Chapter 8

Rouen Revisited

8.1 Overview

This chapter describes *Rouen Revisited*, an interactive art installation that uses the techniques presented in this thesis to interpret the series of Claude Monet's paintings of the Rouen Cathedral in the context of the actual architecture. The chapter first presents the artistic description of the work from the SIGGRAPH '96 visual proceedings, and then gives a technical description of how the artwork was created.

8.2 Artistic description

This section presents the description of the Rouen Revisited *art installation originally written by Golan Levin and Paul Debevec for the SIGGRAPH'96 visual proceedings and expanded for the* Rouen Revisited *web site.*

Rouen Revisited

Between 1892 and 1894, the French Impressionist Claude Monet produced nearly 30 oil paintings of the main façade of the Rouen Cathedral in Normandy (see Fig. 8.1). Fascinated by the play of light and atmosphere over the Gothic church, Monet systematically painted the cathedral at different times of day, from slightly different angles, and in varied weather conditions. Each painting, quickly executed, offers a glimpse into a narrow slice of time and mood.

The *Rouen Revisited* interactive art installation aims to widen these slices, extending and connecting the dots occupied by Monet's paintings in the multidimensional space of turn-of-the-century Rouen. In Rouen Revisited, we present an interactive kiosk in which users are invited to explore the façade of the Rouen Cathedral, as Monet might have painted it, from any angle, time of day, and degree of atmospheric haze. Users can contrast these re-rendered paintings with similar views synthesized from century-old archival photographs, as well as from recent photographs that reveal the scars of a century of weathering and war.

Rouen Revisited is our homage to the hundredth anniversary of Monet's cathedral paintings. Like Monet's series, our installation is a constellation of impressions, a document of moments and percepts played out over space and time. In our homage, we extend the scope of Monet's study to where he could not go, bringing forth his object of fascination from a hundred feet in the air and across a hundred years of history.

The Technology

To produce renderings of the cathedral's façade from arbitrary angles, we needed an accurate, three-dimensional model of the cathedral. For this purpose, we made use of new modeling and rendering techniques, developed by Paul Debevec at the University of California at Berkeley, that allow three-dimensional models of architectural scenes to be constructed from a small number of ordinary photographs. We traveled to Rouen in January 1996, where, in addition to taking a set of photographs from which we could generate the model, we obtained reproductions of Monet's paintings as well as antique photographs of the cathedral as it would have been seen by Monet.

Once the 3D model (Fig. 8.5) was built, the photographs and Monet paintings were registered with and projected onto the 3D model. Re-renderings of each of the projected paintings and photographs were then generated from hundreds of points of view; renderings of the cathedral in different atmospheric conditions and at arbitrary times of day were derived from our own time-lapse photographs of the cathedral and by interpolating between the textures of Monet's original paintings. The model recovery and image rendering was accomplished with custom software on a Silicon Graphics Indigo2. The Rouen Revisited interface runs in Macromedia Director on a 166-MHz Pentium PC, and allows unencumbered exploration of more than 12,000 synthesized renderings.

Figure 8.1: Rouen Cathedral: Full Sunlight, 1894. By Claude Monet. Louvre, Paris.
This is one of a series of nearly thirty paintings of the West façade of the cathedral done by Monet in the years 1892 to 1894. Executed in colors ranging from brilliant yellow to drizzly blue to smouldering red, the series is a definitive impressionist study of the interaction of light and matter. For further reading on this series of paintings see [18, 34].

Figure 8.2: The Rouen Revisited kiosk, shown with the artists Golan Levin (left) and Paul Debevec (right). The brass plate in the middle of the kiosk is a touch pad that lets the user control the angle of view, time of day, amount of fog, and the rendering mode of the cathedral façade. The three rendering modes let the user view the cathedral as it appears in 1996, as it appeared in 1894 at the time Monet's paintings were made, or as it would appear with a Monet painting projected onto the cathedral's 3D geometry. The kiosk design and construction were done with the help of Warren H. Shaw, David Kaczor, Shane Levin, Joe Ansel, Charles "Bud" Lassiter, Scott Wallters, Chris Seguine and Bernie Lubell.

The Presentation

Rouen Revisited is presented in an arch-shaped maple cabinet, seven feet three inches tall (Fig. 8.2). Its front face is articulated by three features: Near the top, a backlit stained-glass rosette (whose design is based on the rosette of the Rouen Cathedral) acts as a beacon for passers-by. Below that, a 17-inch color monitor, configured on its side, provides users with a view onto the cathedral's surface. Finally, a projecting wedge-shaped block at waist-level provides the interface controls for operating the kiosk.

Users explore the surface of the Rouen Cathedral by touching one of three force-sensitive regions exposed within a brass plate mounted on the interface wedge. Each region affords the user with control of a different dimension of the façade:

- Touching the corners of the upper, triangular region of the brass plate allows users to select between renderings of Monet paintings, archival photographs from the 1890's, or new photographs from 1996. Dragging one's finger along this groove creates a blend between these modes.
- Moving one's finger left and right inside the central, upside-down-T-shaped region of the brass control plate allows users to change the time of day. Moving one's finger up and down the vertical groove of this control changes the level of fog. This control is disabled for the archival photographs, for which time-series and fog-series source stills were unavailable. Nevertheless, this control is active for the new photographs and Monet paintings, and permits users to draw comparisons between the actual appearance of the cathedral (given certain lighting conditions) and Monet's interpretation of the cathedral so lit.
- Dragging one's finger across the rectangular, bottom region of the brass plate allows users to change their point of view around the Rouen Cathedral.

8.3 The making of Rouen Revisited

This sections describes, from photography to model recovery to rendering, how the techniques and software presented in this thesis were used to create the imagery for the Rouen Revisited art installation.

8.3.1 Taking the pictures

The pictures of the cathedral used to create the exhibit were taken from three positions in front of the cathedral over a period of three days, January 14-16, 1996. For reference, we named the three positions alpha, beta, and gamma. The alpha position was located to the right of the front of the cathedral, near the corner of the Rouen tourist office, which is the building from which Monet painted many of the paintings. The beta position was located directly across the square from the cathedral in front of a clothing store. Because of the location of the store, it was impossible to fit the entire cathedral into a single photograph with the lenses we had available. Consequently, we took two overlapping photographs for each beta shot. The gamma position was located to the left of the front of the cathedral, just under the awning of a department store. Photos from the alpha, beta, and gamma positions are shown in Fig. 8.3.

