EUROGRAPHICS Workshop on Graphics and Cultural Heritage (2024)
M. Corsini, D. Ferdani, A. Kuijper, and H. Kutlu (Editors)

Short Paper

Reconstructing Lost Altarpieces: A Differentiable Rendering Approach

A. Pagès-Vilà[1] and I. Munoz-Pandiella[2]

[1]Universitat de Girona, Spain
[2]Universitat Politècnica de Catalunya, Spain

Figure 1: *Using a historical photograph of an altarpiece and knowledge of its original context (left), we can determine the camera position and create a lighting environment map. With this information, we leverage differentiable rendering to generate a 3D model that represents the volumes shown in the historical photograph (right).*

Abstract

Studying works that have completely or partially disappeared is always difficult due to the lack of information. In more fortunate scenarios where photographs were taken before the destruction, the study of the piece is limited by the viewpoints captured in the available photographs. In this interdisciplinary research, we present a new methodology for reconstructing lost altarpieces from a single historical image, utilizing differentiable rendering techniques. We test our methodology by reconstructing some reliefs from the altarpiece of Sant Joan Baptista (Valls, Spain), which was destroyed in 1936. These results are valuable for both experts and the public, as they facilitate a better understanding of the relief's volumetrics and their spatial relationships, representing a significant advancement in the virtual recovery of lost artifacts.

CCS Concepts

• ***Applied computing*** → ***Arts and humanities;*** • ***Information systems*** → *Multimedia content creation;* • ***Computing methodologies*** → *Rendering; Shape modeling;*

1. Introduction

Throughout history, the disappearance of heritage objects due to war, urban development, or mere chance has been common. The loss of the object directly affects our understanding of it, making it hard to fully grasp its details and relationship with its environment. Sometimes, we can get an idea of its appearance if photos were taken before its destruction. As is the case of the object we intend to reconstruct.

In this study, we propose the usage of a modified pipeline of differentiable rendering techniques for virtually reconstructing lost altarpieces from the Renaissance and Baroque periods, utilizing preserved photographs as references. We demonstrate its efficacy through the reconstruction of the lost altarpiece from the Sant Joan

DOI: 10.2312/gch.20241245

Baptista church (Valls, Spain), enriching archival photographs with 3D information via partial 3D reconstructions. To achieve this, we first approximate the illumination of the altarpiece by manually modeling the church and creating an environment map. Next, we determine the camera positioning, aligning it with a prism that represents a simplified version of the relief we intend to reconstruct. Finally, we use this information, along with the processed historical photo, as the initial configuration and reference image in an inverse rendering reconstruction process (see Figure 1).

We have combined art history research with the latest computer graphics techniques in differentiable rendering to develop a method for reconstructing segments of a lost altarpiece from a single black and white image. We verify our approach by creating an ideal case scenario and investigate how photograph contrast impacts the reconstruction. Finally, we propose a method to reduce these effects.

Our work represents considerable progress in the reconstruction of lost cultural heritage objects, offering valuable insights in both art history and computer graphics.

2. Previous work

3D digital models in Cultural Heritage (CH) are typically created using scanning technologies like photogrammetry or laser scanning, and 3D modeling tools when direct scanning is not possible (e.g. missing objects). For the latter case, promising alternatives are being studied, such as deep learning-based 3D model generation techniques [SS23, HLB21]. The creation of 3D models from a single image has also been investigated, with Mohan and Mani [MM11] and Fahin et al. [FAZ21] presenting state-of-the-art methods that focus on deep learning approaches. However, no prior work has targeted high-detail 3D model generation from a single historical photo, a crucial gap for CH where historical photos are often the only available information about the geometry of a lost artwork.

Regarding reconstructions of altarpieces, Isidro Ot et al. [OMS20] used archive photographs as guides to manually model the church of Martorell, Spain, where they then positioned photos of all the lost altarpieces in their original place as planes. Although they were digitally located in their original place, they appear as flat objects. Pagès-Vilà et al. [PVPMP24] worked on the partially lost Rosary altarpiece of the Sant Pere Màrtir church in Manresa, Spain. They combined manual modeling of the altarpiece's lost structural geometry with photogrammetric techniques for the preserved fragments to reconstruct it.

