

# Architectural Epistemology, Existing Buildings, and Single-Image Modeling

Mike Christenson, Kallie Self

## Abstract

This paper examines an alternative approach to creating a 3-D digital model of an existing building on the basis of a single photograph. Rather than prioritizing comprehensive coverage or geometric accuracy, the method aims to highlight how the modelmaking process itself can generate architecturally specific knowledge. The paper describes modeling the Nishiki Market in Kyoto using principles of projective geometry and reverse perspective. By problematizing the construction process, the method discloses how a model may embody subjective interpretations and choices. The paper argues that this approach legitimizes producing models counter to prevailing conventions, as the process itself constitutes a form of situational architectural understanding, registered through traces of the modelmaker and the photograph's perspective. Rather than foregrounding usefulness, this method values uncovering hidden assumptions and exposing the contingencies involved in constructing architectural knowledge.

Keywords: digital modeling, single image, existing buildings, architectural epistemology, representation, photography.

## 1. INTRODUCTION

Consider the process of constructing a digital model of an existing building. Research across disciplines tends to privilege existing-building model construction processes that prioritize metrics, such as comprehensiveness of coverage, semantic consistency, geometric accuracy, and support for interoperability (Edwards 2017; Janisio-Pawlowska 2021). The arguments in favor of these metrics are far from esoteric. For example, architects, engineers, contractors, and archaeologists all benefit

from having access to complete and accurate information for reuse or conservation projects (Jouan and Hallot 2020; Son, Bosché, and Kim 2015). Facilities managers depend on comprehensive and geometrically accurate existing-building models to support maintenance pre-planning, and to produce accurate predictive simulations (Cheng et al. 2016; Wong, Ge, and He 2018). Multidisciplinary teams, analyzing heterogeneous and complex historical assets, rely on robust existing-building models to exchange information across disciplinary

Stable URL: <https://arcc-journal.org/index.php/arccjournal/article/view/1216>  
DOI 10.17831/enq:arcc.v18i1.1216

Corresponding Author: Mike Christenson <mike001@umn.edu>

Published by the Architectural Research Centers Consortium under the terms of the Attribution-NonCommercial-ShareAlike 4.0 International license

boundaries, including non-geometrical information (Bruno and Roncella 2019). In a more subtle, but no less consequential way, geometrically accurate existing-building models can disclose perceptual aspects of architecture that are otherwise inaccessible or even actively suppressed in less geometrically accurate, more conventional representations (Piotrowski 2015).

This research stakes out a somewhat different path. Instead of emphasizing how accurate building models can be used across different fields, their compatibility, or their epistemological effects, the main focus here is on how the modelmaking process can generate knowledge. Specifically, the project examines the problem of making a model of an existing building using a single photo to create a “reverse perspective” (Agnello et al. 2022; Kane 1977).

Because various photographing subjects capture buildings from different perspectives, any given existing-building model constructed on the basis of a single photograph can be expected to embody, at most, a geometrically incomplete or fragmentary representation of the referent site. Incompleteness and fragmentation might conventionally be considered liabilities for existing-building models. Alternatively, such fragmentary representations may instead be positioned as opportunities to map the differences between “spaces of photography” corresponding to different photographing subjects (Christenson 2011). In a similar alternative approach, the present research is aimed at highlighting the architectural relevance of *emergent semantics*, that is, the production of meaningful relationships that depend on the specifics of modelmaking construction processes.

In this context, the work examines a method of existing-building model construction that problematizes the conditions of its own production, with the aim of enabling ways of generating knowledge that may otherwise be obscured. In this sense, the work is a part of the larger project of *architectural epistemology* (Christenson 2019). It focuses on using digital tools to explore how photography and architecture are connected and how they can inform the construction of architectural knowledge. Simply, the question is this: how, in constructing a digital model on the basis of a single image, do unique opportunities arise for the production of architecturally specific knowledge? To examine this question, we propose a constructive method invoking principles of projective geometry and reverse perspective.

### 1.1. State of the Art

The general problem in this project is to interpret a single photographic image to inform the construction

of a three-dimensional digital model. This problem has long been recognized as a highly ill-posed task in computer vision (Chen et al. 2020; Nishida et al. 2018). The issue is well-documented in the historical literature (Dzwierzynska 2016, 2017; Guillou et al. 2000; Horry et al. 1997; Liebowitz et al. 1999; Parodi and Piccoli 1996; Sturm and Maybank 1999; Van den Heuvel 1998). Despite its long history in computational research, it remains a pressing contemporary concern, particularly in structure-from-motion (SfM) research (Klodt and Vedaldi 2018; Schonberger et al. 2015) and in deep learning (Fu et al. 2021). The task’s computational difficulty is largely related to the phenomenon of visual occlusion, itself a well-known problem in computer vision (Chandel and Vatta 2015; Hoiem et al. 2005; Hoiem et al. 2007). Depth estimation constitutes another well-known major obstacle to computational solutions (Mertan et al. 2022). Nishida et al. (2018) suggest the “significant user effort,” i.e., manual effort as distinct from automated procedures, that is required to extract semantically meaningful information from single images.

Faced with the apparent challenges and difficulties in single-image reconstruction, the approach would hardly seem to recommend itself whenever simpler, faster, or more accurate techniques are available to existing-building model researchers. If a researcher can visit a site in person, they might choose to use methods that provide wider coverage and more precise data, such as scan-to-BIM (D’Amico and Curra 2017). Yet, with this as context, this project aims to explore the subjective nature of digital model-making processes, uncovering hidden approach-specific values, temporarily setting aside the pursuit of geometric accuracy.

## 2. PROCESS

### 2.1. Candidate Images

In this section, we aim to identify the optimal qualities of candidate images for use in our method.

Our method of single-image reconstruction begins with a photograph (Figure 1) and results in a three-dimensional wireframe model of the space depicted in the photograph. We chose to focus our efforts on minimally distorted photographic images. Within this scope, we recognize three closely related processes for achieving reconstruction, i.e., single-point perspective, two-point perspective, and three-point perspective. This paper narrows its focus to specifically single-point perspective images. Photographic images that incorporate visible lens distortion (e.g., barrel distortion or pincushion distortion) are understood to visibly “warp” the image, depicting straight lines as curves (López-Antequera et al. 2019). Images of this kind are not addressed in this paper.



Figure 1: Photograph of Nishiki Market.



Figure 2: Horizontal centering of vanishing point in photograph of Nishiki Market.

The following points summarize the requirements for a candidate image:

1. **Image Centering.** Centering the image horizontally is crucial for this method: the vanishing point must be centered horizontally within the image, as shown in Figure 2. Vertical centering is counterproductive, as will be discussed.
2. **Vertical Line Alignment:** Ensure that vertical lines in the image are truly vertical, i.e., not subject to perspectival convergence. As mentioned previously, lens distortion can result in the depiction of straight lines as curved, i.e., bent inward, as seen in Figure 3.
3. **Horizontal Line Alignment:** Horizontal lines must be truly horizontal. Two aspects matter: the camera angle causing convergence, which can be mitigated by finding undistorted images, and a perpendicular image angle to the ground (i.e., a lack of roll).
4. **Vertical Height:** An ideal candidate image is produced by a camera that is placed at approximate eye level relative to the scene (i.e., about 5 or 6 feet off the ground). There are two major effects

A similar effect can be caused by an angled camera (tilted up or down). Candidate images should be free of this distortion to ensure accurate alignment.





