

Digitizing the Wooden Scale Model of a Twentieth Century Urban Utopia for the Development of a Full-Scale Virtual Tour Application Using a Game Engine

Steven Pigeon¹, Marianne Bérubé-Dufour^{1,2}

¹ Département de mathématiques, informatique et génie,
Université du Québec à Rimouski, Québec, Canada – steven_pigeon@uqar.ca

² Office d'habitation Rimouski-Neigette, Rimouski, Canada – marianne.berube-dufour@ohrn.ca

Keywords: Scale Model, Photogrammetry, Depth of Field, Ground Sampling Distance, Video Game Engine, Texel Density

Abstract

We created a virtual tour application using photogrammetry, GIS data, and a game engine to allow a greater audience to discover in an immersive way the wooden scale model (scale 1:1000) of a Utopian version of the city of Rimouski imagined by the architect Luc Laporte. In this paper, we show how we determined, based on the quality requirements for video game assets (texel density), what are the conditions to be met during photogrammetry survey (GSD) of such a scale model intended to a virtual tour application. We explain how creating a 3D version of Laporte's wooden model is a problem very similar to air-based photogrammetry of urban spaces but, contrary to typical aerial photogrammetry where the distance d is much greater than the focal length f , we are strongly limited by the depth of field as, for the scale model, the distance is only somewhat greater than the focal length, yielding a reduced depth of field.

1. Introduction

Before computer graphics and modeling software, urban designer and architects made extensive use of scale models in order to test and display their ideas. Le Corbusier, for example, created a large number of such models and his archives still keep 54 models (de La Cova, 2019, Fondation Le Corbusier, 2025). Models also serve to capture time-frozen snapshots of premises: Duberger-By's relief map showing Québec City in 1806 (Charbonneau, 1981) or Bigot's Plan de Rome (Madeleine and Fleury, 2024, Fleury and Madeleine, 2011) are famous examples.

At the turn of the century, the architect Luc Laporte (1942–2012) imagined a Utopian version of the city of Rimouski, rebuilt on the Île Saint-Barnabé, a $5.5\text{ km} \times 400\text{ m}$ island set in the Saint-Lawrence River, just a few kilometers from the city (fig. 1) (Laporte and Perrault, 2000). In order to transform this typical mid-size American-style city in a dense all-pedestrian city of the future, Laporte drew inspiration, ironically, from the great European cities developed before the automobile. Laporte, trained as an architect before the advent of computer assisted design (CAD), relied on traditional techniques: hand-drawings, sketches, blueprints, and scale models. For his vision of Rimouski rebuilt, Laporte and his team created an approximately 6 m long wooden scale model of the city. Commissioned by the Musée Régional de Rimouski, the artwork, shown in fig. 2, is kept in the museum’s archives and is rarely displayed as it requires a lot of space (Bérubé-Dufour, 2024).

To allow a greater audience to discover the scale model—that many consider as being one of the best example of Laporte’s architectural legacy—in an immersive way, we created a virtual tour application using photogrammetry, GIS data, and a game engine. Instead of simply offering a bird’s-eye view of the 3D model of the city, shown in fig 3, as is often the case with 3D models, we created a “game” in which the player/spectator can wander to his will through the streets as one would in a real city.



Figure 1. Rimouski and Île Saint-Barnabé (St-Barnaby's Island).

and discover the vistas it offers (fig. 4). The process of creating such a game is not an entirely trivial one: we had to import GIS data for the larger scenery, use photogrammetry to acquire the 3D model, use mesh editing software to edit the mesh where software failed to provide a correct reconstruction, and, lastly, import all the data into a game engine.