Archival photographs have been used as the main source to reconstruct a lost or partially lost object in several occasions with different approaches. Klaus Hanke et al. [HRM15] has only two images of the Capuchin monastery's altarpiece in Kitzbüel, Austria, in order to reconstruct it, they model it manually employing CAD tools. Condorelli & Morena [CM23] added a manual step to their photogrammetry workflow to extract feature points missed automatically, due to the insufficiency of the 7 available photographs for generating a point cloud of Centrale Station in Caltanissetta, Italy. Finally, Wiedemann et al. [WHA00] used the identical repetition of geometric features in the facade of the lost Stadtschloss building (Berlin, Germany) to reconstruct it, overcoming the challenge of incomplete information provided by the archival photos.

However, we cannot rely on any of these approaches as none of them meet our requirements.

3. Case Study

To validate our technique, we chose to reconstruct two of the reliefs of the altarpiece of the church of Sant Joan Baptista in the city of Valls (Spain). It was located in an elevated chancel, behind the main altar of the single nave church, where it stood at 24 meters in height, exhibiting the work of some of the best 17^{th} century Catalan sculptors and painters. But as happened in almost all the churches in Catalonia at the beginning of the Spanish Civil War (1936-1939), the altarpieces were burned down. Only the alabaster panels, some caryatids and columns survived. In 1954, the sculptor Josep Busquets remade the altarpiece based on historical photographs, but it differs significantly from the original.

We conducted historiographical research in photographic archives and discovered many photographs of the altarpiece taken by various people using unknown cameras and techniques. While most of these photos were not suitable for reconstruction, we did find a few close-up images that we can use.

4. Differentiable rendering for 3D reconstruction

One method that has been used for 3D reconstruction and fills our requirements is inverse rendering, a technique that can handle large optimization problems such as mesh reconstruction. In our work, we utilized the inverse rendering algorithm proposed by Nicolet et al [NJJ21], which is integrated in Mitsuba 3 rendering system [JSR*22] This technique has proven to be effective in reconstructing shapes from renders.

The central concept of our work is to use the archival photo as the reference image instead of the target object, as the former is our desired outcome. To our knowledge, differentiable rendering has not been previously been used to reconstruct missing objects from a photograph.

5. Best case scenario validation

We ran a preliminary test to determine the viability of our goal by building an ideal scenario with an accurate environment map and a photo of a single-colored, non-reflective altarpiece's relief from the same time period.

We placed our test object in a room and created a 360º panoramic HDR photograph, by stitching together multiple images, as our environment map. Additionally, we took six photos of the object and processed them in Metashape, following a traditional photogrammetric pipeline, from which we got the camera's alignment and a point cloud. To match the results from Metashape with Mitsuba, we set up a shared coordinate system in Blender. We aligned the cameras and point cloud so that their origins corresponded with the location of Mitsuba's initial object, the prism, which will deform to match the relief's shape. From the available cameras, we selected the one that had been taken from the most similar viewpoint to the

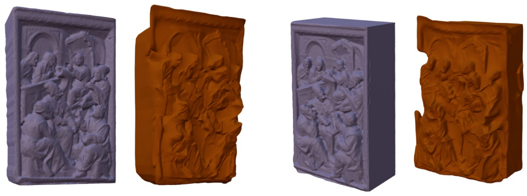

Figure 2: *The result of the differentiable rendering reconstruction of the 3D print using a single photograph as input. Purple represents the scanned 3D printing, while orange indicates the reconstruction.*

archival photographs we will use in the following Section. We configured the camera in Mitsuba using the intrinsic parameters from the camera.

In a standard differentiable rendering pipeline, the method deforms the initial object into the target object using the initial and target objects, camera, and environment map. However, instead of using a target object, we substituted it with our photograph. To do so, we first rendered the scene with the initial object and then overlaid our photo to cover the object while preserving the background from Mitsuba's render. The pipeline then continued with iterative deformations of the initial object, comparing it to our adapted render.

We concluded that, our methodology demonstrates being applicable in an ideal scenario with a high-quality photograph and a precisely accurate environment map (see Figure 2).

6. Input data preparation

6.1. Environmental illumination

To estimate the altarpiece illumination, we created an environmental map. First, we modeled the church in Blender and, then, created a render with Blender's panoramic equirectangular camera, located where the relief we want to reconstruct was situated.