**Figure 3:** Nishiki Market viewed with a tilted camera, resulting in perspectival convergence of vertical lines. Photo: Tom Bricker, [www.travelcaffeine.com](http://www.travelcaffeine.com).

5. **Reference Availability:** An ideal candidate image will be contextualized by additional reference information, allowing the modelmaker to explore the space in comparison to other information, and in order to assess sizes and relative positioning. Having only one image to work from, and nothing else, makes this task significantly more difficult. This also raises a problem we discuss in the conclusion.
6. **Ground Visibility.** When constructing the wireframe, having clear ground lines (e.g., the edges of a floor) helps with alignment and placing gravity-affected objects. These lines also aid in scaling the model. Compare Figure 2, in which the ground is visible, with Figure 3, in which the ground is not visible.
7. **Clear Lines.** Clear lines in the space, especially those along walls or roofs pointing to the vanishing point, simplify the model-making process.
8. **Repetition of features at regular intervals** aids in estimating distances and scale, although it becomes less critical when other requirements, like vertical height and ground visibility are met.
9. **High Resolution.** Higher-resolution images make it easier to place lines accurately at the space's edges and corners.
10. **Adequate Lighting.** Brighter images provide clearer edges, making line drawing easier.
11. **Visible End of Space:** Ensure that the end of the space is visible or within the image frame. If the bottom edge of the space is below the image frame, the overall frame for reference cannot be fully drawn.

The requirements listed here are for optimal application of the method described below. The method can be followed even if all conditions are not met, but it will be less effective and far more difficult. The practicalities of why one might use this process, considering these requirements, are discussed more in the conclusion.

## 2.2. Methods

Nishiki Market consists of a narrow street, several blocks in length, with a sheltering structure covered by a multicolored roof window. The modeling method begins with a photograph of the Nishiki Market in Kyoto, Japan, focused on a typical area, i.e., one that shares visual characteristics with the market as a whole (Figure 1).

The photographed space can be broadly characterized as a simple rectangular volume, designated  $V$ . Only a portion of  $V$  is visible in the original photograph (Figure 4).

Figure 4(a) illustrates the original photograph. In Figure 4(b), broadly-apparent features in the photograph, e.g., edges and corners, are manually traced. These traced lines are then extended past the photograph's edges in Figure 4(c) to produce an image of the rectangular volume  $V$ . However, because the referent space, as photographed, is evidently not a simple or straightforward rectangular box, the tracing process required simplifying assumptions. For example, what appear as parallel and perpendicular lines, subject to the visual phenomenon of perspective convergence, are assumed to be sufficient to determine a horizon and a vanishing point (Figure 5). These preliminary determinations provided a basis for constructing, in three-dimensional modeled space, a bounding box  $V_1$ .

Creating a reverse perspective using a traditional projective-geometry method requires establishing a specific viewing point in space and a flat surface perpendicular to the line connecting the viewing point and the apparent vanishing point (Dzwierzynska 2016). The photograph itself, placed in the Rhino environment as a rectangular, bounded planar surface, makes an obvious candidate for the surface, i.e., the *picture plane*. Having already determined where the vanishing point is in relation to the photograph, it is simple to draw a line of sight ( $L$ ) from that vanishing point ( $P_2$ ). This line proceeds perpendicular from the photograph, and the station point ( $P_1$ ) must exist along that line (Figure 6).

However, there seems to be no obvious way to unambiguously locate the station point on  $L$ : the distance  $D$ , measured along  $L$  between the station point ( $P_1$ ) and the vanishing point ( $P_2$ ), itself located on the picture plane, is a variable with no obvious maximum value. Nevertheless, the placement of  $P_1$  cannot be

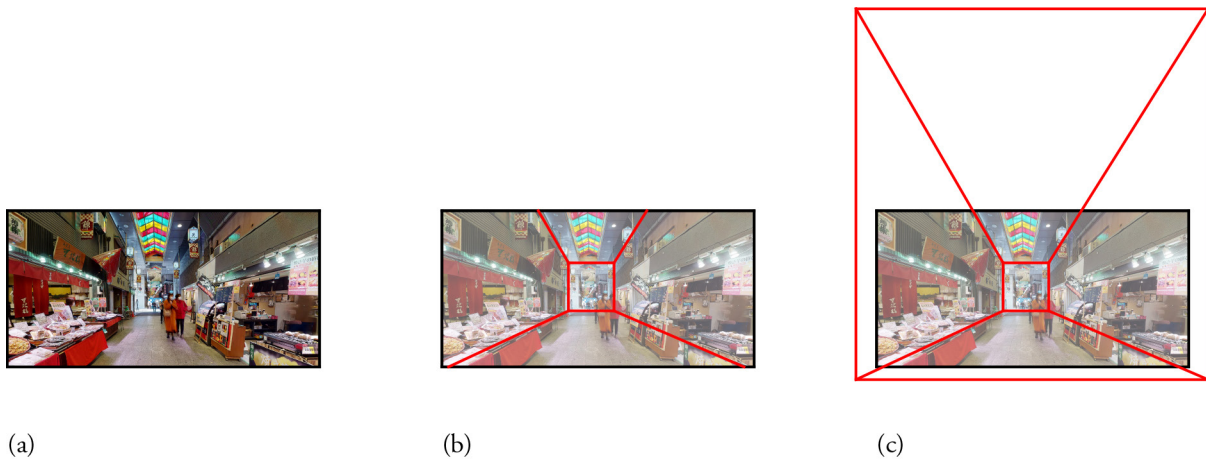
arbitrary, as it has observable consequences on the model's configuration: if  $P_1$  is "too close" to  $P_2$ , the result is a visibly foreshortened model; while placing  $P_1$  "too far" from  $P_2$ , results in a model that appears excessively elongated (Figure 7).

Considered from a purely topological standpoint, fixing the location of  $P_1$  relative to  $P_2$  does not modify the subsequent stages of the modeling process. However, if the final model is to reflect not only topological, but also proportional consistencies with the referent space, an external factor must be introduced to determine the location of  $P_1$ , for example, a distance derived from ground-truth data. For example, if the real-world length of a feature visible in the photograph was known, this data could be used to unambiguously choose between the alternatives illustrated in Figure 4. Without such information, the location of  $P_1$  may be left as a variable parameter, resulting in a geometrically flexible model (as suggested by Figure 7). At this stage in the modelmaking process, the location of  $P_1$  was fixed at a reasonably intermediate working value, as shown in Figure 5(b), allowing the model to remain flexible and potentially responsive to new information.