The 3D model of the city was inset into its surrounding area using GIS data. The GIS-based mesh represents an area of approximately 625 km², centered on the island. The data was integrated into the 3D model via the ArcGIS API as a basemap with elevation data (ESRI, 2022, Ministère des ressources naturelles et des forêts du Québec, 2025). To reconstruct the 3D model using Structure From Motion (SFM) from a large

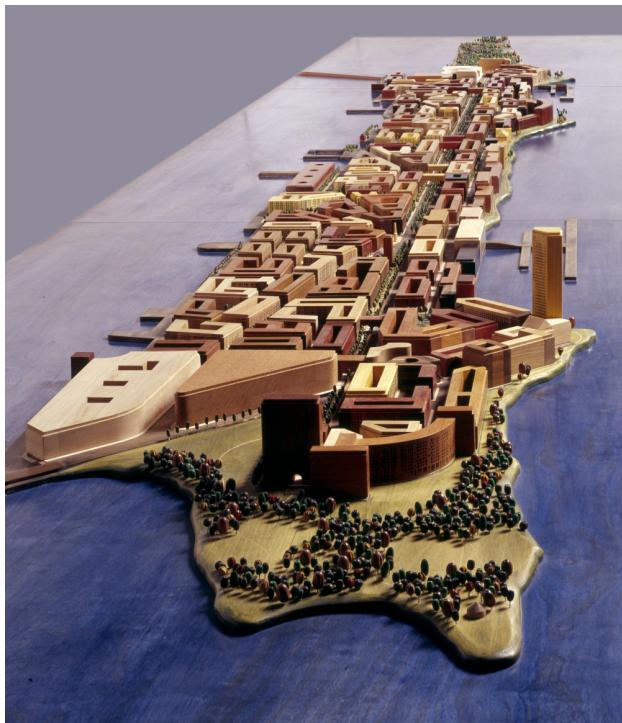


Figure 2. The scale model created by Laporte and his team.
Photo: André Cornellier, with permissions (2000).

number of photos, the open-source 3D reconstruction software Meshroom (Griwodz et al., 2021) was chosen. The set of 798 pictures, captured by IDHP, an external firm, was taken along a series of strips orthogonal to the model. For topological errors, such as holes in buildings and distorted trees (fig. 5), we used 3ds Max (Autodesk, 2025) to correct both mesh and textures (fig. 6). Finally, the corrected mesh and textures were imported in the Unity game engine, where, using our game, the player can move around the city using either mouse or joystick (Unity Technologies, 2025, Bérubé-Dufour, 2024). Fig. 7 shows the relation between the software used in the creation of the game, while fig. 8 demonstrate the Unity development interface.

While the 3D model obtained by SFM from the set of photos is satisfactory when viewed from a distance, say from the viewpoint of an observer standing next to the model in an exhibition hall, a closer inspection reveals a number of defects, especially when the model is viewed from inside the game, from the point of view of the player. Fig. 5 shows such defects where SFM failed to reconstruct the object geometry correctly. The mesh had to be hand-corrected—a painstaking and time-consuming task—to obtain better results, shown in fig. 6.

In this paper, we explore the difficulties of using photogrammetry to create a realistic 3D version of a 1:1000 wooden scale model in a video game. We will discuss the various parameters one must control to achieve a satisfactory image capture, and what texel density must be attained. We will also discuss future work.

2. Texels, Ground distance, and Depth of Field

Creating a 3D model of a wooden scale model of an entire city usable in a first person video game is neither the usual problem of scanning an object to make a video game asset, nor is it the usual problem of using photogrammetry to reconstruct large

urban areas. Indeed, this problem draws from both, but is more complicated than either.

In a typical first person shooter (FPS) game, we see the world on screen as the character would see it, and close views, even to objects of interest (called “assets”), are typically set at about 1 m from the player in the simulated world. If the game is played as a third person game (TPS), the scene is shown from a somewhat greater, but still moderate distance, as if spectated by someone behind, or above, the player. Fig. 9 show the relation between TPS and FPS. However, in video games, regardless of whether it is a FPS or a TPS, the goal is not to be *accurate* to the objects, but merely to create a *sufficient illusion* so that the player’s disbelief is suspended. Therefore, the texel (texture element) density, or how many texels there are given a certain length along the object in-world (that we will note p/l for p texels given length l), doesn’t need to be very high for video game objects (in the order of one texel per object mm), as the objects will likely be rendered on a small region of the screen, which will cause texels to merge onto fewer pixels. Reduction is also applied to the object meshes, which do not need to be overly detailed, as, again, illusion suffices: bump maps, shading and texture will provide what the mesh cannot and should not provide.