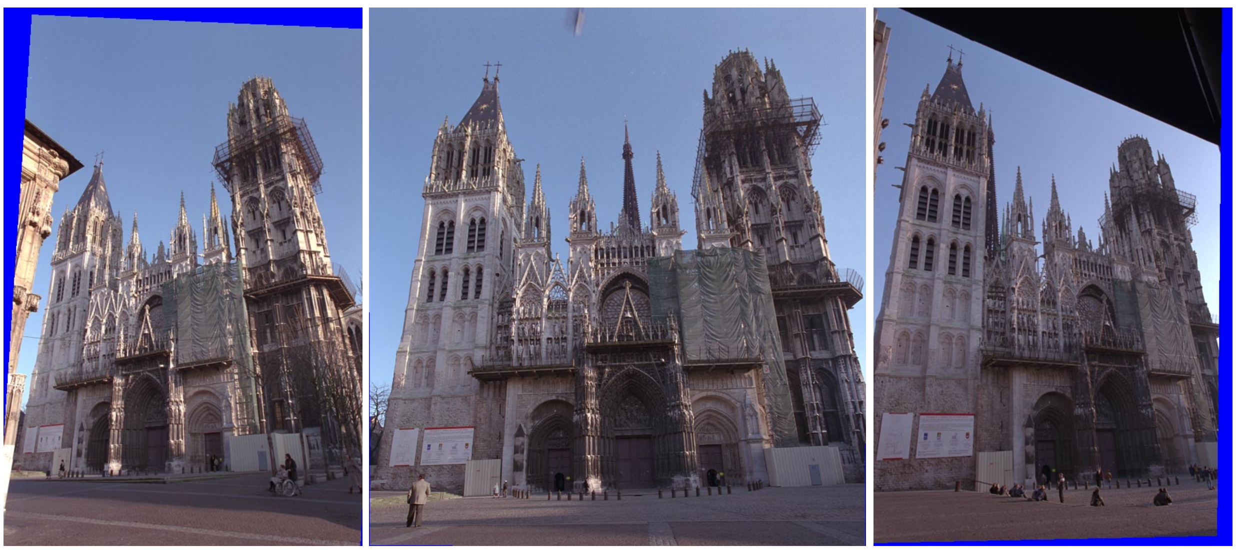


Figure 8.3: Three contemporary photographs of the Rouen Cathedral from the alpha, beta, and gamma positions. The beta photograph was assembled from two separate photographs of the bottom and top of the cathedral.

The photos were all taken with a Canon EOS Elan 35mm camera with a 24mm Canon lens on Kodak Gold 100 ASA print film. Since all of the photography was shot with a single camera, we had to shuttle the camera and tripod between the alpha, beta, and gamma positions for the duration of our stay. To guarantee that the camera was placed in the same location for each shot, we marked the position of the tripod feet on the pavement at each location using a black magic marker. This allowed us to reproduce the three camera positions to within an inch during the three days of shooting. However, since the camera needed to be rotated differently in each of the locations, there was no way to precisely reproduce the rotation of the camera for shots taken in the same location. As a result, the images had to be digitally rotated into consistent positions later.

Because the photographs were taken under greatly varying lighting conditions, it was impossible to properly record each photograph with the same amount of exposure. We took all the photographs with the same aperture since the aperture setting can affect the sharpness of the image and the amount of vignetting. We chose an aperture of f/8 since that is where the lens was tested to produce the sharpest images with the least vignetting.

Since the aperture setting was fixed, we achieved different exposure amounts by varying the shutter speed. The exposure was determined by manually spot-metering a particular area on the cathedral façade before each exposure. When possible, the same exposure was used for each triple of photographs from the alpha, beta, and gamma positions. The exposure times ranged from six seconds for the darkest conditions to $\frac{1}{60}$ of a second for the brightest. For each photograph taken, the time of day and length of exposure was recorded.

The lens was left focussed at infinity for every photograph taken in order to ensure a consistent focal length across the series of photographs. Most lenses, and the Canon 24mm lens we were using in particular, change significantly in focal length depending on the distance at which they are focussed.

The film was taken back to the United States, developed, and digitized using the Kodak PhotoCD process. Normally, the PhotoCD scanner applies a “Scene Balance Adjustment”, or SBA, during the scanning of each photograph. This SBA algorithm takes the scanned 12-bit-per-channel data, analyzes the overall brightness, contrast, and color balance in each image, and then applies a customized corrective mapping to the data before writing it to the PhotoCD in 8-bit-per-channel format. Since we had taken great care to capture the photographs with as consistent an exposure as possible, we ensured that the SBA function was disabled when scanning the film from Rouen.

8.3.2 Mosaicing the Beta photographs

As mentioned, it was impossible to capture the entire cathedral from the beta position in one photograph due to the limited field of view of the camera lens. Instead, two horizontal pictures were taken of the cathedral for each beta shot: one of the doors and the tower bases, and one looking up at the tower tops and spires. While the two photographs could have been used separately in Façade, this would have required estimating an extra camera rotation for each beta pair based on the marked edges.

Instead, we digitally combined the photographs into a single image, as shown in Fig. 8.4. Four corresponding points were located semi-automatically in the two images, and the homography[1] that mapped the four points in the bottom image to the four points in the top image was computed.

8.3.3 Constructing the basic model

A basic model of the West façade of the Rouen Cathedral (see Fig. 8.5) was built from a particular set of alpha, beta, and gamma photographs (Fig. 8.3) using the Façade photogrammetric modeling system presented in Chapter 5. The three photos used to construct the model were referred

[1] A homography is a 3×3 8-degree-of-freedom projective transformation that operates on planar points. Homographies, which model the effects of projecting in perspective a set of coplanar points onto an arbitrary plane, preserve the straightness of lines.

