] and even cities [ASS*09], but unlike earlier methods these often create sparse point clouds. Furukawa et al. [FCSS10] applied dense matching to large data collections and improved the density of the final point-clouds. The combination of resolution limitations and the problem that densely matched point clouds are noisy in comparison to laser-scanned data, means that it is difficult to recover high quality meso-structure. Our implementation uses the method of Liu [Liu07] to obtain gross-scale geometry, but in principle any of the cited methods could be substituted in our pipeline. These techniques are flexible, mostly automatic, and provide good models of gross-scale 3D geometry. But, they also involve a trade-off between the resolution of the recovered model and the size of the scene. This motivates our approach to recover meso-scale detail through a complementary approach, and to combine the gross-scale and meso-scale models.

The problem of **capturing surface appearance** at meso-scale has spawned several approaches. The gold standard is the approach of Dana et al, who digitised small patches of rough, textured surfaces, creating a variation of the BRDF called the Bidirectional Texture Function (BTF) [DMP*00]. This involves capturing images for many different viewing and lighting directions, requiring special equipment, a carefully controlled environment, and resulting in a high storage cost. This is impractical for our purposes. A simpler alternative is to use photometric stereo, which provides both albedo and surface normal orientation from a single viewpoint [Woo80], and has inspired a number of similar approaches [RB99, PCF05, WG09]. These usually require a camera/multi-flash rig for data capture. A more straightforward approach using a single flash, but delivering perceptually validated surface appearance models consisting of both per-pixel depth and albedo was published by Glencross et al [GWJ*08]. These methods capture only relatively small, globally flat samples of surfaces, and/or require controlled lighting, which makes them impractical for direct application to large façades. Our paper presents an extension of these previous methods to overcome these limitations.

The third stage of our reconstruction pipeline is **texture recovery and material appearance transfer**. Obtaining the best texture from multi-view images is a common problem in image-based reconstruction systems [BMR01, CCRS02, LI07, BFA07]. The objective is to use the best view per triangle while minimizing the seams between views. Mosaic composition is a similar problem, where images from different viewpoints are combined together to produce a seamless image. Agarwala el al. [ADA*04] presented a Markov Random Field formulation for image stitching and mosaic composition, where an energy function is minimized. In our work, we apply a similar formulation to Agarwala et al [ADA*04], but to the multi-view texturing problem. Texture alone cannot capture material appearance, except under the specific lighting conditions of image capture. To achieve estimates of albedo and surface depth for entire façades we draw inspiration from previous work in texture transfer [EF01,MKC*06] and texture synthesis [HB95,DB97] together with the work of Glencross et al [GWJ*08]. The latter work demonstrated how to transfer shading and albedo from a sample to a single image of a similar material at a similar scale. This forms the basis of the approach in this paper, but we further extend it to function with large scale changes, multiple materials, and non-orthogonal views. We are not aware of any approaches dealing with the specific problem of transferring albedo and depth over a large scale change. In the following three sections we describe our system details.

4. Gross-scale Geometry Recovery

Our system uses a wide-baseline 3D reconstruction approach which solves three specific problems: estimating 3D depth, combining small reconstructions into a larger model, and surface fitting. We briefly outline each of these in turn.

The input to our system is an unordered collection of wide-baseline images. Capturing suitable input image sequences is straightforwardly performed with a hand-held camera. There is a trade-off between the number of images used, image resolution, performance, and maximizing features matched between images. Our system requires each region to be visible and matched in at least three images to be reconstructed. We typically capture successive images with 70% overlap. For depth estimation, our system uses fully automatic algorithms to first find feature correspondences [Low99]. To these we apply a robust multi-view stereo method to obtain all camera parameters, and a 3D point cloud that forms the basis of our gross-scale geometry [HZ03]. In our implementation, the initial 3D point cloud acquired by using the originally matched features is used to seed a phase analogous to dense correspondence yielding a quasi-dense representation [Liu07]. In principle, the incremental nature of the reconstruction algorithm can handle an arbitrary number of input images; in practice however, this is memory-limited. The bundle adjustment stage, where 3D point positions and camera parameters are optimised iteratively, is the most memory-demanding component in the reconstruction pipeline, increasing linearly with respect to the

total number of 3D points and number of cameras. This limits the number of points that can be constructed in a single sequence to around 1 million resulting in a total memory footprint of around 4GB. Since the SLR camera we use provides 3900×2616 pixel resolution, we scale these images down to 20% or 30% to reconstruct the gross-scale geometry, allowing us to use around 20 to 40 images for a typical reconstruction.