As video game developers use more and more often close-range photogrammetry than the more traditional modeling pipeline to create realistic assets, the steps have been explored in some detail. The steps to create assets in video games are (Lachambre et al., 2017, Statham, 2018):

- Image acquisition, mesh reconstruction, and texture mapping, typically using SFM, yielding high resolution textures and a dense mesh with a high number of points and polygons;
- Mesh reduction, where the number of polygons is reduced so that, at given resolution, the shape is mostly preserved (it may also include meshes with different levels of detail according to the viewing distance);
- Texture baking, that is, folding the high resolution texture onto the lower precision mesh;
- A final adjustment of texture resolution taking into account the final rendering conditions, such as viewing distance and screen resolution, while avoiding blurry or pixelated textures (Dries, 2023).

However, for our game, where the source object is a scale model, the needed texel density is *much* higher, since the object implies a scale parameter, m . This scale parameter is $m = 1000$, since the scale of Laporte’s model is 1:1000, that is, 1 mm on the scale model represents 1 m in the real world. Taking the scale parameter into account we get a texel density of p/lm .

Creating a 3D version of the wooden model is a problem very similar to air-based photogrammetry of urban spaces, where not all surfaces are easily observable, and where careful planning of view-taking is necessary (Leberl et al., 2012, Gao et al., 2023, Adami et al., 2023). Indeed, because of the scale parameter of the model, our problem becomes essentially equivalent to reconstructing a 3D model of an actual city.

But to do so, we must get a very high texel density, which, in the context of aerial photogrammetry becomes the ground

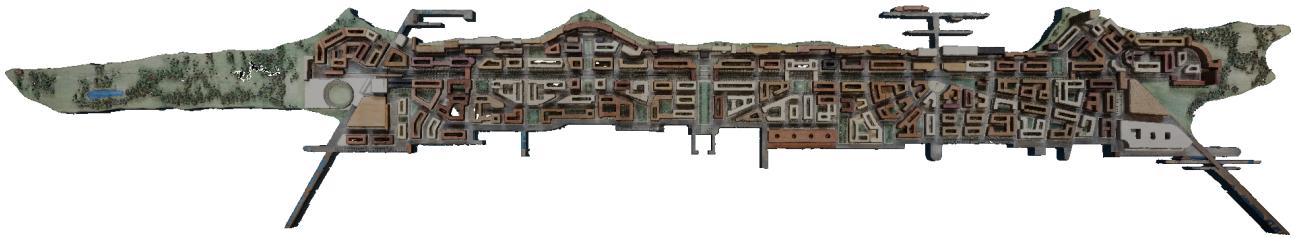


Figure 3. 3D model of the city corrected using 3D modeling software.



Figure 4. The surrounding landscape seen at the end of a street.

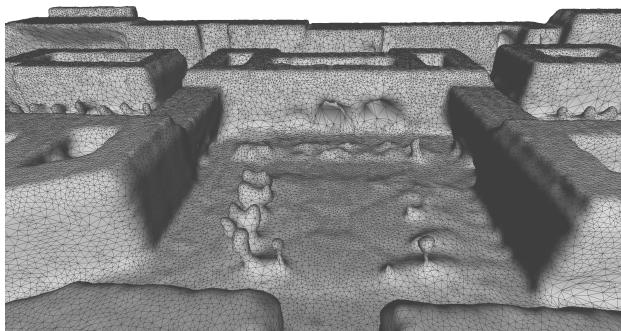


Figure 5. Holes in the reconstructed mesh.

sampling distance, or GSD. The GSD measures the distance on the ground corresponding to one pixel of the captured image¹. Fig. 10 shows a typical arrangement for aerial image capture. The GSD is determined by the lens characteristics (its focal length f and aperture number N), the sensor size S , its number s of pixels in the direction as S , and the distance to the object, d . The GSD is therefore given as

$$\text{GSD} = \frac{Sd}{sf} = g \quad (1)$$

where g is the ratio of ground distance to pixel.