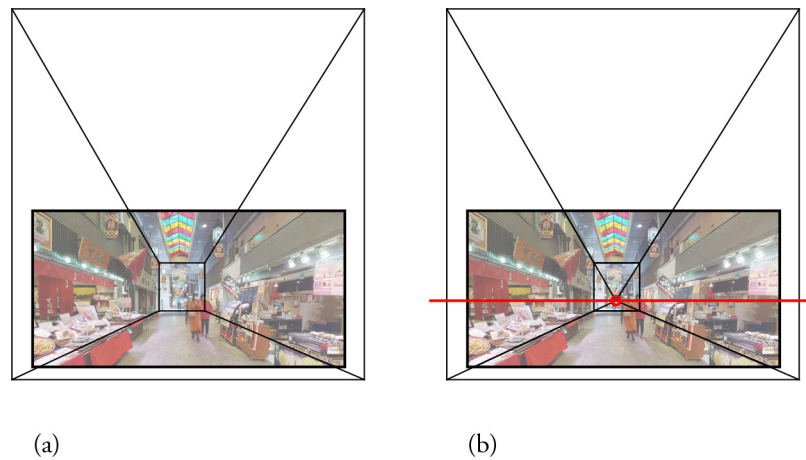
Having determined a working location for  $P_1$  – i.e., the station point – lines from  $P_1$  were projected through the picture plane to establish the near surface of  $V_1$ . This involves an additional determination, i.e., the distance  $d$  from  $P_1$  to the near surface. Testing possible (arbitrary) locations for the near plane quickly reveals that its location has an equiproportional impact on the resulting model, i.e., changing the location of the near plane maintains the model proportions (Figure 8), a consequence easily attributed to triangle similarity. The location of the modeled space's distant plane follows as a direct consequence of this determination.

Thus, this determination is inconsequential to the subsequent modeling process; a location may be arbitrarily selected for the near plane, in turn, fixing the location of the far plane. If ground-truth data, such as measurements of width, length, or height, were accessible, this information could be used now or later in the process to scale the entire model correspondingly. Without such data, the process would result in a consistently proportionate, but difficult-to-measure model. Alternatively, a working unit could be chosen with the aim of maintaining future flexibility. (Using specific units, even if estimated, can help to provide a "reality check" on the model during the construction process.)

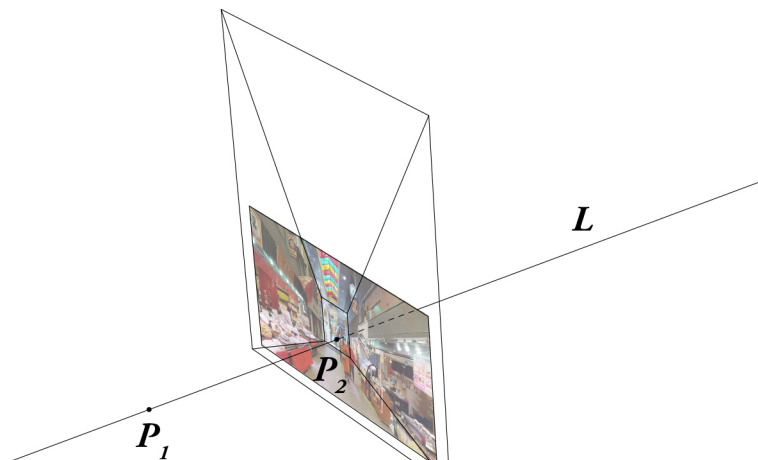
Once  $V_1$  is established in the 3-D environment, the process of constructing individual objects within  $V_1$ , and surfaces parallel and perpendicular to  $V_1$ 's surfaces, can begin. Objects and surfaces of interest must be



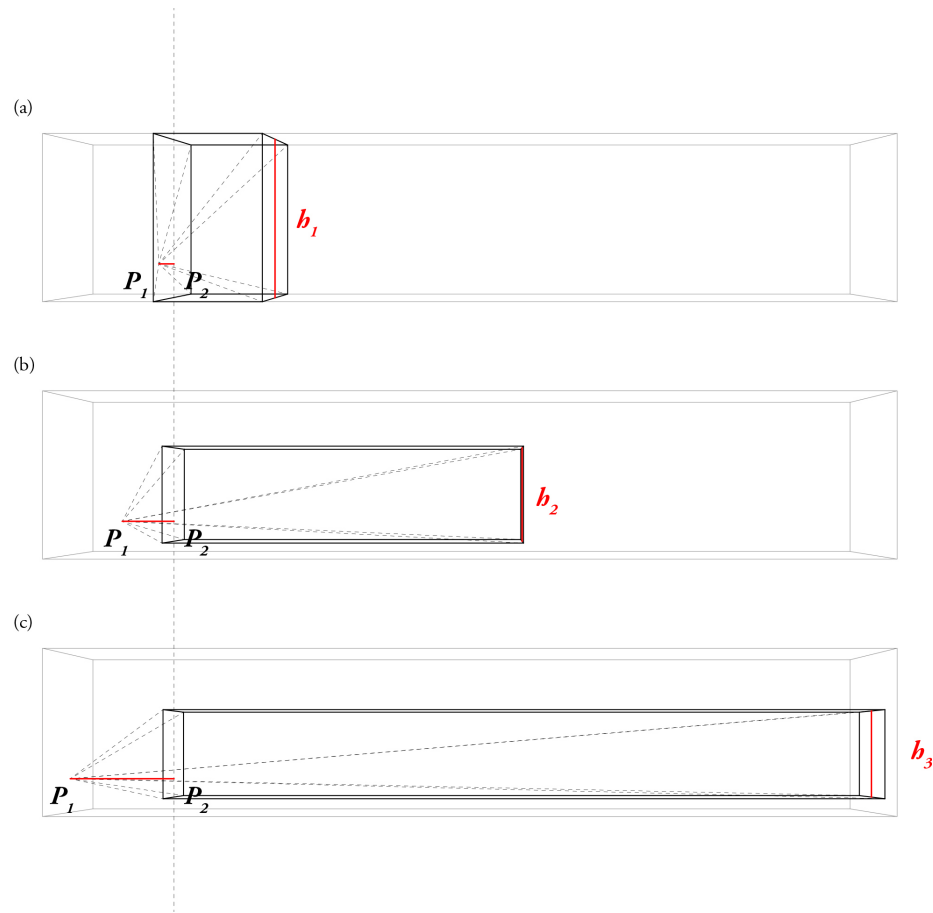
**Figure 4:** (a). Original photograph of Nishiki Market. (b) Manual tracing of broadly-apparent features. (c) Extension of traced features beyond photograph edges, producing 2-D image of rectangular volume V.