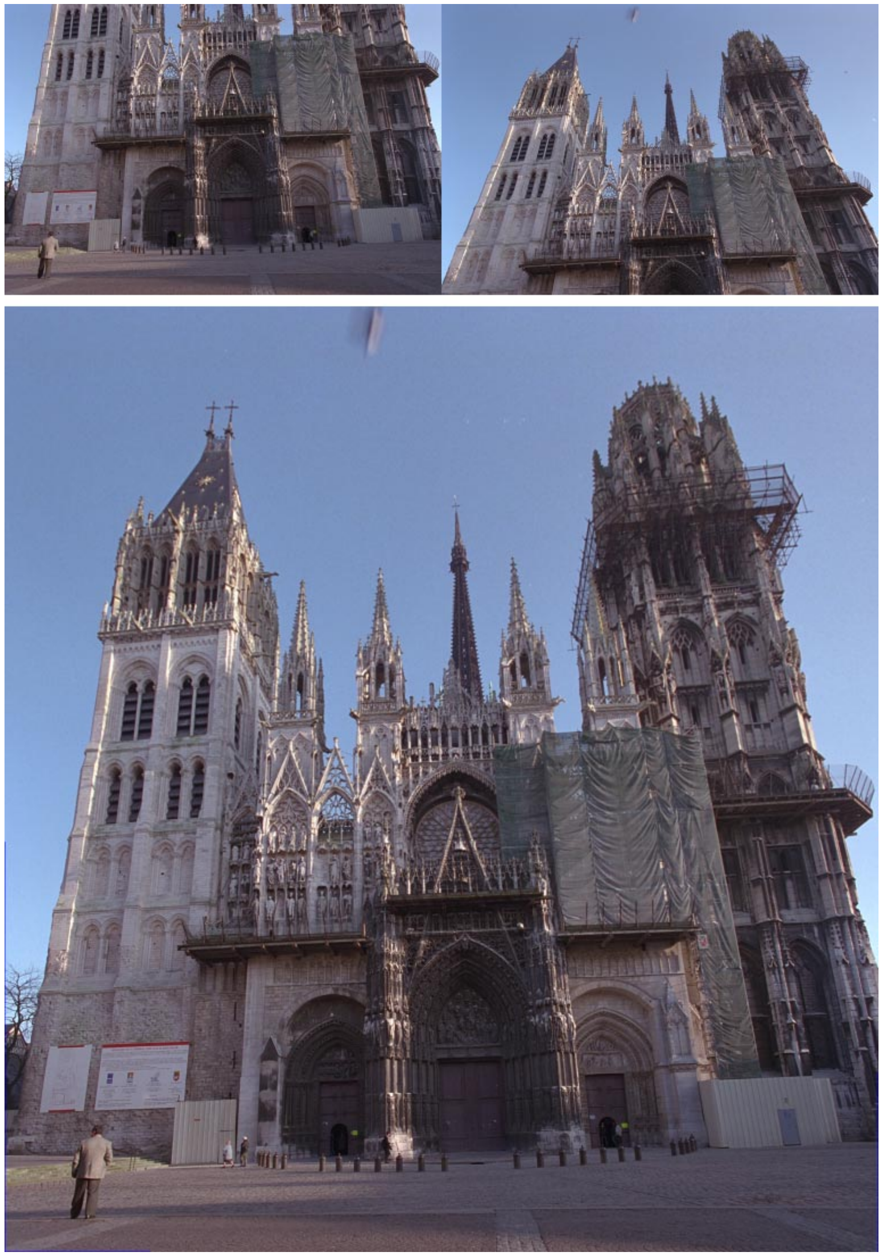
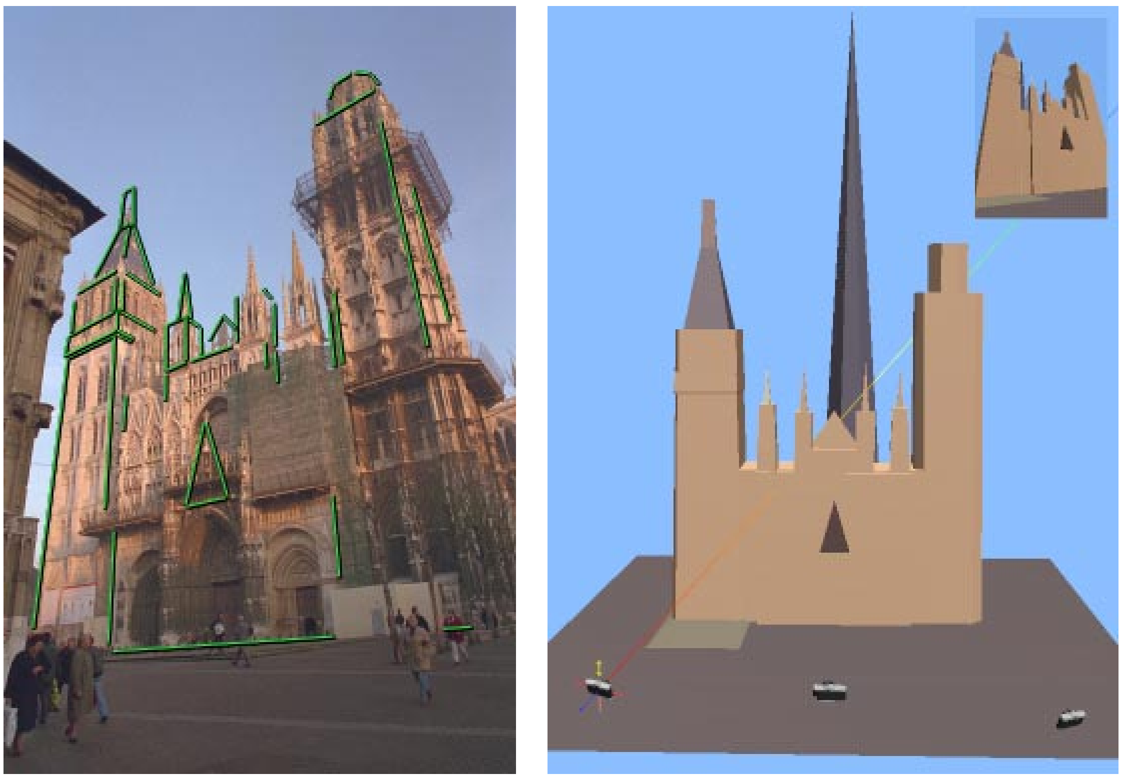


Figure 8.4: Assembling two photographs taken in front of the cathedral into a synthetic image with a wider field of view. The top two images show real photographs, taken from the same location, of the bottom and top portions of the Rouen cathedral. Because of buildings in front of the cathedral, it was impossible to stand further back and photograph the entire cathedral in one shot. The bottom image shows the results of mosaicing these two photographs into a single, wider-angle image of the cathedral. Because of perspective effects, it was necessary to apply a homography to the upper image before compositing. This homography was determined from four manually located corresponding points in the two images. The images are cross-faded in the region of overlap to ensure there is no visible seam in the composite image.



(a) (b)

Figure 8.5: Reconstruction of the West façade of the Rouen Cathedral for the *Rouen Revisited* installation in the SIGGRAPH'96 Art Show. **(a)** One of three original photographs, with marked edges indicated. **(b)** Model recovered from the three photographs. The recovered alpha, beta, and gamma camera positions are indicated right to left in front of the cathedral.

to as the *master set*. The photographs were first undistorted according to the radial distortion information for the camera lens, and the calibrated intrinsic parameters were entered into the modeling system. Polyhedral elements were photogrammetrically placed to model the Tour St. Romain, the older Tour de Beurre[2], the four front steeples, and the main steeple at the center of the cathedral. The geometry of the entire front entrance of the cathedral, including the three doors, the columns of saints, the central triangle, and the recessed Rose window, was modeled as a single flat polygon. This intricate geometry was left to be recovered by the model-based stereo algorithm.