However, for our game, we must also take the scale m of the model into account and eq. (1) is modified to

$$\text{GSD} = \frac{Sdm}{sf} = g_m, \quad (2)$$

as the physical distance d between the camera and the scale model would be a distance dm in the real world. This also explains how the texel density should become p_{lm} .

¹ In the older aerial reconnaissance literature, the GSD is known as the *representative fraction*, linking distance on film and distance on the object (Avery, 1968, p. 37–38), and sometimes characterized in terms of film lines/mm, a measure akin to pixels for analog film (Strandberg, 1968, p. 15).



Figure 6. Texture applied on the corrected mesh.

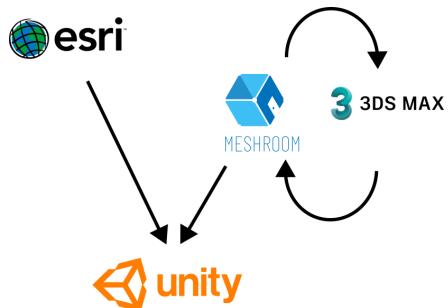


Figure 7. Dependencies between the various software used to develop the “game”.

But, contrary to typical aerial photogrammetry where the distance d is *much* greater than the focal length f , we are strongly limited by the depth of field as, for the scale model, the distance is only *somewhat* greater than the focal length, yielding a reduced depth of field. Consider the general arrangement of fig. 11. A point-like source of light will go through the lens and hit the sensor. However, due to the imperfections of the lens, that point-like source will not resolve as a point, but as a (hopefully) small fuzzy disk, the circle of confusion. All non point-like objects small enough to resolve inside the circle of confusion determine the depth of field around these objects (shown in green in fig. 11). The total depth of field at distance d , given focal length f , aperture f/N , and circle of confusion c , will be given by

$$\text{DoF} = \frac{2f^2Nd^2}{f^4 - (cNd)^2}, \quad (3)$$

a formula derived, for example, in (Ray, 1988, p. 182ff).

Depth of field is crucial for the correct operation of SFM, as SFM searches images for corresponding “features” to match object points across images (Bianco et al., 2018, Gruen, 2012, Griwodz et al., 2021, Remondino and Elhakim, 2006, Schonberger and Frahm, 2016, Jiang et al., 2020). However, if a feature is sharp and in focus in one image and blurry and out of focus in another, SFM algorithms will fail to recognize

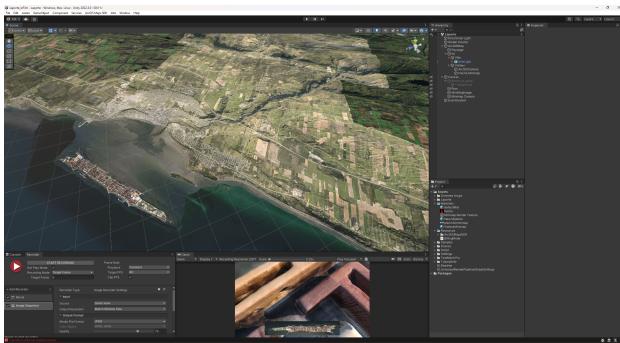


Figure 8. The Unity development platform.

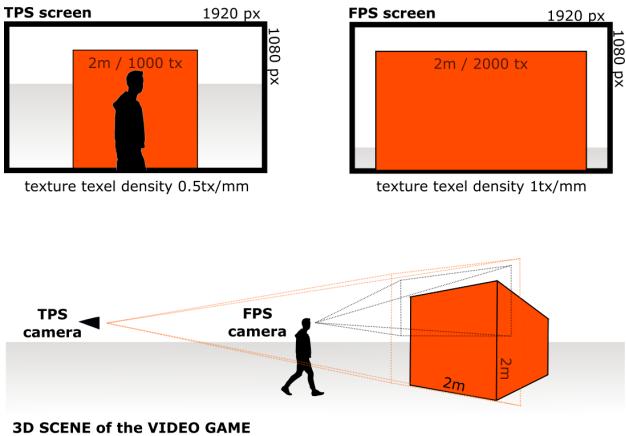


Figure 9. Texel densities vary according to the view.

the match (Nicolae et al., 2014, Menna et al., 2014). Therefore, one must make sure that all features are simultaneously in focus across all images, which implies that the depth of field must be sufficient to encompass all of the objects into view for each picture. We can only achieve this result by having a sufficient depth of field.