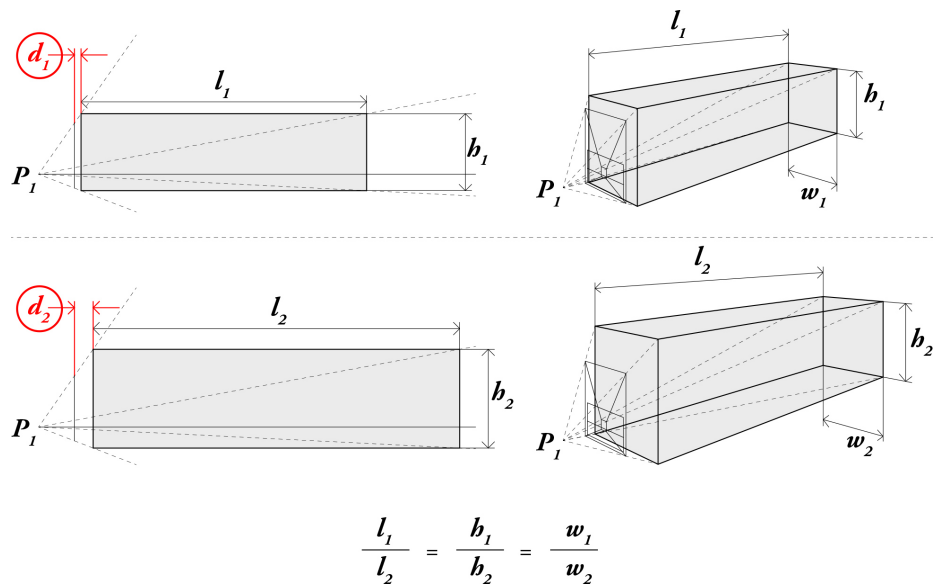
**Figure 5:** Location of apparent horizon and vanishing point.



**Figure 6:** Establishing station point  $P_1$  on line of sight  $L$ .



**Figure 7:** Changing the location of  $P_1$  changes the proportions of the modeled space. In (a), with  $P_1$  and  $P_2$  “too close,” the model appears visibly foreshortened and excessively tall for its length. In (c), with  $P_1$  and  $P_2$  “too far” from each other, the model appears excessively elongated, and too short. (b) appears to strike a reasonable balance.



**Figure 8:** Varying the near-plane location has an equiproportional effect on the model.

sufficiently distinct to be easily outlined, such that their outlines can be projected onto the overall bounding box. The process of finding these outlines is exactly the problem of occlusion as mentioned earlier. Figure 9 illustrates the basic process, using one of the human figures visible in the photograph.

We begin by manually outlining the figure's silhouette within the photograph and establishing a planar rectangular bounding box  $f_1$  around the silhouette. The baseline of  $f_1$  is extended until it intersects the traced line representing the edge of  $V_1$ , establishing point  $P_3$  (Figure 9), all on the surface corresponding to the original photograph.

Next, a ray is produced in three dimensions, originating at the station point  $P_1$ , extending to the previously determined point  $P_3$  (Figure 10).

Figure 11 shows the ray extended (in three dimensions) past  $P_3$ , intersecting the edge of  $V_1$ , hence establishing point  $P_4$ .  $P_4$  is used as the baseline to establish line  $l$ , drawn perpendicular to the sides of  $V_1$ .

Figure 12 illustrates the production of four additional rays emanating from the station point  $P_1$ , proceeding through the four corner points of  $f_1$ , and extending until they intersect with the vertical plane through line  $l$ , establishing frame  $f_2$ . Frame  $f_2$  is, thus, a projection, in 3-D space, of frame  $f_1$ . (Recall that  $f_1$  is the rectangular bounding box of the original figure traced from the photograph).

Frame  $f_2$  is sufficient to locate the figure in 3-D space. Thus, a copy of the traced figure is placed within frame  $f_2$  (Figure 13).

This process is repeated for other elements in the photograph to produce a three-dimensional wireframe model of the observable space (Figure 14).

### 3. RESULTS AND DISCUSSION

#### 3.1. The Completed Model

The completed model consists of a volumetric wireframe trace representing the edges of the volume  $V_1$ , including lines representing edges of major openings, projections, and subsidiary volumes, such as the roof structure. Additionally, the original photograph is superimposed on the picture plane, and the station point ( $P_1$ ) is shown along with lines of sight emanating from it through the picture plane. Rhino's ontology allows model elements, such as lines, to be categorized into layers. Thus, elements in the model are placed into layers corresponding to whether they represent the edges of observable features, such as the edges of  $V_1$ , or conjectured features, such as the extension of  $V_1$  beyond what is visibly apparent in the photograph.

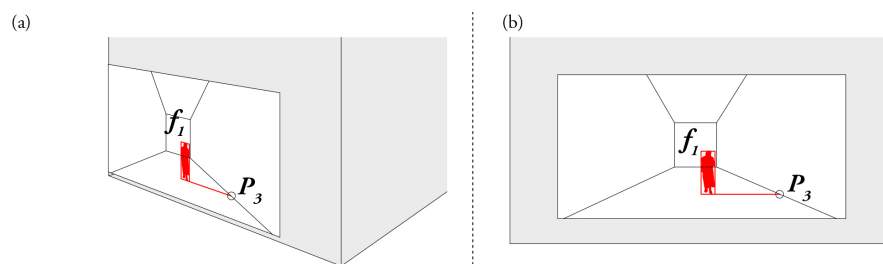
The model can be said to exhibit biases, inflections, and interpretive decisions originating from several sources. Specifically, there are *method biases* related to the tools and procedures used in model construction. There are also interpretive decisions associated with visibility, such as occlusion and view geometry, reflecting the hypothesized vantage point of the photographer within the model. Finally, we recognize that the completed model differs substantially in quality and kind from a model resulting from a remote-sensing process, such as photogrammetry or point-cloud scanning. We discuss these points further in the following sections.

#### 3.2. Semantic Interpretation of Method Biases

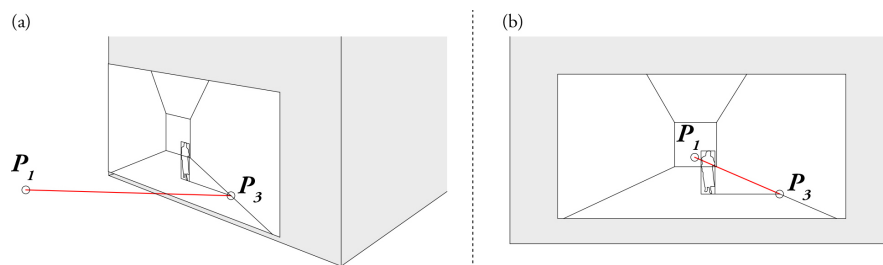
In general, semantic relationships can arise from disciplinary knowledge related to the tools and methods of a model construction process. Semantic relationships that arise in this way can be said to involve method biases, that is, biases that derive from a particular working method. In the broadest context, method biases simply reflect the widely-acknowledged view that tools used for existing-building model construction are not neutral; they necessarily embody unique capabilities and limitations (Castelo-Branco, Caetano, and Leitão 2022; Christenson 2019). In the case of the current project, method biases concern software (specifically Rhino) and the conventions of perspective drawing.