8.4 Recovering additional detail with model-based stereo

With the basic model assembled, the finer geometric detail of the cathedral was recovered using the model-based stereo algorithm described in Chapter 7. Stereo was run four times on the master set of images: between alpha and beta; beta and alpha; beta and gamma; and gamma and beta (see Fig. 8.6). To achieve the best possible renderings for the art installation, the disparity maps were post-processed using a nonlinear smoothing process and retouched using image editing software to remove the remaining stereo errors. All four disparity maps were converted to depth maps, and the two depth maps for the beta image were averaged into a single depth map.

8.4.1 Generating surface meshes

In contrast to the rendering method presented at the end of Chapter 7, the three depth maps were triangulated into 3D surface meshes for the scene. Renderings would later be created by projecting the images onto these surface meshes, rather than using the depth maps to splat pixels into

[2]Literally, "Tower of Butter". The construction of the South tower was largely financed by the sale of indulgences that allowed people to eat butter during Lent.

novel views. As a result, the rendering could be done more efficiently and with better image resampling.

Note that the three surface meshes for the three camera positions were not merged into a single, consistent geometry for the cathedral. Instead, renderings were produced by projecting each image onto its own opinion of the scene geometry, and then optically compositing these image using view-dependent texture mapping. It would be an interesting exercise to generate comparison renderings using a single model geometry, assembled by merging the various depth maps using a technique such as [7], for all the images.

8.4.2 Rectifying the series of images

Since the camera rotation was not precisely the same for successive pictures taken from the same position, the cathedral did not appear at the same places in each of the photographs. This created a problem since all the photographs from the different times of day would need to be projected onto the model in the same manner, but camera rotation information had only been recovered for the master set of photographs used to construct the model. To solve this problem, we applied a homography based on four point correspondences to each of the remaining photographs to virtually align them with the master set. The homography, which is more general than necessary to correct for a rotation, also compensated for the translation of the film during the scanning process. Fig. 8.7 shows the complete set of photographs taken from the alpha location, all registered to each other. Fig. 8.8 shows a collage of the assembled and registered beta photographs. Both the assembly and registration of the beta images were accomplished in the same operation by composing the two homographies, eliminating an extra image resampling step.

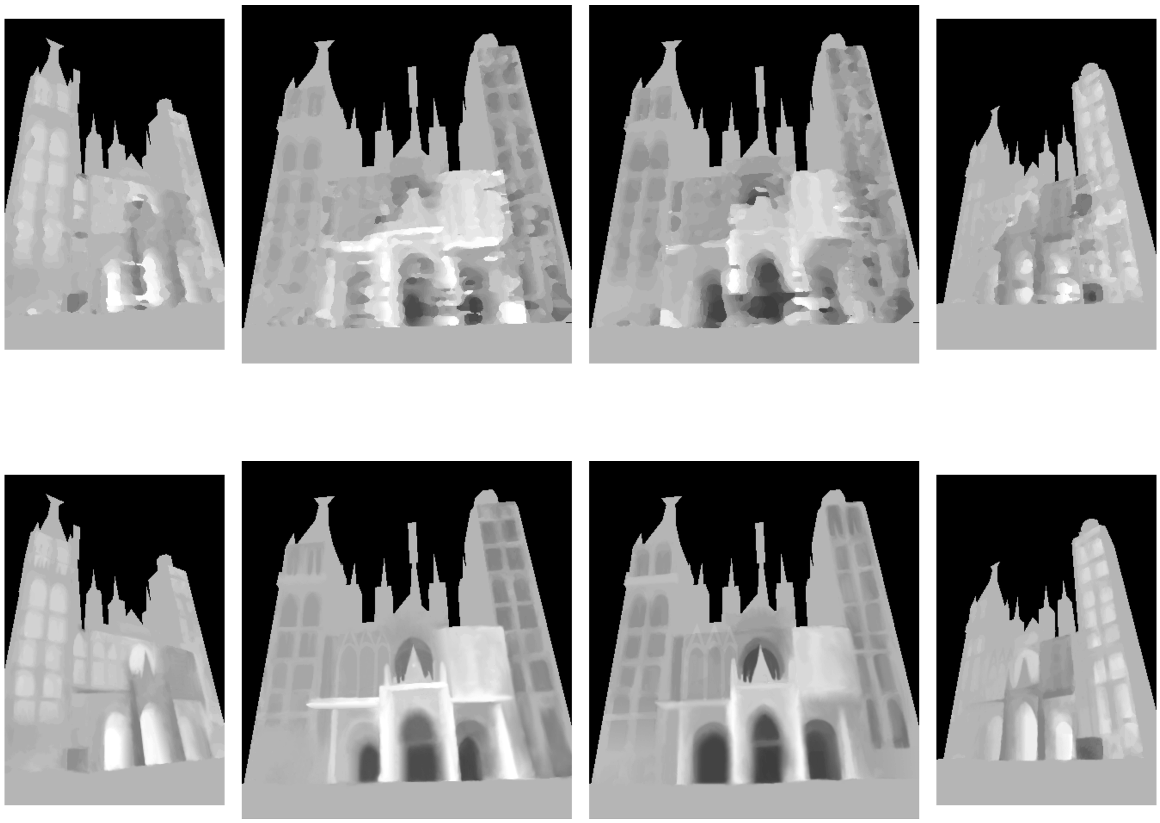


Figure 8.6: **Top:** The top row of images shows raw disparity maps computed for the cathedral façade by the model-based stereo algorithm. The left and right disparity maps are keyed to the gamma and alpha images using the beta image as the offset. The center two images show two different disparity maps computed for the beta image, using the alpha and gamma images as the offsets. Later, these two disparity maps would be merged to create a single depth map for the beta image. In all four images, disparities were calculated only for the front faces of the towers and the front entrance area; the rest of the cathedral was assumed to lie on the model. **Bottom:** The bottom row of images shows the disparity maps after being processed by the anisotropic diffusion algorithm and having their remaining stereo mismatches corrected by hand in an image editing program. Note that these are model-based disparity maps, so the disparity values indicated are relative to the model and are not directly related to depth. In these images, a middle gray color indicates that the scene is coincident with the polyhedral model, and brighter or darker colors indicate pixels which are recessed or protrude from the model.

Figure 8.7: This series of photographs shows the entire set of photographs taken from the alpha camera location over the period of January 14-16, 1996. The images have all been digitally rectified so that cameras all have the same rotation as well as position. As a result, any one of these time-of-day photographs can be projected onto the model using the camera pose information from the master alpha image. The beta and gamma series of images were similarly processed.