To achieve a sufficient depth of field, we must adjust the lens parameters. The focal length f may be variable or fixed given the chosen lens, while the aperture number N and the distance d are adjustable. The circle of confusion c is a characteristic of the lens and could be considered constant, but N affects its size: as N grows, the aperture is reduced and so is c ; but if we have $c \leqslant \frac{S}{s}$ (which may not be achievable with the lens), we can assume that the image is in perfect focus.

3. A Better Capture

For our scale model and our virtual tour “game”, we not only must have sufficient GSD and depth of field, but a sufficient high texel density to render the scale model in our FPS game with FPS conditions (see fig. 9).

The wanted texel density is about one texel per *real world* millimeter, and so we have to solve eq. (2) for $g_m \approx 1$ with $m = 1000$ for d and/or f . For our camera, a Canon EOS 850D (also known as T8i), the sensor size is $S = 22.3$ mm wide, with $s = 6000$ pixels horizontally. Using eq. (2), we find that any combination of d and f such that $\frac{d}{f} \approx 0.269$ will be a solution. We quickly see that with a standard 17–85 mm zoom lens, the distance to the model would need to be between ≈ 4.57 mm and ≈ 22.87 mm to achieve that goal. The problems would be,

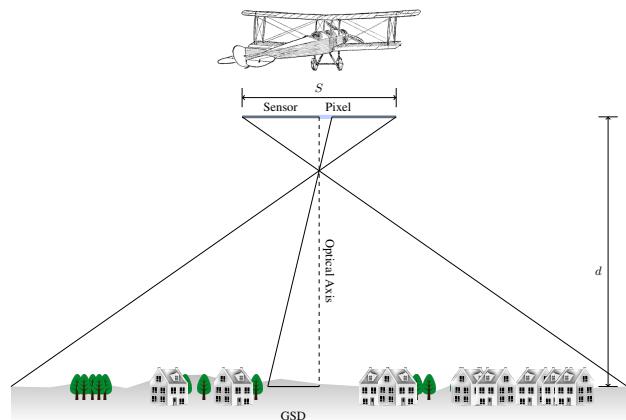


Figure 10. Ground Sampling Distance.

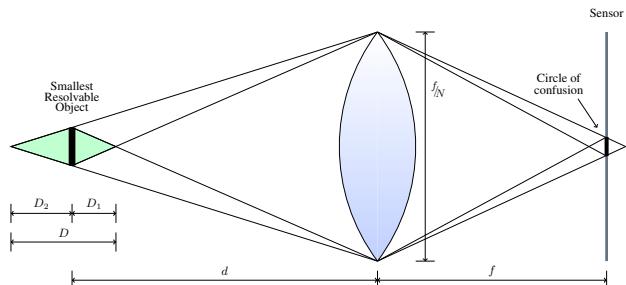


Figure 11. Simplified Depth of Field.

of course, a very narrow depth of field at those distances, would the lens being able to focus that close—typically, the closest possible focus is in the 30 cm range—and having the lens that close to the model limits its mobility and increases the risk—or the need—for contact with the scale model. We might want to solve for $g_m \approx 5$, finding $\frac{d}{f} \approx 0.804$, with distance d varying from ≈ 13.66 mm to ≈ 68.30 mm. If we must solve for f with $g_m \approx 1$ but $d \geq 30$ cm, we find $f \geq 1115$ mm, which is rather cumbersome but not impossible².

Next adjustment is the depth of field. The total depth of field is given by eq. (3), but if we examine fig. 11, we notice that the depth of field is not symmetrical about the focal plane: it extends further behind than in front of the focal plane. Indeed, we can find that

$$D_1 = \frac{cNd^2}{f^2 + cNd}, \quad (4)$$

while

$$D_2 = \frac{cNd^2}{f^2 - cNd}. \quad (5)$$

This suggests a strategy to place the focal plane in order to capture the scene completely in focus. In fig. 12, the focus was obtained using the camera built-in autofocus. Regardless of the specific algorithm used, the camera will select a few regions on the image and try to bring those points only into focus (Baltag, 2015). However, a better strategy would be to not rely on the autofocus but to chose explicitly a focal plane, as shown in fig. 13. In fig. 13, the focal plane is chosen so that given the lens parameters, the depth of field encompasses the whole scene, yielding a completely in-focus image.

