4D Cities: Analyzing, Visualizing, and Interacting with Historical Urban Photo Collections

Grant Schindler, Frank Dellaert
School of Interactive Computing, Georgia Institute of Technology, Atlanta, Georgia
Email: schindler@gatech.edu, dellaert@cc.gatech.edu

***Abstract*— Vast collections of historical photographs are being digitally archived and placed online, providing an objective record of the last two centuries that remains largely untapped. In this work, we propose that time-varying 3D models can pull together and index large collections of images while also serving as a tool of historical discovery, revealing new information about the locations, dates, and contents of historical images. In particular, we use computer vision techniques to tie together large sets of historical photographs of a given city into a consistent 4D model of the city: a 3D model with time as an additional dimension.**

***Index Terms*— 3d reconstruction, structure from motion, historical photographs, image-based rendering**

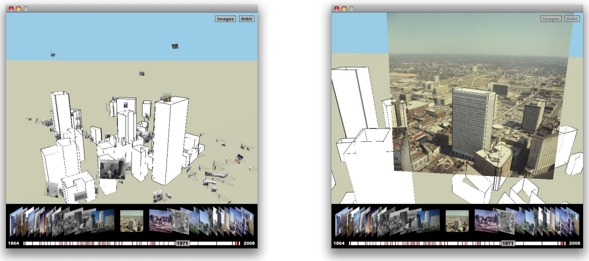

Fig. 1. 4D City Model. A bird's eye view of Atlanta in 1971 (left) and the same model from the viewpoint of a selected 1971 photograph (right). We show that 4D city models serve as an effective means of organizing historical photographs and providing context, both spatially and temporally.

I. Introduction

There is a growing need to organize the vast number of *historical* urban photographs being digitized and put online. We propose that 4D city models – time-varying 3D models of cities – can serve to: (1) organize photo collections, (2) contextualize individual photographs, (3) visualize the past, and (4) uncover historical details. Through these applications, we demonstrate the potential of 4D city models to contribute to the goal of understanding and preserving our world's urban centers and their architectural heritage.

Organizing Photo Collections: Historical photographs are currently distributed across a wide variety of internet locations. If we consider a city like Atlanta, Georgia which played an important role in the American Civil War, such photos can be found through the Atlanta History Center, the Library of Congress, the Digital Library of Georgia, and Flickr, just to name a few. Currently, these collections are completely separate entities, with no efficient means of finding images of the same building captured at different times, or images taken from different locations at roughly the same time, for example. The collections do not know about each other, and it is difficult for a person to form a solid understanding of a given city in a given decade without spending a significant amount of time in each collection. By registering images to a 4D city model using computer vision techniques (Figure 1), we show how the barriers between image collections fall and one can easily get a birds-eye view of all the images from a given era, or transition between two historical views of the same building that come from different sources.

Contextualizing Photographs and Visualizing the Past: When viewing a photograph of the past, it can be difficult to get a sense for where the image was *really* taken even if one is familiar with the equivalent modern day location. In some cases, the entire city has changed beyond recognition. In other cases, it can be difficult to decide if the photo is looking north or south on a given street. We show that by registering the image to a 4D model, the context of the photograph becomes clear, both in space and time. Buildings that were not even in the original photograph become visible around the edges of the image. The viewer can look down to discover that the photographer was standing on a rooftop not visible in the image. In the case of old photographs sharing no common scenery with the present, one can even make visible the 3D models of modern buildings to get a sense of where the photo is positioned with respect to modern day structures.

Uncovering Historical Details: We also show that 4D city models can serve as a tool of historical discovery. By ensuring that every photo is registered to the same model in a mathematically consistent way, we can reveal information about the precise locations, dates, and contents of photographs that would have been unrecoverable without such a 4D model. As an example, we have been able to determine the precise latitude and longitude of a set of 1864 images captured by Civil War photographer

Manuscript received October 15, 2011; revised February 29, 2012; accepted April 5, 2012.