Recovering larger more complicated façades may require more input images, and to deal with such scenes we align partial reconstructions. This is also a common problem for scanned data [Fit03, IGL03, BL04]. In our case, we simplify the problem by including a common image in consecutive reconstructions, which provides the correspondence between sequences by comparing the reprojected points in 2D from both sequences in the common image. Given these correspondences, we find the rigid transformation between both models. We use the robust RANSAC method (Random Sample Consensus) to obtain a good initial estimation and reject outliers. Then we use non-linear least-squares minimisation to find the transformation that best fits the remaining data. By merging partial reconstructions we are able to reconstruct large scenes with appropriate point density on a desktop PC.

Since the raw output (a 3D point cloud) from the reconstruction process is somewhat noisy and not uniform, we obtain a gross-scale model by applying Poisson surface reconstruction to obtain a triangulated implicit surface [KBH06] enforcing continuity and smoothness. Normally we need to perform some simple manual cleaning of the resulting mesh, to crop the areas that we are interested in, using Mesh-Lab [CCR08].

5. Capturing Surface Appearance

Since the gross-scale 3D reconstruction is not able to reliably capture albedo and meso-scale surface detail, we acquire a set of material appearance models that are representative of the textures contained in the façade. We use the acquisition method detailed in Glencross et al [GWJ*08]. Although this provides models containing albedo, shading and per-pixel depth, we only make use of the histograms of the albedo and shading images for our application. Input data for generating these texture exemplars is captured in RAW format at an accessible location and where the gross-scale geometry is flat. We use a standard digital SLR camera mounted on a tripod and an attached shoe-mounted flash unit. A no-flash (ambient) capture is first obtained, and then a corresponding photograph from the same viewpoint is taken with the flash unit fired, ensuring correct exposure. For further details of how to capture and process exemplars we refer the reader to the original publication. Figure 3 shows an example surface appearance model for one of the test case sites reconstructed to evaluate our work.
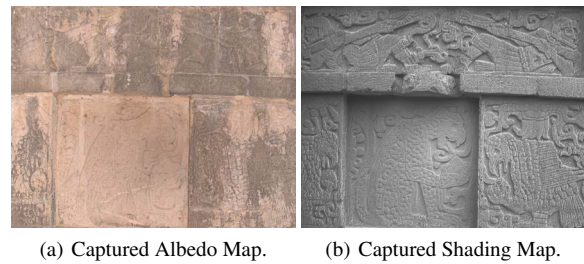

(a) Captured Albedo Map. (b) Captured Shading Map.

Figure 3: *A typical exemplar captured for the Eagles and Jaguars reconstruction model using Surface Depth Hallucination*

6. Texture Recovery and Material Appearance Transfer

In this section we describe the three novel steps that we introduce into our 3D reconstruction pipeline. These are multi-view texturing, transfer of shading and albedo and meso-scale geometry hallucination. Since our intention is to obtain estimated albedo and shading images for an entire façade, and from this to infer meso-scale detail, we need to solve a number of problems. The first is how to construct a texture map for the façade which incorporates the best views for depth hallucination. The second is how to transfer the statistical characteristics of the albedo and shading images of the exemplar to the texture map for the façade. The final problem is how to combine both the gross-scale and recovered meso-scale into a coherent model.

6.1. Multi-view Texturing

To obtain the best possible texture map from all the viewpoints available in the original images we adapt multi-view texturing methods to suit our application. In common with other applications, we want a texture parameterization that minimizes distortion and scaling of the triangles. We use the free software *Graphite* [GRA09] to first create a flat parameterization of our model, allowing for cuts in the geometry to improve the mapping. Next we create a geometry image [GGH02] of the surface mesh. This is the process of sampling the geometry regularly in the parameterization space, encoding the geometry in an image, where RGB channels codify XYZ coordinates. Having the geometry in such a regular structure has the benefit that we can back-project the geometry to every view and obtain a per-view texture map for the complete model, accounting for visibility.