The presence of software-related method biases means that the process of constructing the model is one of discerning relevant information from the process of interpreting the photograph, determining its appropriate organization, while operating within the capabilities and limitations associated with the software. Critically to our process, Rhino's work environment, like that of AutoCAD and unlike that of Revit, allows for the construction of geometry unassociated with conventional building semantics. In other words, Rhino allows, but does not require, model elements to link to real-world objects or categories. For example, our Rhino model includes modeled entities categorized as *curves* (a category which includes straight line segments). In Rhino, curves can be drawn to represent real-world objects, such as walls and openings, and our model makes extensive use of this interpretation. Perhaps most significantly, curves can be drawn to indicate physical or conceptual connections that may not necessarily exist as real-world objects. Our model includes this use of curves in almost every step of construction, as when lines of sight are projected from a station point. In contrast, Revit, consistent with a BIM approach, enforces a clear structure of families, types, and instances, creating a "semantic layer" in the representation (Castelo-Branco, Caetano, and Leitão 2022). While Revit allows for the construction of entities that do not directly correspond to real-world objects (e.g., guidelines and reference planes), as a general rule, Revit objects are always related to buildings.

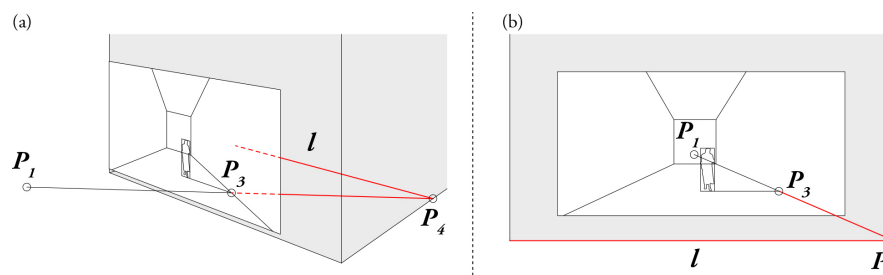




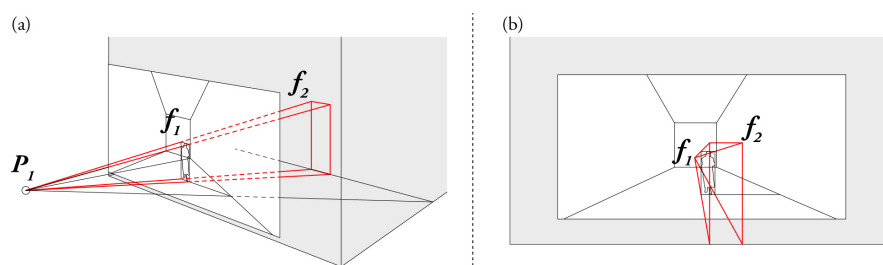
**Figure 9:** Tracing the figure silhouette to establish frame  $f_1$  and point  $P_3$ . (a) Perspective view of model. (b) Front view of model.



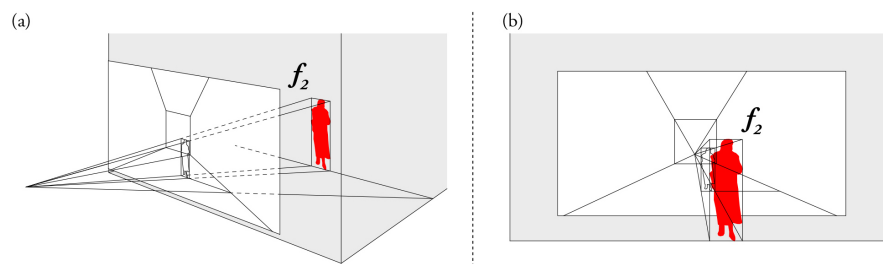
**Figure 10:** Producing a ray from  $P_1$  to  $P_3$ . (a) Perspective view of model. (b) Front view of model.



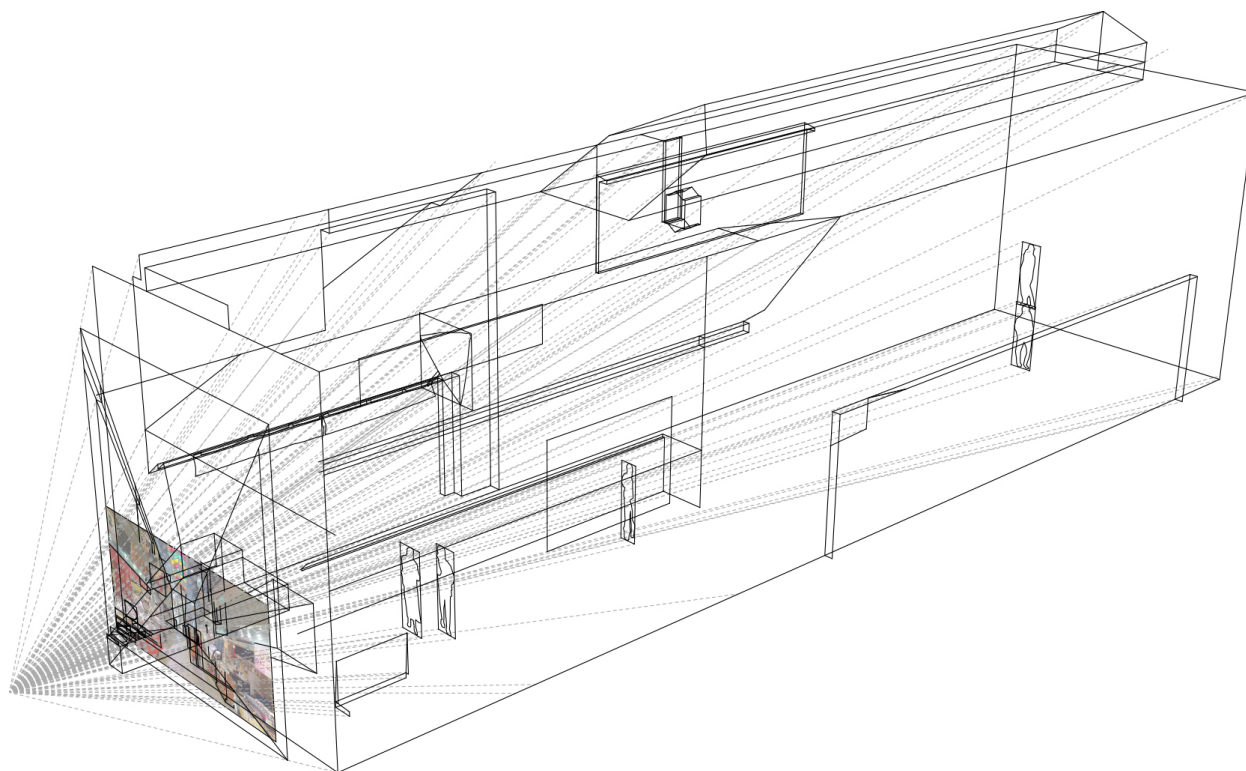
**Figure 11:** Establishing point  $P_4$  and baseline  $l$ . (a) Perspective view of model. (b) Front view of model.



**Figure 12:** Establishing frame  $f_2$  as a projection of frame  $f_1$ . (a) Perspective view of model. (b) Front view of model.