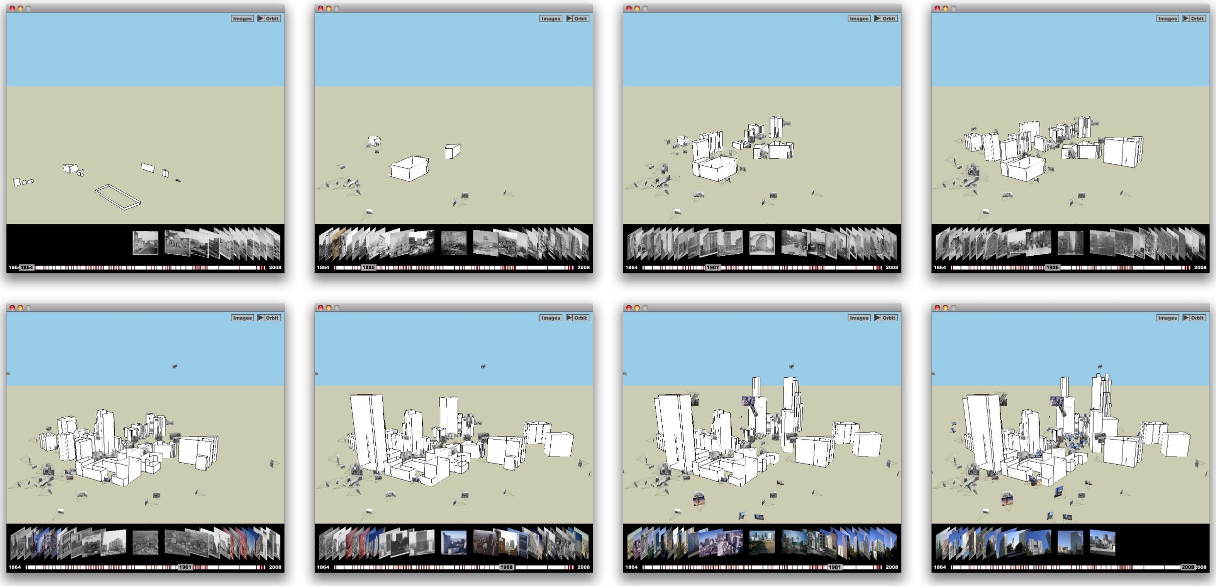

Fig. 2. Time-Varying 3D Model. A user can drag a time-slider to see the 3D model of the city at any point in time. As the user does so, buildings rise and fall and images flicker in and out of existence to reflect the changes that have taken place over time. Here we see a model of Atlanta depicted in successive time periods from 1864 to 2008.

George Barnard, in spite of the fact that no structures (either natural or man-made) pictured in those 1864 photos still exist today. However, because at each stage of history, there has been overlap between the structures of one decade and the next, we are able to register these 1864 photographs to the same coordinate frame of our modern photographs which are easily referenced to a global coordinate system via maps or GPS.

In the remainder of this work, we define 4D city models (Section III), discuss how to construct them directly from photographs (Section IV), develop new visualization techniques using 4D city models (Section V), and define new methods of user interaction for such models (Section VI).

II. Related Work

We formulate the problem of 4D model construction in the context of a Structure from Motion (SfM) framework. Structure from motion is the process of recovering the 3D geometry of a scene (structure) as well as the internal and external camera parameters (motion) associated with a set of images of a scene. Structure from motion is a well-studied problem [7], [4].

Several recent methods have applied SfM to large internet image collections [14], [8], mostly to reconstruct famous landmarks and other well-photographed locations. However, all such approaches have ignored issues of changes over time.

In a related area, a number of recent approaches to large-scale urban modeling from images have produced impressive results [9], [1], [16], though none have yet dealt explicitly with time-varying structure. In [13], a historical Ansel Adams photograph is registered to a reconstructed model of Half Dome in Yosemite National Park, but there is no notion of time in this process – only the location of the image is recovered. Additionally, since we are dealing with historical photographs, approaches that rely on video [9], densely captured data [16], or additional sensors are not directly applicable to our problem.

Interestingly, recent work in civil engineering *has* begun to use SfM to track 3D changes in building sites over time and compare them to planned models of the same site [5].

Finally, [13] defined a number of novel user interaction mechanisms for navigating between photos of a reconstructed 3D scene. Our work expands upon this set of interactions to deal explicitly with photographs taken over large time scales and in which the structure of the scene has changed drastically over time.

III. 4D City Models

We define a 4D city model as a time-varying 3D model of a city. Formally, the model consists of a number m of 3D geometric objects $O = \{O_i | i \in 1 \dots m\}$, each with an associated time interval (a_i, b_i). The geometry of the scene O changes over time only due to objects (such as buildings) beginning and ceasing to exist, but the geometry itself never moves through space. Thus, a 4D city model could contain a 3D point that lasts for five minutes, or a polygonal model of a building that exists for 100 years. Note that we first define 4D city models without considering the types of algorithms we will use to build them. Later, we will see that such models can be built either manually or automatically.

Essential to the idea of 4D cities, in this work, is the concept of a set of photographs geometrically registered to the geometry of the scene, and taken at different points in time. Formally, the geometric objects in the scene O are observed in a set of n images $I = \{I_j | j \in 1 \ldots n\}$, where each image has an associated time t_j. For every image, we must also know the 3D pose of the camera and internal parameters such as focal length, in addition to the date and time at which the photograph was taken. Though it is true that time-varying 3D models can exist independently of any set of photographs, when we talk about a 4D city model in this work, we are assuming that such a set of photographs is present for two reasons. Firstly, we see one of the most important functions of a 4D city model as organizing the historical photographic record of a given city, and enabling new ways of understanding historical photographs in their spatial and temporal context. Second, we show in the next section that such a set of photographs has enough information to enable the construction of 4D city models from images alone.