Having every view in the same parameter space, allows us to define an energy function to be minimized. This energy function consists of two cost functions: the data cost and the smoothness cost. The first term represents the cost of assigning a camera view to a texel. We encode in this term the quality of the texel for each view. The closer to fronto-parallel the view, the lower the cost. We use the inverse of

the area of the pixel in the reprojected view to measure this, which accounts for perspective and distance to the camera.

The second term determines the cost of assigning different cameras to neighboring texels, encoding the Euclidean distance in RGB space between the two texels as in Agarwala et al [ADA*04]. This term has the effect of minimizing the seams due to texture changes arising from different photographs. To avoid discontinuities in the texture map we define the pixel neighbourhood using an indirection map [LH06] that takes account of the separation between charts in the texture map. We use graph-cut optimization [BVZ01] to obtain a close-to globally optimal solution. By combining these two energy terms, we strike a good balance between optimal views and the avoidance of seams. Figure 4 shows the reconstructed texture for the Eagles and Jaguars model.

Figure 4: *Combined Texture map from multi-view images.*

We typically create a final texture mosaic of 7000×7000 pixels but we allow the process to create an arbitrary resolution for the final texture map. During the optimization process we need to instantiate in memory the data cost maps for every view. Since the number of views can be large, we run the optimization at a lower resolution depending on the number of views, and then we scale up the association map and sample the texture map at full resolution.

6.2. Transferring Shading and Albedo

Armed with exemplars of the different materials present in the façade and a unified texture map, we need to first segment the texture mosaic and associate the different segments with their respective materials. We employ standard interactive techniques to separate the different regions and create a

simple text file to associate every segment with its material exemplar. The user takes the exemplars as references and select regions in the texture map containing similar global appearance and meso-structure. When the boundaries between materials are well defined, semi-automatic segmentation techniques such as graph-cuts [BVZ01] can be used. When materials are not easily differentiable, a closer user supervision is required to adequately segment the texture. We used GIMP † to generate the results for this paper.

Next, we automatically histogram-match segments with their exemplars [HB95]. This is a key stage in our process since we solve two problems. First, we provide the correct material appearance statistics for local shape estimation via surface depth hallucination, and second we estimate albedo, providing a close approximation to diffuse reflectance. Our approach is motivated by the idea that the image statistics of similar surfaces under diffuse lighting encode information about shading and albedo, and that appearance and depth under such lighting conditions are closely correlated.

Histogram Matching equalizes the shape of the histogram of an image to approximate the shape of another histogram – in this case provided by the exemplar – by matching the their cumulative distribution functions. We use a 256-bin histogram and match the different RGB channels separately. When transferring shading we use the luminance channel of the texture map. This process matches the statistics for each material to those of its associated exemplar, enabling subsequent shape and albedo approximation. The validity of this recombined histogram-matched albedo image depends on how well the global statistics of the segments match the statistics of the exemplar. It is advisable to create a larger number of smaller segments to adequately match the exemplars. In the models included in this paper, an average of 17 segments were used.

6.3. Meso-scale geometry hallucination

The shading image estimated for the full façade is used to compute a per-pixel depth map that we superimpose on the gross-scale model. Applying this technique directly, as previously published, results in some artifacts visible as excessively bumpy depth maps. There are two reasons for this. First, even under diffuse lighting conditions, large features in the building can cast soft shadows on the façade. Such features are already captured by the gross-scale geometry, so should be omitted from the depth map. Second, areas with dark appearance may be due to dark albedo but may instead be interpreted as areas in shade. If these areas are large they also need to be removed from the depth map.

We observe that both artifacts correspond to low frequencies in the texture map. Consequently, we correct the shading image by filtering out the low frequencies, achieved by

---

† http://www.gimp.org

dividing the shading image by a Gaussian blurred version of itself. This filtering preserves the required meso-scale detail. The depth computation, following the multi-scale model from [GWJ*08], is then applied resulting in a per-pixel depth map that we render as a bump map over the gross-scale model.