**Figure 13:** Placement of human-figure silhouette within 3-D space. (a) Perspective view of model. (b) Front view of model.



**Figure 14:** Completed model.

Similarly, method biases arise with the drafting of a perspective view. When constructing the model, assumptions must be made about how the photographed scene relates to conventions such as a singular station point, a horizon line, a picture plane, and a vanishing point. These assumptions have implications for how features, such as parallel lines, qualities of perpendicularity and flatness of surfaces, and other characteristics of the modeled space are interpreted for modelmaking purposes. More specifically, as we worked to trace the source photograph, our interpretations were guided by our own assumptions in at least two ways. First, we made assumptions about perspectival projections, e.g., about how parallel lines would appear in a perspective projection (convergent on a vanishing point), or how objects could visually occlude each other (according to their depth in the photographed space). Second, we made assumptions about objects and surfaces in the real world, based on how they appeared in the photograph, e.g., surfaces that appeared flat were traced and modeled as flat, and surfaces that appeared mutually perpendicular were modeled, so as to be actually perpendicular.

Finally, all of these assumptions are conditioned by Rhino's drafting and modeling environment. As one example, consider that the modeling procedure involved creating an overall bounding box to serve as

a reference for locating individual points within the scene. The choice of a bounding box depends on both the image's particulars and the assumptions concerning the implied rectangular volume. In Rhino, locating individual points and mapping them to the bounding box becomes a geometrical exercise rather than an exercise in associating visible features with nameable "real world" objects.

### 3.3. Semantic Interpretation of Boundaries and Occlusions

Spatially dispersed entities share a capacity to visually occlude; it is this that visually distinguishes them as discrete entities, giving them individuality and identity. Occlusion relationships, as seen in the original photograph, arise as a consequence of the photographing subject's viewpoint, the direction of view, and the angle of view, with respect to the spaces and objects being photographed. Throughout the model-construction process, the geometric specifics of occlusion relationships were interpreted according to a manual outline-detection and interpretation process. Yet, to the extent that occluding objects are individually identifiable, they may or may not be identifiable with pre-existing semantic categories. For example, some entities within the model can be unambiguously labeled as tables, windows, or people, while others



Figure 15: Vector tracing of pixel image.

can only be classified as shapes. In the overall context of the process, the fact of an occluding shape being semantically identifiable as “a person” or “a tray of fish” is inconsequential in comparison to the fact of occlusion itself. Stated differently, objects in the photograph do not require unambiguous semantic identification in order to be outlined and placed in the model: decisions made during the modelmaking process about apparent object boundaries are not necessarily informed by unambiguous knowledge about the objects being traced.

Furthermore, modelmaking decisions are influenced not only by the modelmaker’s need to interpret objects in a meaningful way, but also because of the need to interpret a pixel-based image into a vector trace, introducing a level of geometric accuracy that isn’t present in the original photograph (Figure 15).

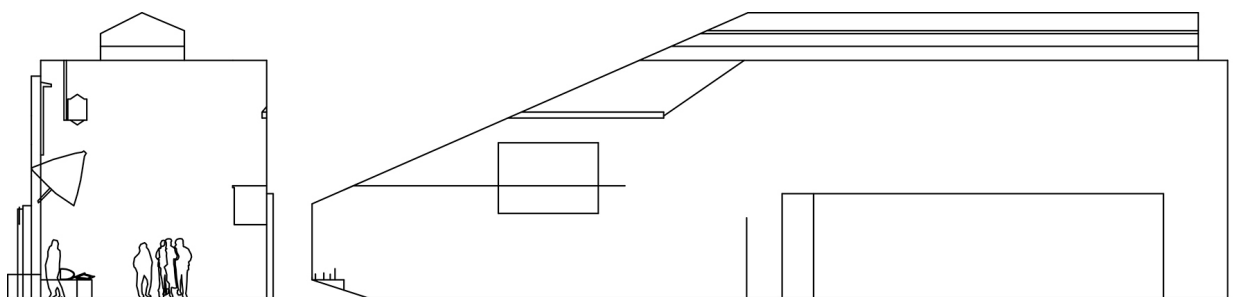
In all of these ways, the modeling process asserts independence from preexisting semantic classifications for objects (such as furniture, walls, people, and market goods), replacing them not by new classifications, but by procedural questions related to identifiability of outline. Considered from an architectural perspective, occlusion is epistemologically interesting not because of its technical or procedural specifics, but rather because of how it uniquely places the modeled space and objects into discourse.



### 3.4. Semantic Interpretation of the Photographer’s Viewpoint

From an epistemological perspective, the knowledge that emerges through the model construction process is *situated* knowledge, i.e., its specificity depends on the photographer’s viewpoint. To illustrate this, consider that the model construction process is contingent on the geometric projection of presumed “lines of sight” emanating from a hypothesized station point, the modeled location of which is hypothesized on the basis of (a) our interpretation of the visual evidence in the photograph, and (b) our interpretation of volumetric proportionality as described above and as illustrated in Figure 7. Point (a) required us to interpret evidence in terms of perspectival construction (i.e., making decisions about the horizon line, vanishing point, etc.). The second point, (b), does not impact the model’s *topological* configuration, but does impact its *proportional* configuration; hence, it has a direct impact on the model’s measurable correspondence with the market space.

Furthermore, the model’s construction process depends on our interpretation of visible clues within the photograph (e.g., recognizable objects, shapes, shadows). As such, the process reveals relationships between what are known to be spatially dispersed objects and their geometrical representations. Some of these relationships—such as those allowing us to



**Figure 16:** Derived views (elevations) of the finished model.

determine observed height with respect to apparent ground level—arise only when the model is observed from a specific perspective. Consequently, the model is a consequence of our interpretations based on a specific (hypothesized) viewpoint, as registered by visible clues and issues of perspectival construction such as eye level, center line, and other parameters.

It follows that the model—and more significantly, the knowledge constructed during the modelmaking process—would be constituted differently if the scene were photographed from a different station point. Objects otherwise obscured would become visible in whole or in part, and vice versa; shapes and shadows that appeared ambiguous with respect to real-world interpretations could resolve themselves to simple explanations, and new ambiguities could arise. In short, repeating the process using a photograph taken from even a slightly different vantage point would result in a different model.

Significantly, the completed model may be viewed in ways that differ from the original photographer's perspective. As in BIM software, this ability makes it possible to produce an arbitrarily large number of *derived views*, such as perspectives, plans, and elevations. Derived views of the Nishiki Market model reveal the model to be *inflected* in ways that emphasize and privilege the photographing subject's viewpoint (Figure 16).

An obvious example of this inflection appears in Figure 16, where the side view of the model (at right in the figure) incorporates a prominent diagonal line that effectively represents the boundary of vision as recorded in the original photograph. Anything above and to the left of this diagonal line is not visible in the photograph, and, hence, must be modeled purely conjecturally, if indeed at all.