IV. Constructing 4D City Models

Constructing a 4D city model, at its core, involves identifying corresponding points across multiple images in a Structure from Motion (SfM) framework. 4D city models, like the ones pictured above, can be constructed either automatically or interactively with a user guiding the process. In [10], we introduced a completely automated method of constructing 4D city models, including the estimation of both geometric and temporal information about the scene. Though modern computer vision techniques are capable of performing this task automatically for certain image collections, there are distinct advantages to allowing users to manually specify corresponding points with the help of a user interface designed for this task.

Primary among the advantages of interactively constructing 4D models is that humans can identify corresponding points despite enormous changes in appearance that takes place over time. For the case of Atlanta, by using manual point correspondences we are able to produce a user-constructed 4D model spanning the dates 1864 to 2008 across 212 images (see Figures 1 and 2). In contrast, a method based on automated correspondences produces a 4D model of Atlanta spanning only 1956 to 1975 (across 102 images), a much shorter span of time, due to the inability to detect corresponding SIFT features across the entire database of images.

A second advantage to putting humans in the loop is that it allows the creation of simplified solid building geometry. Though in [10] we presented several methods of automatically segmenting and triangulating the point cloud resulting from automated SfM methods, the resulting building models can be incomplete and noisy, and may split or merge buildings incorrectly. Interactive modeling methods avoid this problem.

Therefore, we have created a 4D city construction tool which consists of an interface:

- to specify corresponding points between two images,
- to define a building by joining a series of points,
- and to specify a date for each image and a time interval for each building in the scene.

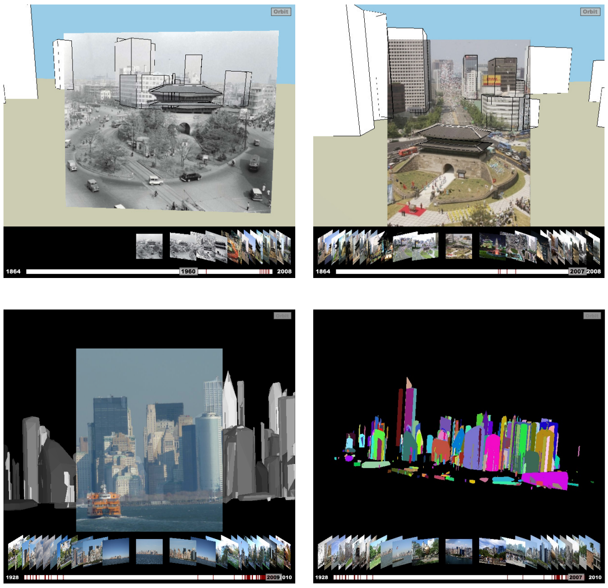

Fig. 3. 4D City Models. The 4D city model of Seoul, South Korea (top) was constructed using manual point correspondences, while the model of lower Manhattan (bottom) was constructed automatically using the method of [10]. In the top left image of Seoul, in order to visualize the changing skyline, we display the outlines of buildings which had not yet been constructed in 1960 when the photo was taken. The famed Namdaemun gate pictured here was burned down in a 2008 incident, which highlights the importance of preserving this historical imagery in an accessible way.

To recover camera parameters and 3D scene geometry, we use a custom bundle adjuster based on Levenberg-Marquardt within an automatic differentiation framework [6]. After specifying point correspondences, a user can choose to optimize structure parameters, optimize camera parameters, or optimize all parameters simultaneously using bundle adjustment. This is the only operation we require for 3D reconstruction, as we depend upon user input to initialize cameras and points to reasonable values.

Model construction begins with an initial pair of images in which the user specifies correspondences for a specific set of 4 points which define the origin, scale, and coordinate axes of the world. Additional prior terms are added to the bundle adjustment in order to constrain these points during subsequent optimization. In addition to these constrained points, the user specifies a height for the ground plane which is used in the interactive viewer. Knowledge of the direction of gravity and a ground plane enables buildings to be defined by simply specifying an ordered set of points along the roof of a building which are to be connected into a polygon and extruded to the ground. Such simple building models are useful during subsequent user interaction with a 4D model to determine when a user has clicked on a given building in any image.

Once the geometric parameters of the initial images and points have been solved, the user alternates between adding additional 3D points to the model (initialized by back-projection into existing images) and adding additional images (initialized with the pose of an existing camera in the reconstruction). This procedure enabled the creation of the models for Atlanta, Georgia and Seoul,

Fig. 4. Visualizing a 4D City Model. We juxtapose different eras in the same photograph, rendering buildings from the 20th century and inserting them into an 1864 photograph of Atlanta. Since we know the internal and external camera parameters for the original 1864 photograph (bottom left), we can render a 3D model of the city from the same viewpoint (bottom right), and pull textures for this 3D model from two other photographs taken in 1966 and 2008. As a result, we get context for the 1864 photograph that is lacking in the original photograph.