7. Results

The final models acquired with our method contain a high level of detail in the recovered geometry, and approximate albedo, providing rich appearance under novel lighting conditions. This improves the results of both gross-scale geometry acquisition, and surface depth hallucination, used in isolation. Our method is able to capture interesting surface features, like the reliefs in the façades in Figures 1 and 7, that would be labour intensive to generate manually. Estimates for the breakdown of effort for the models included in this paper is presented in Table 1. The complete process takes around one day of work including capture and processing, the segmentation process being the most time consuming.

Table 1: *Breakdown of effort. York₁ and York₂ = Respectively, Exterior and Interior Façades of Clifford's Tower, E&J = Eagles and Jaguars Platform. (A) = Automatic. (I) = Interactive. Processing over $7K \times 7K$ pixels Textures.*

	$York_1$	$York_2$	$E\&J$
N^o Images Gross-scale	33	7	19
N^o Exemplars	5	6	1
N^o Segments	16	17	18
Capture Gross-scale	30min	10min	20min
Capture Exemplars	1h	1h	10min
Process Gross-scale(A)	1.5h	30min	1h
Model Cleaning(I)	1h	20min	1h
Texturing(A)	15min	15min	15min
Segmentation(I)	3h	3h	3h
Transfer(A)	15min	15min	15min
Depth Estimation(A)	1min	1min	1min
TOTAL Reconstruction Time	7h 15min	5h 15h	6h

Figure 5 shows visual comparisons for one of the panels in the Eagles and Jaguars model reconstructed using different techniques. Our combined model (5(a)) is able to recover both global depth and local detail, while the model containing only the gross-scale (5(d)) fails to recover meso-structure. Comparing with [DTC04] – considered the gold standard in automatic multi-view reconstruction – their model (5(c)) provides sharper geometry than our gross-scale model, but introduces artifacts at the fine detail level. Our transfer system provides more detailed and coherent meso-structure. Depth Hallucination (5(b)) captures high-resolution local detail, but is not able to model the global structure due to the flat-surface assumption.
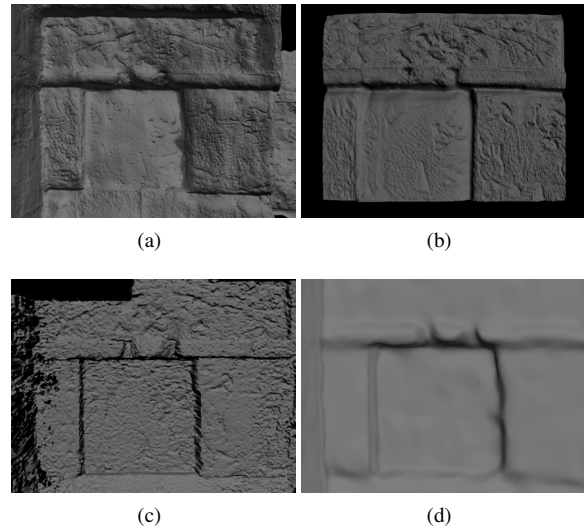

(a)　　　　　　(b)

(c)　　　　　　(d)

Figure 5: *(a) Close-up view of panel reconstructed with our transfer system, (b) using Depth Hallucination [GWJ*08], (c) reconstructed with [DTC04], (d) gross geometry without meso-scale detail.*

Figure 6 shows the plausibility of the appearance recovered with our system by comparing side-by-side a photograph of Clifford's Tower ‡ with a rendering of the model under approximately matched lighting conditions.

Figure 6: *(Left) Photograph of Clifford's Tower downloaded from Internet, (Right) Synthetic Rendering under similar lighting conditions.*