Figure 8.8: This set of photographs shows a sampling of digitally mosaiced photographs taken from the beta camera location. Note that fortuitous weather and vigilant photography allowed the cathedral to be photographed in about every condition evoked in Monet's paintings, from dawn to dusk in sun and clouds and even fog.

8.5 Recovering a model from the old photographs

Since the façade of the Rouen cathedral has undergone numerous changes since the time of Monet, it was necessary to recover a model of the cathedral façade as it stood at when the paintings were made. This model would be used to generate novel historic views, as well as to serve as the projection surface for Monet's paintings. We obtained this model from actual historic photographs of the cathedral, which necessitated working with uncalibrated images. This section describes how the antique images were calibrated and how the historic geometry was recovered.

8.5.1 Calibrating the old photographs

The three historic photographs used to model the nineteenth-century geometry of the cathedral façade are shown in Fig. 8.9. The three photos, purchased at an antique postcard shop during the trip to Rouen, were taken from approximately from the alpha, beta, and gamma positions. However, they appear to be taken from the second or third stories of the buildings across the street rather than at ground level.

To project these photographs onto the model, it was necessary to solve for their camera positions as well as their cameras' intrinsic parameters. In these images, vertical lines in the world remain vertical in the images, although the horizon does not pass through the center of the image. This indicates that the images were taken with bellows cameras to move the center of projection far from the center of the image.

The method used to calibrate the intrinsic parameters of our present-day camera required taking pictures of the calibration object in Fig. 4.8. For the historic pictures, it was not possible to take pictures of this calibration object, but it *was* possible to obtain photographs of a different object of known 3D structure, namely, the cathedral itself. The 3D geometry of the cathedral recovered from the present-day photographs was used to calibrate the historic cameras in the following manner:

The three historic photographs were loaded into the Façade modeling system and initialized with estimates of their intrinsic camera information. These estimates were obtained from observing vanishing points and applying the techniques described in Sec. 4.6. Then, the model recovered from the present-day photographs was loaded in, and all of the free parameters of the model were set to be held constant. Then, correspondences were made between the model and edges in the historic photographs, and the nonlinear optimization was used to recover both the intrinsic and extrinsic camera parameters for the historic photographs. The historic photographs were assumed to have no radial distortion effects.

8.5.2 Generating the historic geometry

With the camera geometries recovered, it was now possible to project the historic photographs onto the model. However, this model was recovered from photographs of the cathedral as it stands in the present day, and as such has minor differences from the geometry of the cathedral as captured in the historic photographs (see Fig. 8.10b). The historic geometry of the cathedral was recovered by explicitly removing the two inner steeples from the model and updating the disparity maps in Fig. 8.6 using a combination of model-based stereo and hand-editing techniques. The surface meshes generated from the historic geometry would be used for projecting both the old photographs and the Monet paintings into 3D.

8.6 Registering the Monet Paintings

To project the Monet paintings onto the model, it was once again necessary to recover the intrinsic and extrinsic parameters of uncalibrated images. But this time the uncalibrated images were not generated by an actual camera; they were generated by an impressionist painter. Surprisingly, it was possible to use the same techniques to calibrate Monet's artistic perspective as we used to calibrate the historic photographs.

8.6.1 Cataloging the paintings by point of view

For the paintings, it was necessary to estimate camera parameters for a series of thirty paintings rather than just three photographs. Fortunately, careful observation of the paintings revealed that only four distinct viewpoints were used in painting the cathedral. The paintings were cataloged according to their viewpoints and a representative painting was chosen for each set of paintings. The other paintings from the same viewpoint were then rectified to coincide with the representative painting.

8.6.2 Solving for Monet's position and intrinsic parameters

For each of the representative paintings, lines were marked on the painting using the Façade modeling system and these lines were corresponded to edges in the historic model of the cathedral. The triangular structure surrounding the clock was a particularly useful source of correspondences. Then, as in the case of the historic photographs, the reconstruction algorithm was used to solve for Monet's eye position and the intrinsic perspective parameters he used while painting.

The recovered positions from which Monet painted were consistent with the historical

New Photographs							
(24 times of day	+	6 levels of fog)	$\times$	200 viewpoints	=	6000 renderings
Historic Photographs							
1 time of day			$\times$	200 viewpoints	=	200 renderings
Monet Paintings							
(24 times of day	+	6 levels of fog)	$\times$	200 viewpoints	=	6000 renderings
Total						12,200 renderings

Table 8.1: Summary of renderings produced for the *Rouen Revisited* art installation.

record, indicating that he painted from the second stories of three different apartment buildings in front of the cathedral. And despite the impressionistic nature of his paintings, the accuracy of the intrinsic calibration showed that Monet's use of perspective was surprisingly accurate. The error of the model projected into the images using the recovered camera geometry was, on average, within a few percent of the dimensions of the paintings.

8.6.3 Rendering with view-dependent texture mapping

To create the final renderings, the original images – the new photographs, the historic photographs, and the Monet paintings – were projected onto the recovered geometry of the cathedral. For the exhibit, the images were pre-rendered from two hundred different points of view. The virtual viewpoints were arranged in a curved lattice twenty views across (from a little beyond the alpha camera position to a little beyond the gamma position) and ten views tall (from ground level to approximately four stories up.) In total, there were 12,200 renderings generated for the exhibit:

The rendering process was slightly different for the three classes of images:

Rerendering the new photographs Synthetic views of the cathedral as it stands in 1996 were made by projecting the new images onto the surface meshes recovered by the model-based stereo algorithm. A basic implementation of view-dependent texture mapping was used to composite the projections of the alpha, beta, and gamma images. For the viewpoints to the right of the center of the cathedral, a linear combination of the alpha and beta images was used. Likewise, for the viewpoints to the left of the center, a linear combination of the gamma and beta images was used. The relative contributions for any given horizontal position are shown in Fig. 8.13. Note that this view-dependent texture mapping method is an approximation to the method presented in Chapter 6 in that the same blending amounts are used for every pixel in a given rendering.

As in the view-dependent texture mapping method described in Chapter 6, surfaces not seen in one image were automatically filled in using the other available image, and regions not seen in any image were filled in using the hole-filling algorithm of Sec. 6.5.3. Fig. 8.10a shows a synthetic view of the modern-day cathedral.

Each point of view was actually rendered thirty different times using sets of alpha, beta, and gamma photographs taken at different times of day. To achieve more efficient rendering, this viewpoint coherence was exploited by pre-computing the alpha, beta, and gamma image coordinates needed to map each pixel of each novel viewpoint.

Rerendering the historic photographs The historic photographs were rendered in essentially the same manner as the new photographs, with the slight difference that the images were projected onto the the historic rather than the current geometry of the cathedral. Although the three available historic photographs were taken in noticeably different lighting conditions, the smooth fading achieved through the view-dependent texture mapping method prevented the differences from being distracting. Figs. 8.10b and 8.11 show renderings of the historic cathedral.