South Korea as depicted in Figure 3. The 4D model of Seoul, consisting of 29 buildings and 88 images, was built by a non-expert (a student unfamiliar with the city) and required roughly 10 hours of work.

V. Visualizing 4D City Models

There are several unique visualization techniques that become possible when we have a 4D city model consisting of images taken of the same scene (at different historical dates) registered to time-varying 3D geometry. We focus here on image-based rendering methods, which involve projecting the original images as textures onto the 3D geometry, as distinct from the real-time interactive visualization in the next section which employs textureless 3D models.

A. *Image-Based Rendering*

Often we see a historical image of city that has changed so much that it is difficult to tell exactly where, geographically, the photo was taken. We might be told that a photograph was taken looking North from a given intersection, but without any structures co-existing in the historical and modern day photographs, there is a lack of genuine understanding of the context of the photograph. We can overcome this problem by rendering modern-day buildings in precisely the location they would have

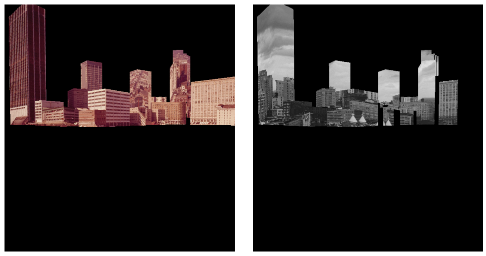

Fig. 5. Failure of Traditional Image-Based Rendering. If we had only static geometry for the city, rather than time-varying geometry, then traditional image-based rendering techniques would fail by projecting image background onto non-existent 3D geometry. By knowing a date for each image and a time-interval for each building, we avoid this problem.

Fig. 6. Animating a transition between two images. We use the known time-varying 3D geometry to morph between two different viewpoints and time-periods. *Photos provided by New York Public Library (left) and Tony Street (right).*

appeared had the historical photograph been taken years later.

In Figure 4, we juxtapose different eras in the same photograph, rendering buildings from the 20th century and inserting them into an 1864 photograph of Atlanta. Since we know the internal and external camera parameters for the original 1864 photograph, we can render a 3D model of the city from the same viewpoint, and pull textures for this 3D model from two other photographs taken in 1966 and 2008. As a result, we get context for the 1864 photograph that is lacking in the original photograph.

The time-varying nature of a 4D model requires a change to traditional image-based rendering techniques (which assume static geometry [3], [2]) to extract textures from the regions underlying the projections of building geometry only for those images which were taken during the period when each building existed according to the 4D model. If we had only static geometry for the city, rather than time-varying geometry, then traditional image-based rendering techniques would fail by projecting the image background (such as sky or background buildings) onto non-existent 3D geometry as in Figure 5. By knowing a date for each image and a time-interval for each building, we avoid this problem.

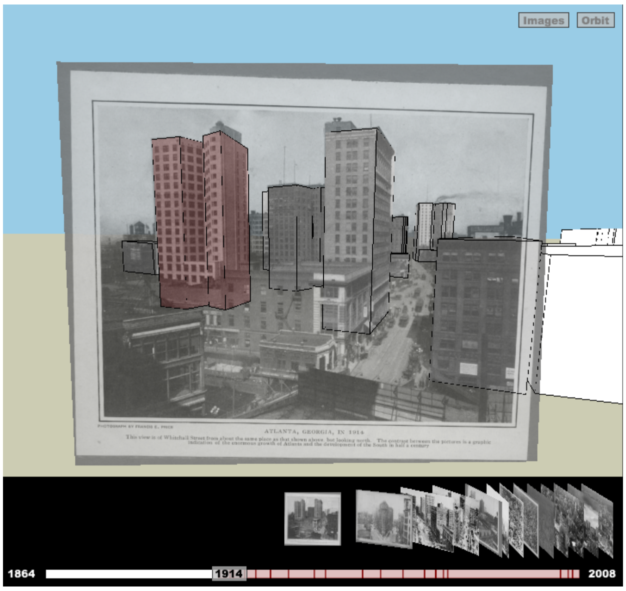

Fig. 7. The Fourth National Bank Building, Atlanta, 1914 (highlighted in red). This is the earliest image we have of this building.