8. Limitations and Future work

For the results in this paper, gross-scale geometry was reconstructed using wide-baseline photographs, offering the advantage of simple data capture. Better results might be obtained from using other systems, such as ARC3D [DTC04] to acquire gross-scale geometry, although these involve other tradeoffs. Fortunately, our transfer methods can be easily applied to models captured with other systems.
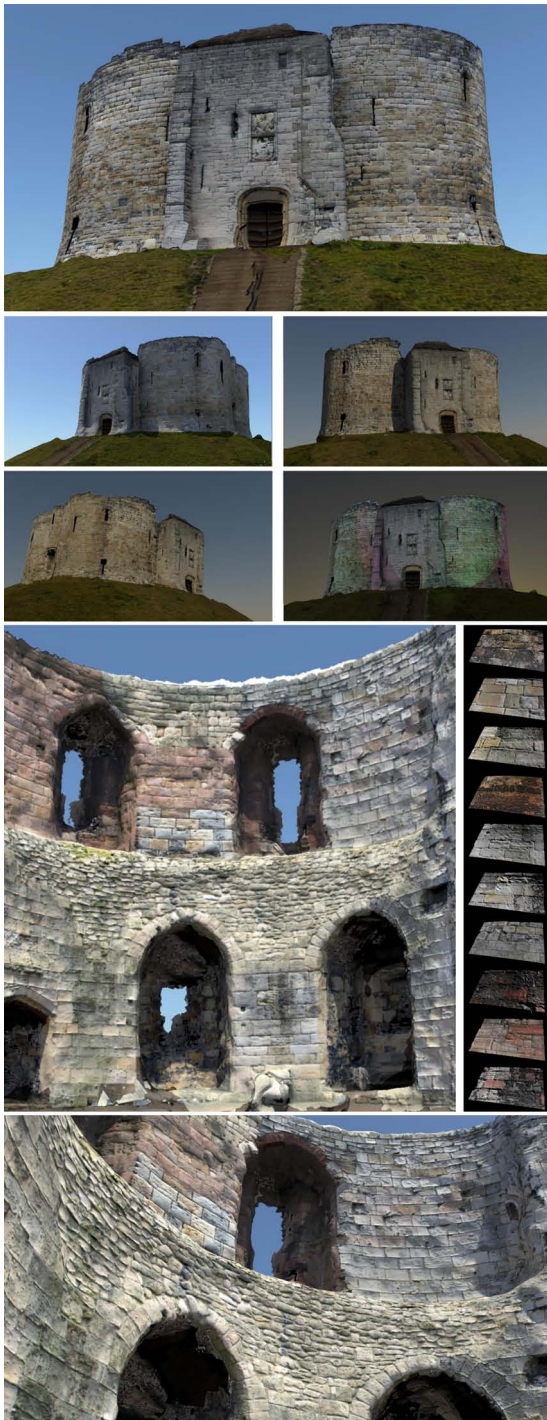
---

‡ http://www.nicolaconforto.com/erasmus/img/york20080927

Figure 7: *Clifford's Tower Model reconstructed with our system, including exterior and interior façades, rendered under different view-points and lightings. We captured 10 materials for these reconstructions shown right-middle.*

In our system, segmentation still requires some effort from the user. More sophisticated methods that perform automatic segmentation by matching against the exemplars are part of our current investigations.

Since our method relies on histogram matching for both albedo estimation and computing the shading image, the effectiveness of the technique depends on the quality of match between the exemplar and associated segment. This is especially important in the albedo computation, since a poor match will result in the wrong color appearance. Other transfer methods that provide local rather than global matching may relax this limitation.

Finally, creating a library of exemplars would further simplify the capture process, by avoiding the need to recapture samples. It would also allow users to enhance existing 3D models.

9. Conclusions

We have described a novel method for constructing models of buildings from photographs. Our method extends existing approaches for geometry acquisition by adding image-based estimation of local depth and albedo. We illustrated this by applying it to capturing both the geometry and detailed appearance of building façades. We showed how the enhanced geometric detail contained in our 3D models enables computation of complex self-shadowing effects. Although this detail is approximate, it is informed by shape information contained in the statistics of the photographs. Our use of histogram matching to obtain albedo is a step forward in separating lighting from geometry in images. Uniquely, the method mitigates the limitations of image-based 3D reconstruction and surface depth hallucination by fusing information captured at different scales. This combination provides relightable 3D models for a range of cultural heritage, visualization and entertainment applications.