Rerendering the Monet paintings The synthetic views of the Monet paintings were generated by projecting the paintings from their recovered viewpoint positions onto the historic geometry of the cathedral. Each painting was projected onto the surface mesh most appropriate for the painting's viewpoint; the result that just one of the paintings was projected onto the beta image surface mesh, while all the others, painted somewhat to the right of the façade, were projected onto the alpha image mesh. No effort was made to extend Monet's brush strokes beyond the original frames of his paintings[3]; instead, the parts of the cathedral not featured in the paintings were in black rendered as the polyhedral model. The black model geometry served to situate the paintings within the context of the architecture.

8.6.4 Signing the work

Lastly, one of the renderings, shown in Fig. 8.14, was used by the artists to sign their work.

[3]Although it would have changed the artistic mission of the installation considerably, at one point we considered using the techniques presented in [20] to model Monet's painting style and rendering the entire façade in brush strokes. In fact, at the same SIGGRAPH conference where *Rouen Revisited* showed, a new method was presented for rendering three-dimensional scenes in a painterly style [28]. Coincidentally, the example renderings from the paper were modeled after another of Monet's famous series of paintings: *Grainstacks*.

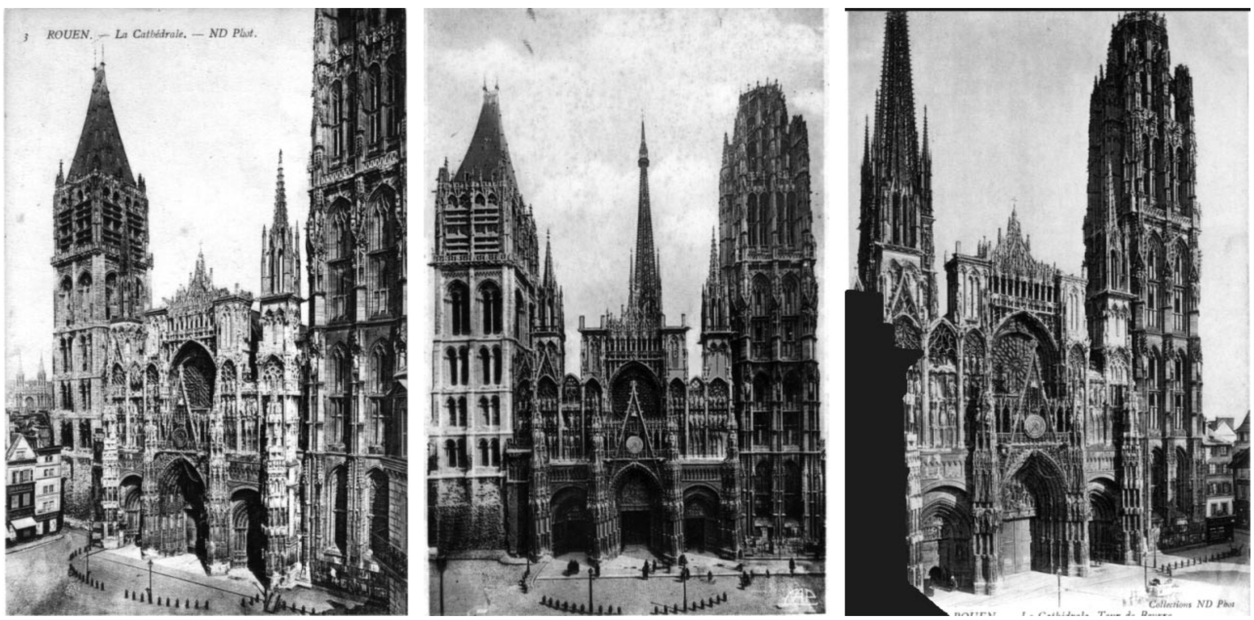
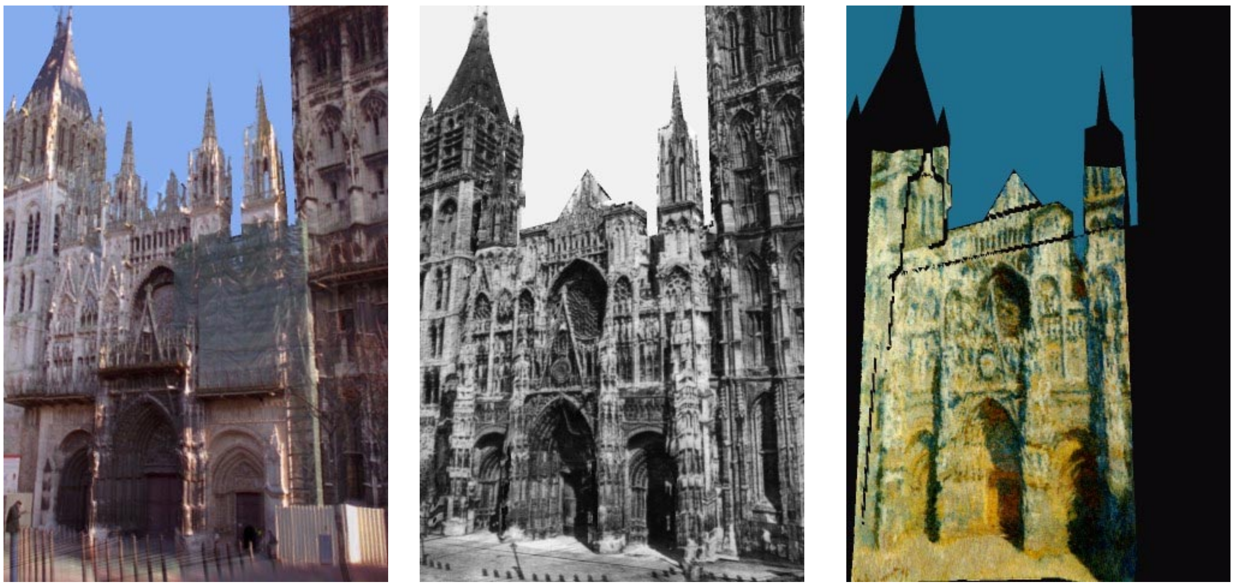


Figure 8.9: Three historic photographs of the Rouen Cathedral showing the West façade in the form seen by Monet during his time in Rouen from 1892-1894. The photo on the right has had an interposed building masked out at left.