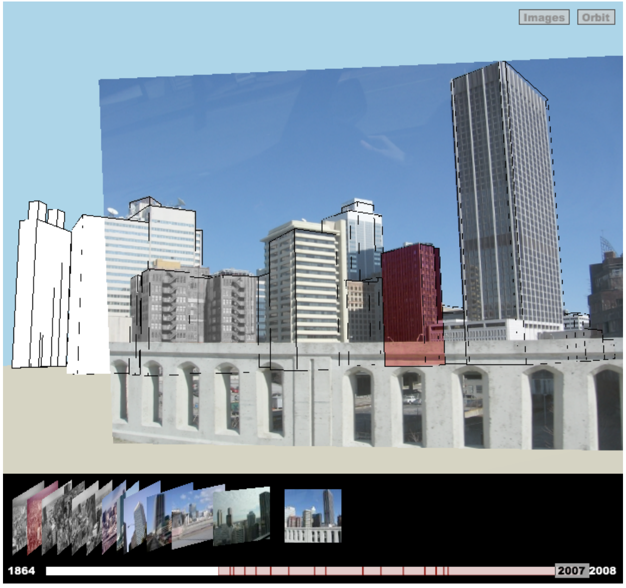

Fig. 8. The Metropolitan, Atlanta, 2007. A modern image showing the same building as the previous figure (formerly the Fourth National Bank Building), identifiable using the 4D model, despite large changes in appearance.

B. Animated Image Transitions

Another powerful tool to communicate a changing scene is an animated transition between two images taken at different historical times. Here, we use an image morphing [12], [15] method to transition between two images taken from a similar viewpoint. The known geometry of the scene is used to create a 2D mapping between pixels in the two images as a virtual camera transitions between the known viewpoints of the two original images. This morphing-based method avoids visual holes where no geometry exists (e.g. in the sky).

In Figure 6, we show a transition between a 1937 image of Lower Manhattan and one from 2001. Some of the buildings remain the same between the two images, while there are also a large number of new buildings that appear. The transition makes it clear which buildings are new, which buildings remain, and how the two viewpoints are related.

VI. Interacting with 4D City Models

Historical and modern images are currently dispersed across a wide variety of online sites, including Flickr, Picasa, the Library of Congress, and numerous smaller collections at various universities, historical societies, and other institutions such as The Atlanta History Center, The New York Public Library, and the Charles W. Cushman Photograph Collection at Indiana University. The goals of a 4D city viewer include:

- to bring together historical and modern photos from a variety of sources
- to place these photos in both their *spatial* and *temporal* context
- to allow a user to see how a whole city, a specific building, or a specific view changed over time

To enable this interaction we require precisely the type of representation outlined above: a set of images with known pose, calibration, and date, and a collection of buildings with known 3D geometry, date of construction, and (if applicable) date of demolition. A list of which buildings are observed in which images is also necessary to take full advantage of our 4D model interaction methods. For the purpose of interacting with a 4D model, as long as we have all these pieces of information, then how we acquire the model is not important (and any of the interactive or automatic methods described above may be used).

A. User Interface

Given a 4D city model, we define a number of ways for the user to interact with this model, which we outline briefly below. To see examples of the visual elements of the user interface described here, refer to Figures 7 through 14 for the best views of the interface elements. The reader may also visit http://4d-cities.cc.gatech.edu/atlanta/ to explore a 4D model of Atlanta using the interface described here.

Timeline: The primary novelty of this interface is a timeline which lets the user set the current time at which to view the model. Tick marks along the timeline indicate dates at which photographs exist in the model, and corresponding thumbnail images are arrayed along the timeline. As the user drags the time slider back and forth, both the 3D model of the city and the displayed images change to reflect the current date.

3D View: The user sees the entire 3D city model from an overhead viewpoint and is able to orbit around the city with a mouse. Along with 3D building models, this view also shows images floating in space at the position and orientation of the associated camera. The user may select any of these floating images, or the images along the timeline, to view the model from the viewpoint of any

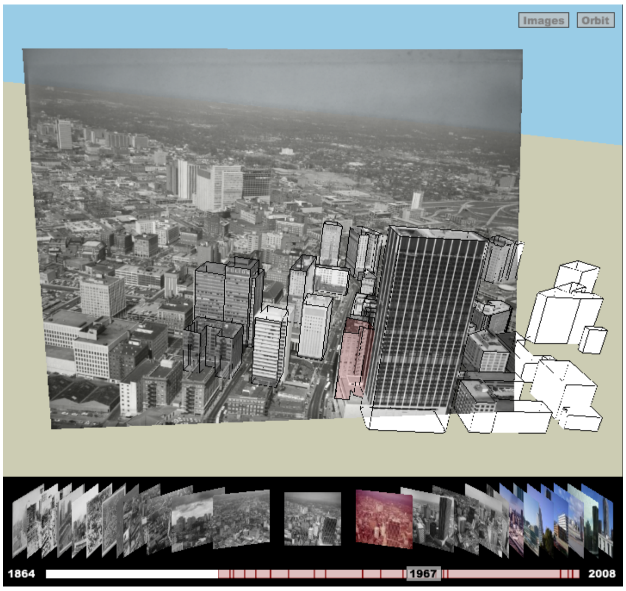

Fig. 9. The Fourth National Bank Building, Atlanta, 1967. The last image depicting the building with its original facade.