References

[ADA*04] AGARWALA A., DONTCHEVA M., AGRAWALA M., DRUCKER S., COLBURN A., CURLESS B., SALESIN D., COHEN M.: Interactive digital photomontage. In *SIGGRAPH '04: ACM SIGGRAPH 2004 Papers* (2004), pp. 294–302. 3, 5

[ASS*09] AGARWAL S., SNAVELY N., SIMON I., SEITZ S., SZELISKI R.: Building rome in a day. In *Computer Vision, 2009 IEEE 12th International Conference on* (sep. 2009), pp. 72 –79. 3

[BFA07] BANNAI N., FISHER R. B., AGATHOS A.: Multiple color texture map fusion for 3d models. *Pattern Recogn. Lett. 28*, 6 (2007), 748–758. 3

[BL04] BAE K.-H., LICHTI D. D.: Automated registration of unorganised point clouds from terrestrial laser scanners. In *In: International Archives of Photogrammetry and Remote Sensing, Vol. XXXV, Part B5, Proceedings of the ISPRS working group V/2* (2004), pp. 222–227. 4

[BMR01] BERNARDINI F., MARTIN I. M., RUSHMEIER H.: High-quality texture reconstruction from multiple scans. *IEEE*

Transactions on Visualization and Computer Graphics 7, 4 (2001), 318–332. 3

[BVZ01] BOYKOV Y., VEKSLER O., ZABIH R.: Fast approximate energy minimization via graph cuts. *IEEE Trans. Pattern Anal. Mach. Intell. 23*, 11 (2001), 1222–1239. 5

[CCR08] CIGNONI P., CORSINI M., RANZUGLIA G.: Meshlab: An open-source 3d mesh processing system, Apr. 2008. 4

[CCRS02] CALLIERI M., CIGNONI P., ROCCHINI C., SCOPIGNO R.: Weaver, an automatic texture builder. In *3D Data Processing Visualization and Transmission. Proceedings. First International Symposium on* (2002), pp. 562 – 565. 3

[DB97] DE BONET J. S.: Multiresolution sampling procedure for analysis and synthesis of texture images. In *Proceedings of the 24th annual conference on Computer graphics and interactive techniques* (1997), SIGGRAPH '97, pp. 361–368. 3

[DMP*00] DANA K., MARSCHNER S., PREMOZE S., RUSHMEIER H., SATO Y., YU Y.: Image-based surface details. ACM SIGGRAPH Course notes, 2000. 3

[DTC04] DICK A. R., TORR P. H. S., CIPOLLA R.: Modelling and interpretation of architecture from several images. *International Journal of Computer Vision 60*, 2 (2004), 111–134. 3, 6

[DTM96] DEBEVEC P. E., TAYLOR C. J., MALIK J.: Modeling and rendering architecture from photographs: a hybrid geometry and image-based approach. In *SIGGRAPH* (1996), ACM, pp. 11–20. 3

[EF01] EFROS A. A., FREEMAN W. T.: Image quilting for texture synthesis and transfer. In *SIGGRAPH* (2001), ACM, pp. 341–346. 3

[FCSS10] FURUKAWA Y., CURLESS B., SEITZ S., SZELISKI R.: Towards internet-scale multi-view stereo. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on* (jun. 2010), pp. 1434 –1441. 3

[Fit03] FITZGIBBON A. W.: Robust registration of 2d and 3d point sets. *Image and Vision Computing 21*, 13-14 (2003), 1145 – 1153. 4

[FZ98] FITZGIBBON A., ZISSERMAN A.: Automatic 3D model acquisition and generation of new images from video sequences. In *Proceedings of the European Signal Processing Conference* (1998), pp. 1261–1269. 3

[GGH02] GU X., GORTLER S. J., HOPPE H.: Geometry images. In *SIGGRAPH '02: Proceedings of the 29th annual conference on Computer graphics and interactive techniques* (New York, NY, USA, 2002), ACM, pp. 355–361. 4

[GHCH03] GIBSON S., HUBBOLD R. J., COOK J., HOWARD T. L. J.: Interactive reconstruction of virtual environments from video sequences. *Computers and Graphics 27*, 2 (2003), 293–301. 3

[GRA09] GRAPHITE: In: http://alice.loria.fr/software/graphite, 2009. 4

[GSC*07] GOESELE M., SNAVELY N., CURLESS B., HOPPE H., SEITZ S. M.: Multi-view stereo for community photo collections. In *International Conference on Computer Vision ICCV* (2007), pp. 1–8. 3