(a) Synthetic view, 1996 (b) Synthetic view, 1896 (c) Synthetic view, Monet painting

Figure 8.10: Synthetic views of the Rouen cathedral from the *Rouen Revisited* art installation. **(a)** A synthetic view created from photographs taken in January, 1996. **(b)** A synthetic view created from historic postcards showing the cathedral around the time Monet executed his series of paintings of the cathedral façade from 1892-1894. Note the architectural differences from 1996: the two missing spires; the clock above the door; the lack of scaffolding, the slats at the top of the St. Romain tower, and the columns of saints flanking the left and right doors which were removed around 1914. **(c)** A synthetic view of a Monet painting (Fig. 8.1) obtained by projecting the painting onto the 1894 geometry of the cathedral. The black regions on the periphery are the areas beyond the frame of the original painting. Monet's original viewpoint was solved for from features in the painting using the Façade photogrammetric modeling system. Note that all three pictures are rendered from exactly the same viewpoint.

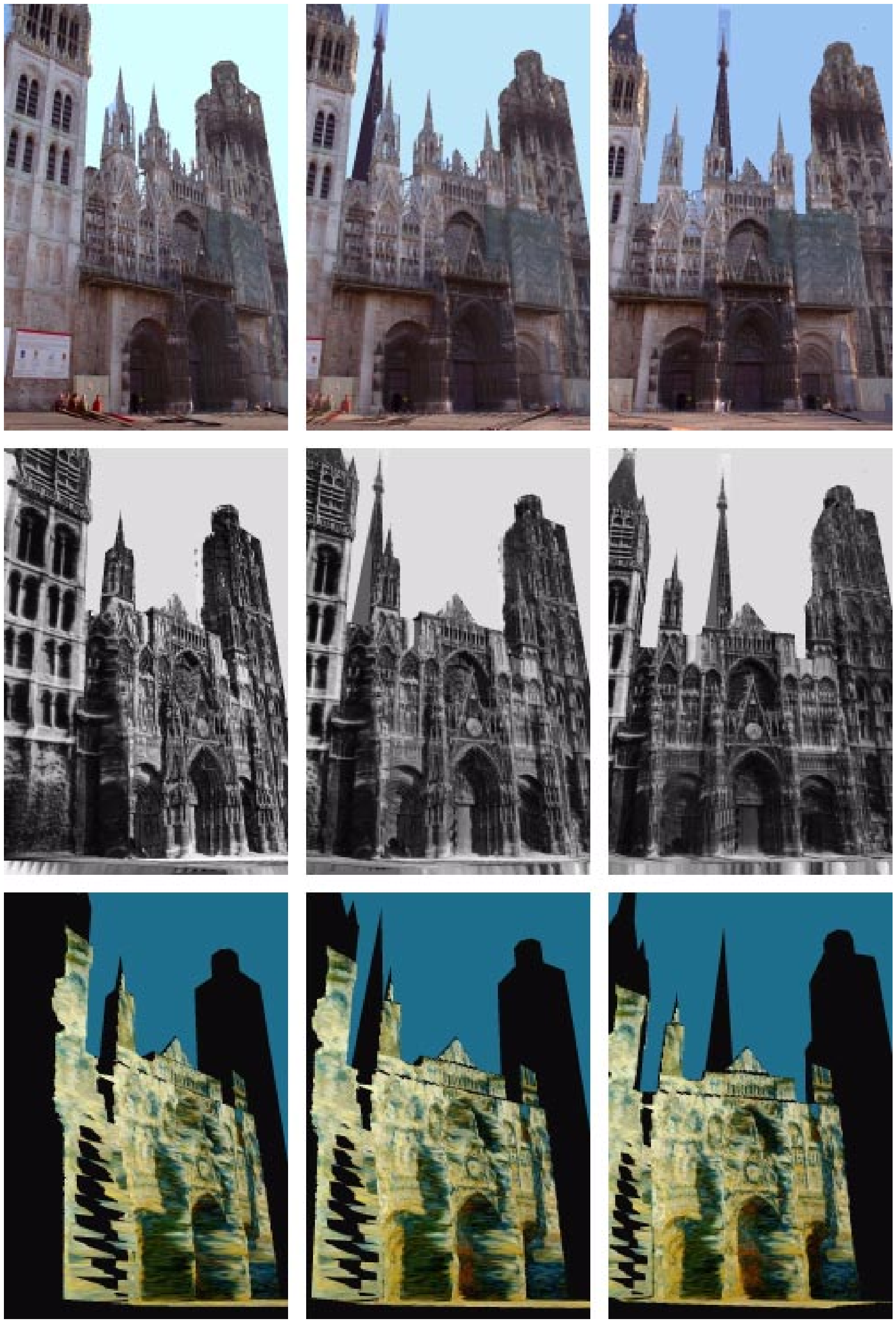


Figure 8.11: This figure, which is continued as Fig. 8.12 on the next page, shows six synthetic renderings from the left to the right of the cathedral. Top to bottom we see the cathedral as it stands today, as it stood one hundred years ago, and as it would look with a Monet painting projected onto it in three dimensions.