6.2. Initial rendering

From the studied altarpiece, there are preserved two black and white images of two reliefs (1924 and 1931) taken with an unknown camera. These photos have scratches, stains and color shifts that will affect the final result. They also have different luminance and contrast as they were developed manually.

We followed the same pipeline as detailed in Section 5 (see Figure 3). As the photos of the altarpiece were taken from a nearly identical viewpoint, we decided to use only one of them as the input for the reconstruction, as omitting the others does not significantly alter the result. Processing them in Metashape produces a flat point cloud not suitable for 3D reconstruction. However, the camera alignment is usable, allowing the resulting point cloud to be aligned with the initial object.

As shown in the reconstruction of the 3D print, our objective is to replace Mitsuba's render of the target scene with part of the historical photograph. In this sense, we configure Mitsuba's camera parameters. We then create the initial rendering with the initial object, the camera, and the environmental map. Following this, we adapt the initial render by cropping the photograph to match the shape of the initial object and overlaying it, as shown in Figure 3 (bottom right).

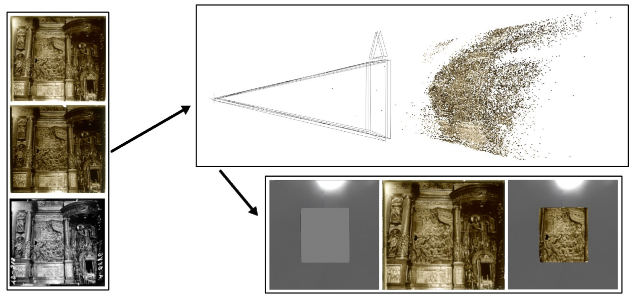

Figure 3: *Photographs of the relief we aim to reconstruct on the left. After aligning these photographs, we obtained a point cloud (right). We generated a render of the target scene (first image at the bottom right) and then cropped the original photograph of the relief to match the shape of the initial object.*

7. Contrast study

We noted that directly using the historical photograph resulted in exaggerated slopes in the reconstructed geometry. We attribute this to the manual development of the photo, which could have been adjusted to produce a more contrasted image for a more appealing look, a common approach in the early 1900s. To validate our hypothesis, we progressively lowered the contrast of the input historical photo in order to see its impact on the reconstruction.

Our analysis shows that contrast significantly affects the outcome, with geometry appearing flatter as contrast decreases, emphasizing the need to alleviate these effects. Due to insufficient information to digitally reverse contrast adjustments, we propose an experimental approach, rendering the 3D printed altarpiece in Mitsuba under the same lighting and camera conditions used for reconstruction. We use the contrast of the rendered image as the reference contrast that the historical photograph should ideally possess for being reconstructed.

To adjust the contrast, we first ensure consistency in hue, converting the images to HSV and computing the histogram of V in both images. Then, we calculate a scaling factor s as:

$$s = \frac{target_{p99} - target_{p01}}{input_{p90} - input_{p10}} \quad (1)$$

Here, $target_{p_i}$ corresponds to the ith percentile of the V distribution of the render and $input_{p_i}$ represents the same for the photograph. We use percentiles to reduce the possible effects of outliers. These percentiles differ between images and have been chosen heuristically, as the render shows a diffuse object while the relief in the photograph is specular and highly contrasted.

Then, for each pixel p, its value p_v ($p_v \in [0,1]$) is scaled as follows:

$$p'_v = (p_v - 0.5) \cdot s + 0.5 \quad (2)$$

8. Results and Discussion

We used a computer equipped with an 11th Gen Intel(R) Core(TM) i7-1165G7 @ 2.80GHz processor, accompanied by 16GB of RAM, running on a 64-bit operating system. Additionally, we employed an NVIDIA T500 graphics card with 4GB of video memory. The photographs used in Section 5 were taken using a SONY DSLR-a100 camera.

The 3D print reconstruction was completed in under an hour with over 100 iterations. To speed up Mitsuba's processing time, we reduced the SONY photos from 3872x2592 to 775x519 pixels. Similarly, we resized the archival photos to 770x735 pixels and 742x740 pixels. The final results, which were computed after 800 iterations, took approximately 8 hours.