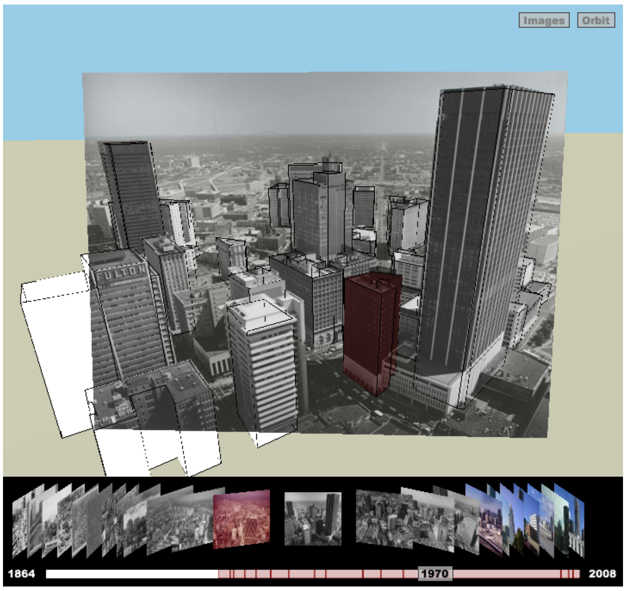

Fig. 10. The Fourth National Bank Building, Atlanta, 1970. The first image depicting the building with its new facade.

individual photograph. When a photo is selected, the date of the time slider changes to the date on which the image was captured, thus changing the displayed 3D buildings.

Image View: From the viewpoint of any image, the photograph itself is overlaid on the 3D model such that the buildings present in the image are clearly outlined by the corresponding 3D building models underneath. The user can rotate this viewpoint with the mouse to look around the scene and see which buildings were present at the time of the photograph, *despite the fact that they lie outside the field of view of the camera*. In addition, from the same image viewpoint, the user can drag the time slider long the timeline to show what the current viewpoint would have looked like in a different era.

Building Selection: The user may click on any building, whether from the orbiting 3D view or image view, in order to highlight it. When this happens, the images on the timeline and in the 3D view are filtered down to only those images that view the highlighted building. In addition, the area of the timeline between the beginning and end dates of the selected building are highlighted as well, and the set of remaining date tick marks on the timeline gives an indication of the periods from which we have images that observed this building.

Note that if any tick mark lies outside the highlighted region of the timeline for this building, we know there is an inconsistency between this image date and the time interval for the selected building. Similarly, if from the viewpoint and date of a specific image a building in the image is not being shown by the 3D model (or vice versa), we know there is a temporal inconsistency in the model as well. These two cases illustrate that, even without any automated dating mechanism, just relating all the images to a 4D city model and visualizing the result is a powerful tool for ensuring consistency between photographic dates and historical building records.

VII. Results

We have used the methods described above to construct a number of 4D city models. The model of Atlanta that consists of images from 1864 to 2008 (212 images) was constructed using manual point correspondences, while automated methods were able to construct a separate Atlanta model using 102 images from 1956 to 1975. In Figure 3, we also show a 4D city model of lower Manhattan, constructed completely automatically, as well as a model of Seoul, South Korea constructed with manual point correspondences. Quantitative evaluations of the 4D city model construction methods can be found in our previous work on the subject [11], [10]. Here, we focus on presenting qualitative results for the kind of historical discoveries enabled through the 4D city model framework we have described above.

A. Historical Discovery

To illustrate the claim that 4D city models may be used as a tool of historical discovery, here we provide two detailed examples of how such discoveries have been made by interacting with our system.

Buildings change both visually and physically over time, which can make it difficult to recognize the same building in both modern and historical photos. This is further complicated by the fact that the spatial context of the building may change as new, taller buildings are built around it. A 4D model makes it easy to find the same building in historical and modern photos simply by selecting the building, which becomes highlighted in all photos in which it is visible.

As an example, using a 4D model of Atlanta, we are able to discover that The Fourth National Bank Building (1914) and the Metropolitan (2007), two buildings which differ vastly in appearance, are in fact the same