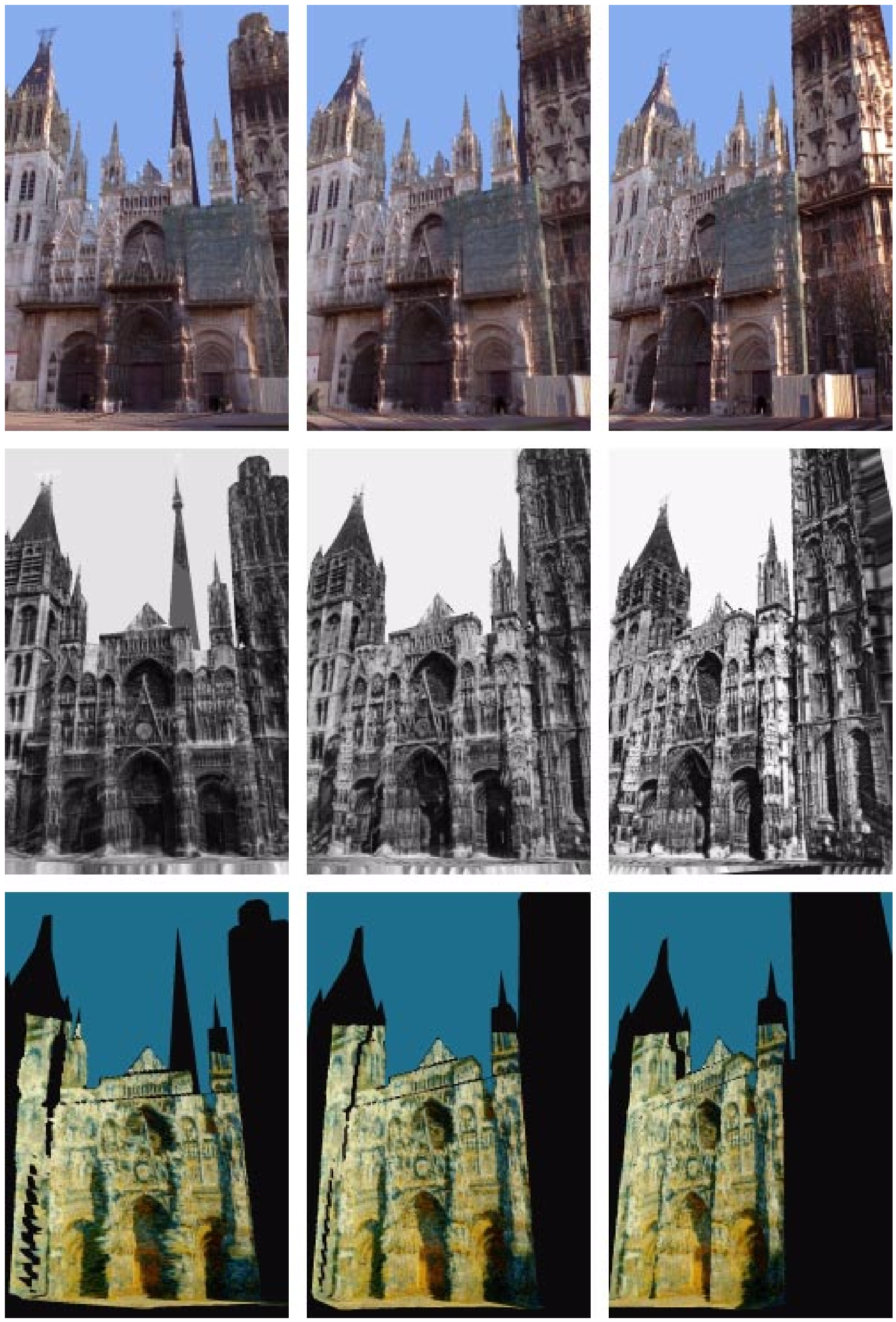


Figure 8.12: This is a continuation of Fig. 8.11 from the previous page.

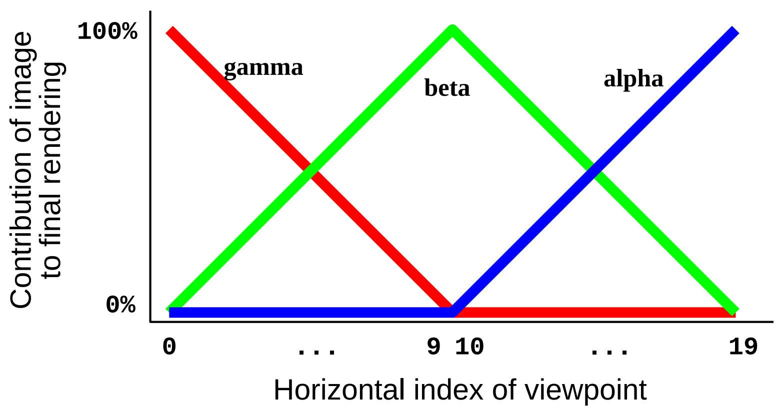


Figure 8.13: This graph shows the explicit form of view-dependent texture mapping used to create the renderings for Rouen Revisited. The two hundred viewpoints were arranged in front of the cathedral in a lattice twenty sites wide and ten tall. The bottom axis is the horizontal index of a given viewpoint. The relative blending amounts used for the alpha, beta, and gamma image rerenderings are indicated for each index. Note that only two of the alpha, beta, and gamma images were composited to produce any given rendering, which increased the rendering efficiency.

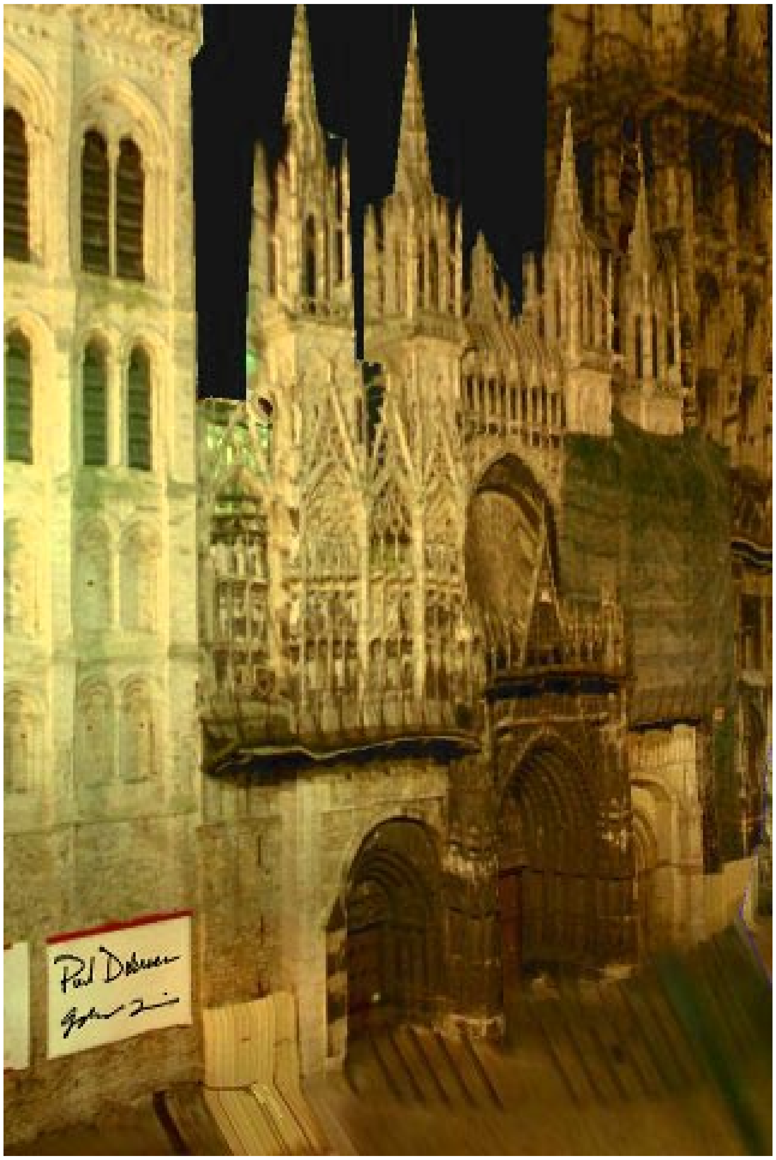


Figure 8.14: In this rendering, one of over twelve thousand in the exhibit, the artists replaced the text of the “under reconstruction” placard to sign their work.