This research demonstrates that reconstructing an object from a single image is feasible when spatial context and lighting conditions are precisely defined. To achieve our goal of visualizing the altarpiece as it once was, we textured the reconstruction with a projection of the historical photograph. Since the original object no longer exists, we can only compare visually, but our reconstruction seems to more closely resemble the original than Busquets' version.

Our contrast study revealed that photograph contrast significantly affects the depth of the resulting geometry, as inverse rendering deforms the geometry to match both illuminated and shadowed areas. Adjusting the contrast provides a promising balance, resulting in an adequate representation of the reliefs' depth that help us understand the shape of the relief, as shown in Figure 4 The loss function for both figures decreases greatly for the first iterations, followed by a more gradual decrease in subsequent iterations, eventually reaching stabilization at 800 iterations, with a value of less than 0.012.

Since the altarpiece had not been cleaned when the historical photos were taken, the gold leaf did not shine as much, which helped mitigate issues related to bright spots. This is advantageous for our methodology, which faces challenges with high specularity. However, some areas in the photos still appear bright, such as the torso and left art of Saint John (kneeling) (see Figure 4, right). Despite this, our methodology proves to be robust handling moderate specularities.

The limited understanding of the reliefs' polychromy remains a challenge to address. In light-matter interaction, a surface's reflective properties affect the colors seen in a photo. If this is in black and white, different colors with similar luminosities can appear as the same shade of gray: a dark area in the photo might be misinterpreted by differentiable rendering as lacking illumination, thus making geometry adjustments until the area is correctly shaded.

9. Conclusions

In this interdisciplinary work, we propose a methodology for reconstructing segments of an altarpiece using a single black and white image. We utilize the differentiable rendering technique supported by the Mitsuba Renderer, making adjustments to the pipeline to achieve our goal. Our approach employs a historical photograph as the target image. To our knowledge, this is the first study to utilize differentiable rendering for shape reconstruction from photographs. We tested our methodology on the altarpiece of the San Joan Baptista church in Valls, Spain.

After its destruction in 1936, the altarpiece can only be seen through black and white photos preserved in archives. Our approach reconstructs reliefs by modeling the church to create an environmental map and using a single photograph as the target image. Given the sensitivity to contrast in black-and-white images, we introduced an automatic contrast adjustment to improve geometry accuracy. Overall, our work marks a significant advancement in the virtual reconstruction of lost altarpieces.

The results reveal the intricate details of the reliefs, providing valuable insights into their artistic quality and clarifying the relationship between the altarpiece and the church. Additionally, they highlight the importance of understanding the spatial context of cultural heritage objects and the methods and techniques used to take the 20^{th} century photographs we rely on for study.

In future research, we intend to investigate techniques to address and potentially reverse conservation issues in photographs in order to understand how the state of preservation affects our reconstruction efforts. Additionally, we aim to study the interaction between light and matter in historical black and white photographs, and how polychromy impacts our reconstruction process. Our goal is not only improving the quality of shape reconstruction but also provide insights for accurate colorization.

Acknowledgements

The authors would like to thank Prof. Xavier Pueyo and Prof. Joan Bosch for their support and advice during the development of the research.

This work has been partially supported by the projects PID2021-122136OB-C21 and PID2021-122136OB-C22 funded by MCIN/AEI/ 10.13039/501100011033 and by FEDER A way of making Europe, by the project PID2020-114339GB-C21 funded by the MCIN/AEI, and by the Department of Research and Universities of the Generalitat de Catalunya (2021 SGR 01035).

References

[CM23] Condorelli F., Morena S.: Integration of 3d modelling with photogrammetry applied on historical images for cultural heritage. *VITRUVIO - International Journal of Architectural Technology and Sustainability 8* (04 2023), 58–69. doi:10.4995/vitruvioijats.2023.18831. 2

[FAZ21] Fahim G., Amin K., Zarif S.: Single-view 3d reconstruction: A survey of deep learning methods. *Computers & Graphics 94* (2021), 164–190. URL: https://www.sciencedirect.com/science/article/pii/S0097849320301849, doi:https://doi.org/10.1016/j.cag.2020.12.004. 2