However, getting a sufficient depth of field might necessitate a larger F-number, that is, a large N . Figs. 14, 15, and 16

² There are commercially available 1200 – 1700 mm zoom lenses.

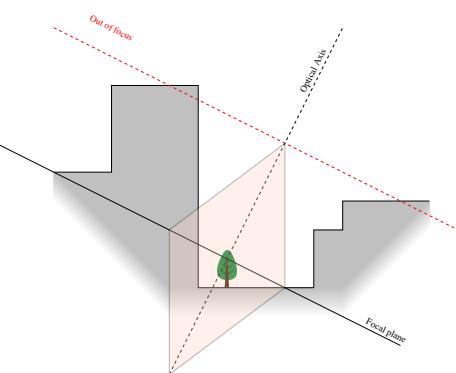


Figure 12. Badly adjusted depth of field.

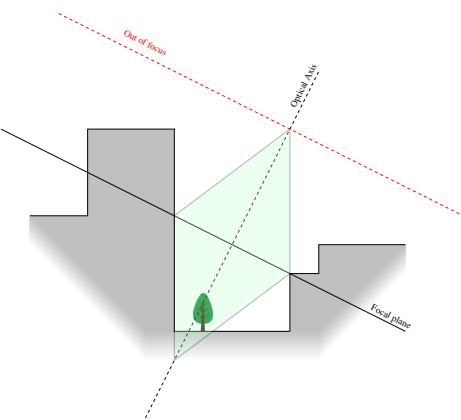


Figure 13. Well adjusted depth of field.

show the effect of N on the depth of field. The lens used was a Canon EF 24 – 70 mm F/2.8-F/22 II USM lens set at $f = 55$ mm with the camera sensitivity set to ISO 250. We see that even if the focal plane is adjusted correctly—about mid-height of the standing blocks—the depth of field is insufficient at $f_{2.8}$ to get the complete scene into focus. Using f_{8} increases the depth of field, but only at f_{22} do we achieve the desired result. But if using a larger F-Number N brings the desired result, we also increased the exposure time quite significantly! From $\frac{1}{30}$ s with $f_{2.8}$, we get an 2 s exposure time with f_{22} . While such long exposure times would be problematic for aerial photogrammetry—as vehicle and ground have a high relative speed—it is also not a situation we encounter: the distance d is so large relative to the focal length f that the depth of field is sufficient even with small N ! For our problem of creating a 3D version of our scale model, this means that the time needed for image acquisition is greatly increased—which may not be a problem with adequate automation!

Therefore, to create a satisfying capture of the scale model, we must use the largest possible focal length f , which will bring us as close as possible to the desired texel density, and the largest depth of field possible, using a larger F-Number and therefore longer exposure times.

4. Future Work

While satisfactory when viewed at what one could call a normal viewing distance—corresponding to an observer standing besides the scale model in an exhibition room—the mesh and texture obtained from the 798 pictures taken in our first experiment, shown in fig. 3, show their limitations in a



Figure 14. Simulated capture: ISO 250, $f_{2.8}$, $\frac{1}{30}$ s.

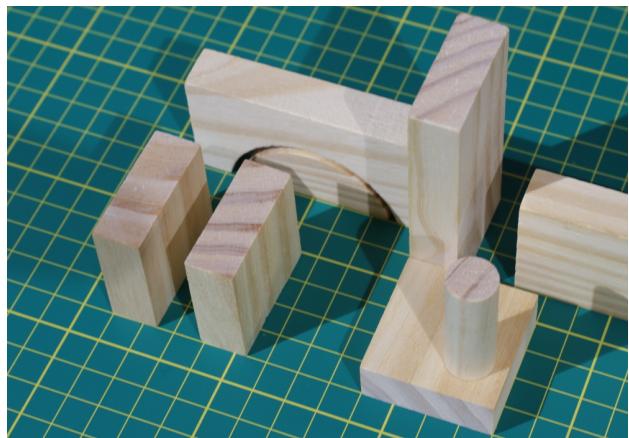


Figure 15. Simulated capture: ISO 250, f_{8} , $\frac{1}{3}$ s.