In such ways, the derived views embody a trace of the model-construction process; they highlight the presence of both modelmaker and photographer. Stated differently, the derived views are not *views of*

*the space*; they are *views of the photographer's view of the space as interpreted by the modelmaker*. In this way, the model-construction process produces knowledge that is specific not only to Nishiki Market, but to the photograph, the photographer, and the modelmaker. By making the model-construction process accessible in a graphical or nonverbal way, the derived views are susceptible to examination and critique in ways that apply equally well to design processes. That is to say, they can be examined and critiqued as products of judgment and prioritization.

### 3.5. Process and Conventional Assumptions

We can reasonably ask—why build an architectural model if it does not accurately and transparently incorporate the “real” conditions of the space being modeled? Of course, the answer is simple: models are always “value-laden tools of representation” that reflect observations and interpretations of spatial conditions underlying architecture's conception and realization (Pérez-Gómez 2005). This phenomenon may be more or less obvious in specific cases. A digital model based on a dense point cloud, capable of supporting fully immersive simulations, may not present itself as an “interpretation” of the referent space; nevertheless, of course, that is exactly what it is.

We could, of course, approach the model construction process differently, so as to maximize its alignment with conventional assumptions. Specifically, instead of using a non-provenanced photographic image, the site could be photographed afresh under controlled conditions. Recording the camera's position relative to a sufficient number of spatially dispersed reference points, and maximizing photographic resolution, could help to ensure greater accuracy. Tracing, whether done manually or automatically, could be supplemented or guided by contextual data (e.g., other photographs or direct observation) in an effort to accurately discern precise shapes and other details. Or, as suggested earlier in this paper, the single-image approach could be passed over entirely in favor of more comprehensive and accurate methods, for example, choosing instead



to produce a high-resolution point cloud. However, the case studied here is precisely limited to what can be interpreted from a single photograph. The process is at no point supplemented with additional photographs, although these are easy to find in abundance online. Similarly, there are no significant obstacles in visiting the site to directly observe, measure, photograph, or scan it to a level of detail and precision that meets any desired standard. In other words, the “input” information is deliberately limited to a practicable minimum in order to test a broader question, stated at the outset, i.e., how might architecturally specific knowledge be constructed on the basis of a single image? In this way, what results is a model for discovery or a model for developing new processes, and not a model of the market in a conventional sense. This observation simply restates a distinction between the model’s epistemic function (what it is “for”) and its representational nature (what it is “of”) (Cannaerts 2009; Gouvea and Passmore 2017).

#### 4. CONCLUSION

Our model of Nishiki Market diverges from conventional expectations for existing-building models, in that it does not aim to precisely depict the site’s measured dimensions. Rather, the model constitutes a type of “tangible speculation” (Graves 1977) that is simultaneously precise and speculative. In this way, the model is neither definitive nor comprehensive, but rather a step in a process of architecturally specific understanding—a process that requires ongoing negotiation between expectations and observations, accompanied by a willingness to embrace inconsistencies and contradictions.

We have discussed how conventional expectations for digital models of existing buildings include comprehensiveness of coverage, semantic consistency, geometric accuracy, and support for interoperability. These expectations, as we have acknowledged, are not esoteric or impractical: they are genuinely rooted in a cross-disciplinary desire for reliable information exchange. It is precisely the expectation of reliability that makes it possible for digital models of existing buildings to be used effectively in collaborative projects (e.g., renovation, restoration, documentation, etc.). Conventional approaches to existing-building modeling, understood in this way, are largely directed towards usefulness or utility, whether in the context of speculative reconstruction, practical efforts toward conservation, or facilities management. In this context, existing-building model qualities of interoperability, interactivity, and immersion are recognized as paramount (Banfi 2021).

Nevertheless, there are epistemological risks inherent in the conventional approach as we have defined it. For example, insisting on comprehensiveness of coverage

could imply that *knowledge* is in some way unaffected by *ways of knowing*, i.e., that the model reflects a reality “out there” waiting only to be disclosed, irrespective of whether that reality comes to be known through direct experience, photographs, or drawings. Yet, existing-building modeling projects can reveal that different ways of knowing are profoundly influential on processes and results (Sullivan and Snyder 2017; Leung, Davies, and Ching 2018). Similarly, an insistence on geometric accuracy could have the effect of delegitimizing models that may be knowingly *inaccurate* with respect to highly precise measurements. The need for such models can arise in cases where researchers are interested in relating the perceptual effects of different forms of representation, or when they study transformations of existing structures over time (Adami et al. 2021: 170–18).

In these ways, considered relative to more conventional approaches for constructing existing-building models, our approach is attitudinally distinguished toward accuracy and consistency. Our attitude effectively abandons conventional expectations for the finished model. For example, our modelmaking process proceeded with no expectation that the completed model would be useful for producing photorealistic renderings. If there was any expectation for the finished model—as distinct from the process of constructing the model—it was only that it should operate to disclose specific tactical limits on the construction of architectural knowledge. In short, our expectations for the finished model were only that it should visibly register the decision-making process leading to its final form. Indeed, this is exactly what the London Charter seeks to recognize through its emphasis on paradata, seeking to make transparent the model-making decision process with the aim of highlighting existing-building models’ scientific legitimacy (Bentkowska-Kafel 2016; Denard 2016).

While a single-image approach, like that presented here, may be appropriate or even necessary in the context of historical reconstruction (i.e., in cases where only limited photographic information is available concerning a subject site), our goal in this work was theoretical rather than practical. Through this project, we aim to call conventional assumptions into question, and to legitimize ways of producing existing-building models that might be counter to prevailing approaches.

Yet, questions remain: To what extent does an existing-building model gain its legitimacy due to its usefulness? In what ways does the usefulness or utility of a final model relate to the assumptions made during the model-construction process? To what extent does usefulness emerge from model-construction processes as distinct from the product of the final model? Supposing an

existing-building model is constructed according to an esoteric process, resulting in a highly subjective model, how useful can we expect the model to be—and hence how should we gauge its legitimacy? Once the model-construction process itself is acknowledged to be capable of producing a legitimate form of knowledge, then the question becomes whether (and how) this knowledge informs the question of usefulness: is an existing-building model legitimately “useful” for research if it results in a newfound perception or a novel way of seeing?