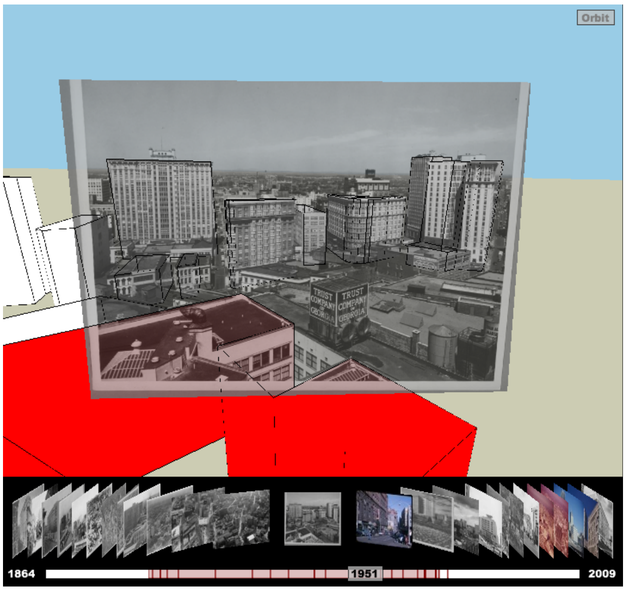

Fig. 11. Downtown Atlanta in 1951. It is unclear from this image exactly where the image was captured from. It appears to have been taken from the roof of the building highlighted in red. (See Figure 12.)

Fig. 12. Downtown Atlanta in 1951. The 4D model reveals that the image in Figure 11 was taken from the rooftop of a building, though not the rooftop visible in the image itself.

exact building. The Fourth National Bank Building in downtown Atlanta appears to have a light stone facade in the 1914 image depicted in Figure 7. Meanwhile, the Metropolitan, a building in downtown Atlanta in 2007, appears to have a black facade decorated with metal strips as shown in Figure 8. The 4D model tells us that, despite the change in appearance, this is the same building depicted in the two images.

Moreover, we can determine when this change occurred by quickly flipping through all images which observe the given building. The building remains highlighted in red as we flip through the images, and we are able to determine that the change occurred between 1967 and 1970. Figures 9 and 10 show the last image depicting the building with its original facade, and the first image depicting the building with its new facade.

Another type of discovery is determining the location from which a particular photograph was taken. Rather than just determining the GPS coordinate of an image, a 4D model can show us that a photographer was standing on top of another building, or looking out the window of another building, when a photo was captured. For example, we see in Figure 11 a photograph of the downtown Atlanta skyline in 1951. There is a rooftop visible in the image, and one might assume the photo was captured from this same rooftop. When we select the building in question (the old Equitable Building) and back out to a wider view, we see that this is not the case. In Figure 12, the 4D model reveals that the image was taken from the rooftop of a different building (the Hurt Building), not the rooftop visible in the image itself. In addition to the historical significance of finding out where a specific photographer stood over 50 years ago, it tells us that we could go take a photograph from this same viewpoint today because the Hurt Building still exists, despite the fact that the old Equitable Building no longer stands.

Figures 13 and 14 show an example of an image which one might assume was captured from a helicopter, but was in fact captured from the rooftop of the newly constructed State of Georgia Building, the tallest building in the Southeast United States at the time. For both examples above, there turn out to be other images captured from the same corners of the same rooftops years later, indicating that these locations were quite popular spots for photographing the skyline.

It is important to note that this type of analysis can be carried out by a user with no domain expertise, since all the relevant information is captured by the 4D model itself. However, we did work with archivists and researchers at the Atlanta History Center to acquire many of the historical images of Atlanta used in this work. Though we have performed no formal user studies, these experts have been enthusiastic about the results of our work. An Atlanta History Center docent said of the historical photos that he had "often been tripped up trying to pinpoint exactly where they were taken", and after interacting with the 4D model of Atlanta that he had "learned more in 5 minutes than I have in the last year."

VIII. Conclusion and Future Work

The natural endpoint of this line of research is to register every photograph in existence, across all time periods, to a common global reference frame and to use this exhaustive photographic record to construct a time-varying 3D model of the world that is as accurate and complete as can be achieved from photographic evidence. Such a comprehensive model would serve as a general reference tool for the visual world, much as Wikipedia or Google Earth are used today. Such a model would not only allow a person to find historical images similar to

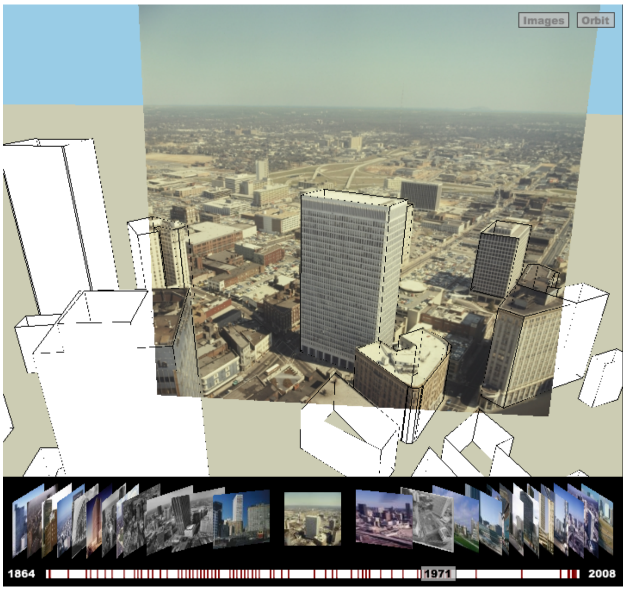

Fig. 13. Downtown Atlanta in 1971. The image appears to be an aerial photograph, perhaps taken from a helicopter. (See Figure 14)