[GWJ*08] GLENCROSS M., WARD G. J., JAY C., LIU J., MELENDEZ F., HUBBOLD R.: A perceptually validated model for surface depth hallucination. *ACM SIGGRAPH 27*, 3 (2008), 59:1 – 59:8. 2, 3, 4, 6

[HB95] HEEGER D. J., BERGEN J. R.: Pyramid-based texture analysis/synthesis. In *SIGGRAPH* (1995), ACM, pp. 229–238. 3, 5

[HZ03] HARTLEY R., ZISSERMAN A.: *Multiple View Geometry in Computer Vision*, second ed. Cambridge University Press, 2003. 3

[IGL03] IKEMOTO L., GELFAND N., LEVOY M.: A hierarchical method for aligning warped meshes. In *3-D Digital Imaging and Modeling, 2003. 3DIM 2003. Proceedings. Fourth International Conference on* (2003), pp. 434 – 441. 4

[KBH06] KAZHDAN M., BOLITHO M., HOPPE H.: Poisson surface reconstruction. In *SGP '06: Proceedings of the fourth Eurographics symposium on Geometry processing* (Aire-la-Ville, Switzerland, Switzerland, 2006), Eurographics Association, pp. 61–70. 4

[LH06] LEFEBVRE S., HOPPE H.: Appearance-space texture synthesis. In *SIGGRAPH* (2006), ACM, pp. 541–548. 5

[LI07] LEMPITSKY V., IVANOV D.: Seamless mosaicing of image-based texture maps. In *Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on* (June 2007), pp. 1–6. 3

[Liu07] LIU J.: *Automatic Scene Reconstruction from Wide-Baseline Images*. PhD thesis, School of Computer Science, University of Manchester, 2007. 3

[Low99] LOWE D. G.: Object recognition from local scale-invariant features. In *Seventh International Conference on Computer Vision (ICCV)* (1999), pp. 1150–1157. 3

[LQ05] LHUILLIER M., QUAN L.: A quasi-dense approach to surface reconstruction from uncalibrated images. *IEEE Transactions on Pattern Analysis and Machine Intelligence 27* (2005), 418–433. 3

[MKC*06] MERTENS T., KAUTZ J., CHEN J., BEKAERT P., DURAND F.: Texture transfer using geometry correlation. In *Symposium on Rendering* (2006), Akenine-Moller T., Heidrich W., (Eds.), Eurographics, pp. 273–284. 3

[Nis05] NISTÉR D.: Preemptive ransac for live structure and motion estimation. *Machine Vision and Applications 16*, 5 (2005), 321–329. 3

[PCF05] PATERSON J. A., CLAUS D., FITZGIBBON A. W.: BRDF and geometry capture from extended inhomogeneous samples using flash photography. In *Computer Graphics Forum* (2005), Eurographics, pp. 383–391. 3

[PvV*04] POLLEFEYS M., VAN GOOL L., VERGAUWEN M., VERBIEST F., CORNELIS K., TOPS J., KOCH R.: Visual modeling with a hand-held camera. *International Journal of Computer Vision 59*, 3 (2004), 207–232. 3

[RB99] RUSHMEIER H., BERNARDINI F.: Computing consistent normals and colors from photometric data. In *Second Conference on 3-D Imaging and Modeling 3DIM* (Ottawa, Canada, 1999), IEEE, pp. 99–108. 3

[SSS06] SNAVELY N., SEITZ S. M., SZELISKI R.: Photo tourism: Exploring photo collections in 3D. *ACM SIGGRAPH* (2006), 835–846. 3

[SSS*08] SINHA S. N., STEEDLY D., SZELISKI R., AGRAWALA M., POLLEFEYS M.: Interactive 3D architectural modeling from unordered photo collections. *ACM SIGGRAPH Asia* (2008). 3

[WG09] WARD G. J., GLENCROSS M.: A case study evaluation: Perceptually accurate textured surface models. In *Symposium on Applied Perception in Graphics and Visualization (APGV)* (2009), ACM, pp. 109–115. 3

[Woo80] WOODHAM R. J.: Photometric method for determining surface orientation from multiple images. *Optical Engineering 19*, 1 (1980), 139–144. 3