[HLB21] Han X.-F., Laga H., Bennamoun M.: Image-based 3d object reconstruction: State-of-the-art and trends in the deep learning era. *IEEE Transactions on Pattern Analysis and Machine Intelligence 43*, 5 (2021), 1578–1604. doi:10.1109/TPAMI.2019.2954885. 2

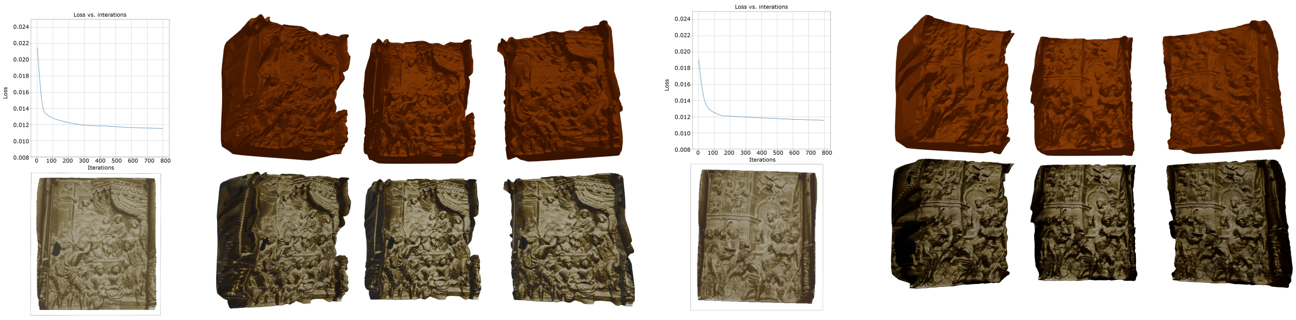

Figure 4: *Reconstruction results for two different reliefs of the altarpiece. On the left, we present both the loss graph and the input image employed for the reconstruction. On the right, the resulting geometry of the reconstruction is displayed with and without the original photo used as the texture (top and bottom, respectively).*

[HRM15] HANKE K., RAMPOLD R., MOSER M.: Historic photos and tls data fusion for the 3d reconstruction of a monastery altar ensemble. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XL-5/W7* (2015), 201–206. doi:10.5194/isprsarchives-XL-5-W7-201-2015. 2

[JSR*22] JAKOB W., SPEIERER S., ROUSSEL N., NIMIER-DAVID M., VICINI D., ZELTNER T., NICOLET B., CRESPO M., LEROY V., ZHANG Z.: Mitsuba 3 renderer, 2022. https://mitsuba-renderer.org. 2

[MM11] MOHAN S., MANI L. M.: Construction of 3d models from single view images: A survey based on various approaches. In *2011 International Conference on Emerging Trends in Electrical and Computer Technology* (2011), pp. 557–562. doi:10.1109/ICETECT.2011.5760180. 2

[NJJ21] NICOLET B., JACOBSON A., JAKOB W.: Large steps in inverse rendering of geometry. *ACM Transactions on Graphics (Proceedings of SIGGRAPH Asia) 40*, 6 (Dec. 2021). doi:10.1145/3478513.3480501. 2

[OMS20] OT I., MAURI A., SOCORREGUT J.: De la Investigación histórica y arqueológica a la divulgación mediante la virtualización. In *I Simposio anual de Patrimonio Natural y Cultural ICOMOS España* (2020), p. 181–187. doi:10.4995/icomos2019.2020.11710. 2

[PVPMP24] PAGÈS-VILÀ A., PUEYO X., MUNOZ-PANDIELLA I.: Digital reconstruction of partially lost altarpieces. the case of the rosary's altarpiece of sant pere màrtir de manresa. *J. Comput. Cult. Herit.* (apr 2024). URL: https://doi.org/10.1145/3652860, doi:10.1145/3652860. 2

[SS23] SAMAVATI T., SORYANI M.: Deep learning-based 3d reconstruction: a survey. *Artificial Intelligence Review 56*, 9 (2023), 9175–9219. 2

[WHA00] WIEDEMANN A., HEMMLEB M., ALBERTZ J.: Reconstruction of historical buildings based on images from the meydenbauer archives. *International archives of photogrammetry and remote sensing 33*, B5/2; PART 5 (2000), 887–893. 2