FPS settings, shown in fig. 4. However, the all images weren't captured in exactly the same way, leaving corresponding regions at different focus in different images. To avoid this problem, a second capture with carefully planned depth of field and focal plane must be performed. Furthermore, a structured sequence of angles, distances, and spacings between pictures must be performed, and necessarily with the assistance of computer-controlled camera movement.

5. Conclusion

Creating a 3D version of a scale model is a good way to not only make the artwork accessible to a greater audience via, for example, a first person game allowing the observer to wander around, but also to help architects and urbanists understand a project from its scale model and other artifacts such as plans and sketches. In the case of Laporte's insular utopia, creating a 3D model helped us not only understand the vision behind the model (Bérubé-Dufour, 2024), but also see its limitations and constraints, and how the city, would it have been built, relate to its surrounding environment: the Saint-Laurence river, the coast, and its mountains. But there are still problems to be solved satisfactorily, for example faithfully capturing the geometry of the model and obtaining textures with sufficient texel density. Ideally, the procedure should be fully automated.

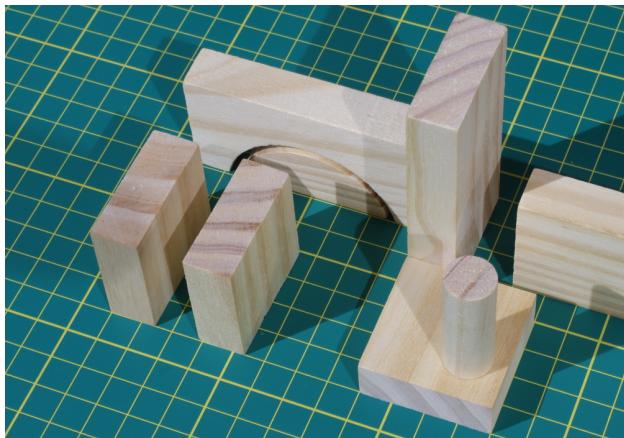


Figure 16. Simulated capture: ISO 250, $f/22$, 2 s.

Acknowledgments

The photogrammetry survey was conducted with the help of the external firm **Image Drone Haute Précision**, and was partially funded by the Government of Canada with the Digital Access to Heritage Museums Assistance Program and by the Université du Québec à Rimouski. The authors also thank the personnel from the Musée Régional de Rimouski for their support and help during the acquisitions. House clip-art in fig. 10 is by Roger Borrell, and is used under the CC Attribution-Share Alike 4.0 license (Wikimedia). Map in fig. 1 b y Service de la cartographie, Ministère des ressources naturelles, Québec (1994).