Fig. 14. Downtown Atlanta in 1971. What seems at first to be an aerial image (see Figure 13) was actually captured from the rooftop of the recently finished State of Georgia Building, the tallest building in the Southeast United States at the time.

their own modern photographs, but with further research, to shoot a video and see what it would look like in a different time period, and even to walk around in the present, using a mobile phone as a window into the past via real-time visualization of one's current viewpoint from any time in history.

There are a number of important problems to be solved that would benefit any attempts to reach this ambitious goal. At the low level, designing features that are time-invariant could greatly improve the ability of images to be incorporated into a 4D model in the first place. Higher up the chain, if we have methods of obtaining more accurate building models from segmented point clouds, we would improve both the accuracy of visibility reasoning and the quality of resulting 4D city visualizations. At the visualization level, a future goal is to be able to really dive into a single image of a city at any point in history, combining the appearance of the parts of the scene visible in the given image and the known 3D geometry of the rest of the scene to create a convincing reconstruction of the world at a moment in time.

Finally, one of the major obstacles to reaching these goals is simply getting access to the historical imagery necessary for constructing 4D city models. It is our hope that, as time goes on, more and more of these historical images will become freely available online and that the methods described in this paper will be used to truly unlock the urban photographic record of our world.

ACKNOWLEDGEMENTS

This material is based upon work supported by the National Science Foundation under Grant No. 0534330. This work was partially supported by Samsung Electronics and by a gift from Microsoft Research. The authors would like to thank the Atlanta History Center for providing the historical images of Atlanta used throughout this work.

REFERENCES

[1] Sameer Agarwal, Noah Snavely, Ian Simon, Steven M. Seitz, and Richard Szeliski. Building rome in a day. In *ICCV*, 2009.
[2] P. Debevec, Y. Yu, and G. Borshukov. Efficient view-dependent image-based rendering with projective texture-mapping. In *Eurographics Workshop on Rendering*, pages 105–116, 1998.
[3] P.E. Debevec, C.J. Taylor, and J. Malik. Modeling and rendering architecture from photographs: A hybrid geometry- and image-based approach. *SIGGRAPH*, 30:11–20, 1996.
[4] O.D. Faugeras and Q.T. Luong. *The geometry of multiple images*. The MIT press, Cambridge, MA, 2001.
[5] M. Golparvar-Fard, F. Peña-Mora, and S. Savarese. Monitoring of construction performance using daily progress photograph logs and 4d as-planned models. In *ASCE International Workshop on Computing in Civil Engineering*, 2009.
[6] A. Griewank. On Automatic Differentiation. In M. Iri and K. Tanabe, editors, *Mathematical Programming: Recent Developments and Applications*, pages 83–108. Kluwer Academic, 1989.
[7] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.
[8] X. Li, C. Wu, C. Zach, S. Lazebnik, and J. Frahm. Modeling and recognition of landmark image collections using iconic scene graphs. In *ECCV*, pages 427–440, 2008.
[9] M. Pollefeys, D. Nistér, J.M. Frahm, A. Akbarzadeh, P. Mordohai, B. Clipp, C. Engels, D. Gallup, S.J. Kim, P. Merrell, et al. Detailed real-time urban 3d reconstruction from video. *International Journal of Computer Vision*, 78(2):143–167, 2008.
[10] Grant Schindler and Frank Dellaert. Probabilistic temporal inference on reconstructed 3d scenes. In *CVPR*, 2010.
[11] Grant Schindler, Frank Dellaert, and Sing Bing Kang. Inferring temporal order of images from 3d structure. In *CVPR*, 2007.
[12] S.M. Seitz and C.R. Dyer. View morphing. In *Conf. on Comp. Graphics and Interactive Techniques*, pages 21–30. ACM, 1996.
[13] N. Snavely, S.M. Seitz, and R. Szeliski. Photo tourism: Exploring photo collections in 3D. In *SIGGRAPH*, pages 835–846, 2006.
[14] Noah Snavely, Steven M. Seitz, and Richard Szeliski. Modeling the world from internet photo collections. *International Journal of Computer Vision*, 80(2):189–210, 2008.
[15] G. Wolberg. Image morphing: a survey. *The Visual Computer*, 14(8):360–372, 1998.
[16] L. Zebedin, J. Bauer, K. Karner, and H. Bischof. Fusion of feature- and area-based information for urban buildings modeling from aerial imagery. In *Eur. Conf. on Computer Vision (ECCV)*, 2008.