References

- Adami, A., Treccani, D., Fregonese, L., 2023. Lessons Learnt from the High Resolution UAS Photogrammetric Survey of a Historic Urban Area: UNESCO Site of Sabbioneta. *The Int. Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLVIII-M-2-2023, 629–635.
- Autodesk, 2025. Autodesk 3ds Max. <https://www.autodesk.com/ca-en/products/3ds-max/overview>.
- Avery, T. E., 1968. *Interpretation of Aerial Photographs*. 2nd edn, Burgess Publishing Company.
- Baltag, O., 2015. History of Automatic Focusing Reflected by Patents. *Science Innovation*, 4(1), 1–17.
- Bianco, S., Ciocca, G., Marelli, D., 2018. Evaluating the Performance of Structure From Motion Pipelines. *Journal of Imaging*, 4(8), 18p.
- Bérubé-Dufour, M., 2024. Une cité virtuelle pour 33 296 habitants: découvrir une œuvre architecturale utopique à l'aide de la photogrammétrie, d'un moteur de jeu vidéo et d'un système d'information géographique. Master's thesis.
- Charbonneau, A., 1981. *Le plan-relief de Québec*. Cahiers des Fortifications de Québec, Parcs Canada.
- de La Cova, M.-A., 2019. Photographie et maquette chez Le Corbusier. Dialogues entre la création et la diffusion. *Les cahiers de la recherche architecturale, urbaine et paysagère*, 5, 1–25.
- Dries, T., 2023. Texel density. <https://www.beyondextent.com/deep-dives/deepdive-texeldensity>.
- ESRI, 2022. ArcGIS Maps SDK for Unity, Version 1.0. <https://developers.arcgis.com/unity/>.
- Fleury, P., Madeleine, S., 2011. Le "Plan de Rome" de Paul Bigot à l'université de Caen et son double virtuel: de l'objet patrimonial à l'outil scientifique. *In Situ: Revue des patrimoines*, 1–10.
- Fondation Le Corbusier, 2025. Collections. <https://www.fondationlecorbusier.fr/collections/>.
- Gao, J., Shi, Y., Cai, Y., 2023. Research on the Application of UAV Oblique Photogrammetry to Lilong Housing: Taking Meilan Lane as an Example. *The Int. Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLVIII-M-2-2023, 629–635.
- Griwodz, C., Gasparini, S., Calvet, L., Gurdjos, P., Castan, F., Maujean, B., Lillo, G. D., Lanthony, Y., 2021. AliceVision Meshroom: An open-source 3D Reconstruction Pipeline. *Procs. of the 12th ACM Multimedia Systems Conference (MMSys '21)*, ACM Press.
- Gruen, A., 2012. Development and Status of Image Matching in Photogrammetry. *The Photogrammetric Record*, 27(137), 36–57.
- Jiang, S., Jiang, C., Jiang, W., 2020. Efficient structure from motion for large-scale UAV images: A review and a comparison of SfM tools. *ISPRS Journal of Photogrammetry and Remote Sensing*, 167, 230–251.
- Lachambre, S., Lagarde, S., Jover, C., 2017. Photogrammetry Workflow. https://unity3d.com/files/solutions/photogrammetry/Unity-Photogrammetry-Workflow_2017-07_v2.pdf.
- Laporte, L., Perrault, M., 2000. *Une cité pour 33 296 habitants*. Musée Régional de Rimouski, Québec, Canada.
- Leberl, F., Meixner, P., Wendel, A., Irschara, A., 2012. Automated Photogrammetry for Three-Dimensional Models of Urban Spaces. *Optical engineering*, 51(2), 021117-1–021117-12.
- Madeleine, S., Fleury, P., 2024. *La Rome antique. Du Plan de Rome de Paul Bigot à la restitution virtuelle*. Presses universitaires de Caen.
- Menna, F., Rizzi, A., Nocerino, E., Remondino, F., Gruen, A., 2014. High Resolution 3D Modeling of the Behaim Globe. *The Int. Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXIX-B5, 451–456.
- Ministère des ressources naturelles et des forêts du Québec, 2025. Lidar: Modèles numériques (terrain, canopée, pente, courbes de niveau). <https://www.donneesquebec.ca/recherche/dataset/produits-derives-de-base-du-lidar>.
- Nicolae, C., Nocerino, E., Menna, F., Remondino, F., 2014. Photogrammetry Applied to Problematic Artefacts. *The Int. Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL-5, 451–456.
- Ray, S. F., 1988. *Applied Photographic Optics: Imaging Systems for Photography, Film and Video*. Focal Press.

Remondino, F., Elhakim, S., 2006. Image-Based 3D Modelling:
A Review. *The Photogrammetric Record*, 21(115), 269–291.

Schonberger, J. L., Frahm, J.-M., 2016. Structure-From-Motion
Revisited. *Procs of the IEEE Conference on Computer Vision
and Pattern Recognition*, 4104–4113.

Statham, W., 2018. Use of Photogrammetry in Video Games: A
Historical Overview. *Games and Culture*, 15(3), 289–307.

Strandberg, C. H., 1968. *Aerial Discovery Manual*. John Wiley
& Sons.

Unity Technologies, 2025. Unity Engine. <https://unity.com/products/unity-engine>.