## REFERENCES

- Adami, Andrea, Luigi Fregonese, Simone Helena Tanoue Vizioli, Laura Taffurelli, Daniele Treccani, Olga Rosignoli, and Jacopo Helder. 2021. “HE. SU. TECH. Laboratory: Academic Research Combined with the Practice Applied in the Documentation, Representation, and Preservation of the Architectural Heritage. The Case of the Palladio Villa Forni Cerato.” *Gestão & Tecnologia de Projetos* 16, no. 3. <https://doi.org/10.11606/gtp.v16i3.172008>.
- Agnello, Fabrizio, Manuela Milone, and Laura Barrale. 2022. “From Image to Investigation: 3D Reconstruction with Perspective Restitution.” In *Congreso Internacional de Expresión Gráfica Arquitectónica*, edited by M.A. Ródenas-López, J. Calvo-López, and M. Salcedo-Galera, 83–92. Cham: Springer International Publishing. [https://doi.org/10.1007/978-3-031-04703-9\\_9](https://doi.org/10.1007/978-3-031-04703-9_9).
- Banfi, Fabrizio. 2021. “The Evolution of Interactivity, Immersion and Interoperability in HBIM: Digital Model Uses, VR and AR for Built Cultural Heritage.” *ISPRS International Journal of Geo-information* 10, no. 10: 685. <https://doi.org/10.3390/ijgi10100685>.
- Bentkowska-Kafel, Anna. 2016. “Processual Scholia: The Importance of Paradata in Heritage Visualization.” In *Paradata and Transparency in Virtual Heritage*, edited by Anna Bentkowska-Kafel, Hugh Denard, and Drew Baker, 245–259. New York: Routledge.
- Bruno, Nazarena, and Riccardo Roncella. 2019. “HBIM for Conservation: A New Proposal for Information Modeling.” *Remote Sensing* 11, no. 15: 1751. <https://doi.org/10.3390/rs11151751>.
- Cannaerts, Corneel. 2009. “Models of / Models for Architecture: Physical and Digital Modelling in Early Design Stages.” In *Proceedings of the 27th Conference on Education and Research in Computer Aided Architectural Design in Europe (eCAADe)*, edited by G. Çağdaş and B. Çolakoğlu, 781–786. Istanbul: eCAADe. <https://doi.org/10.52842/conf.ecaade.2009.781>.
- Castelo-Branco, Renata, Inês Caetano, and António Leitão. 2022. “Digital Representation Methods: The Case of Algorithmic Design.” *Frontiers of Architectural Research* 11, no. 3: 527–541. <https://doi.org/10.1016/j.foar.2021.12.008>.
- Chandel, Himanshu, and Sonia Vatta. 2015. “Occlusion Detection and Handling: A Review.” *International Journal of Computer Applications* 120, no. 10: 33–38. <https://doi.org/10.5120/21264-3857>.
- Chen, Xin, Yuwei Li, Xi Luo, Tianjia Shao, Jingyi Yu, Kun Zhou, and Youyi Zheng. 2018. “Autosweep: Recovering 3D Editable Objects from a Single Photograph.” *IEEE Transactions on Visualization and Computer Graphics* 26, no. 3: 1466–1475. <https://doi.org/10.1109/TVCG.2018.2871190>.
- Cheng, Jack C. P., Weiwei Chen, Yi Tan, and Mingzhu Wang. 2016. “A BIM-Based Decision Support System Framework for Predictive Maintenance Management of Building Facilities.” In *Proceedings of the 16th International Conference on Computing in Civil and Building Engineering*, edited by N. Yabuki and K. Makanae, 711–718. Osaka: The International Society for Computing in Civil and Building Engineering.
- Christenson, Mike. 2011. “On the Architectural Structure of Photographic Space.” *Architectural Science Review* 54, no. 2: 93–100. <https://doi.org/10.1080/00038628.2011.582365>.
- Christenson, Mike. 2019. *Theories and Practices of Architectural Representation*. New York: Routledge. <https://doi.org/10.4324/9781351677790>.
- D’Amico, Alessandro, and Edoardo Curra. 2017. “From TSL Survey to HBIM, Issues on Survey and Information Modeling Implementation for the Built Heritage. The Case Study of the Temple of Baccho, Ravello.” In *Proceedings of the 35th International Conference on Education and Research in Computer Aided Architectural Design in Europe*, edited by A. Fioravanti, S. Cursi, S. Elahmar, S. Gargaro, G. Loffreda, G. Novembri, and A. Trento, 39–48. Rome: Gangemi Editore SpA International. <https://doi.org/10.52842/conf.ecaade.2017.2.039>.
- Denard, Hugh. 2016. “A New Introduction to the London Charter.” In *Paradata and Transparency in Virtual*

*Heritage*, edited by Anna Bentkowska-Kafel, Hugh Denard, and Drew Baker, 83-98. New York: Routledge.

Dzwierzynska, Jolanta. 2016. "Reconstructing Architectural Environment from the Perspective Image." *Procedia Engineering* 161: 1445-1451. <https://doi.org/10.1016/j.proeng.2016.08.608>.

Dzwierzynska, Jolanta. 2017. "Single-Image-Based Modelling Architecture from a Historical Photograph." In *IOP Conference Series: Materials Science and Engineering* 245, no. 6: 062015. <https://doi.org/10.1088/1757-899X/245/6/062015>.

Edwards, John. 2017. "It's BIM—But Not as We Know It!" In *Heritage Building Information Modelling*, edited by Y. Arayici, J. Counsell, L. Mahdjoubi, G. Nagy, S. Z. Hawwās, and K. Dweidarpp, 6-14. Abingdon (UK): Routledge.

Fu, Kui, Jiansheng Peng, Qiwen He, and Hanxiao Zhang. 2021. "Single Image 3D Object Reconstruction Based on Deep Learning: A Review." *Multimedia Tools and Applications* 80: 463–498. <https://doi.org/10.1007/s11042-020-09722-8>.

Gouvea, Julia, and Cynthia Passmore. 2017. "'Models of' versus 'Models for': Toward an Agent-Based Conception of Modeling in the Science Classroom." *Science & Education* 26: 49–63. <https://doi.org/10.1007/s11191-017-9884-4>.

Graves, Michael. 1977. "The Necessity for Drawing: Tangible Speculation." *Architectural Design* 47, no. 6: 384–394.

Guillou, Erwan, Daniel Meneveau, Eric Maisel, and Kadi Bouatouch. 2000. "Using Vanishing Points for Camera Calibration and Coarse 3D Reconstruction from a Single Image." *The Visual Computer* 16: 396–410. <https://doi.org/10.1007/pl00013394>.

Hoiem, Derek, Alexei Efros, and Martial Hebert. 2005. "Automatic Photo Pop-Up." In *Proceedings, ACM SIG-GRAPH 2005*, edited by M. Gross, 577–584. New York: Association for Computing Machinery. <https://doi.org/10.1145/1186822.1073232>.

Hoiem, Derek, Andrew N. Stein, Alexei A. Efros, and Martial Hebert. 2007. "Recovering Occlusion Boundaries from a Single Image." In *Proceedings, 2007 IEEE 11th International Conference on Computer Vision*, 1-8. IEEE. <https://doi.org/10.1109/iccv.2007.4408985>.

Horry, Youichi, Ken-Ichi Anjyo, and Kiyoshi Arai. 1997. "Tour into the Picture: Using a Spidery Mesh Interface to Make Animation from a Single Image." In *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques*, edited by G. Owen, T. Whitted, and B. Mones-Hattal, 225–232. New York: ACM Press. <https://doi.org/10.1145/258734.258854>.

Janisio-Pawłowska, Dorota. 2021. "Analysis of the Possibilities of Using HBIM Technology in the Protection of Cultural Heritage, Based on a Review of the Latest Research Carried out in Poland." *ISPRS International Journal of Geo-Information* 10, no. 10: 633. <https://doi.org/10.3390/ijgi10100633>.

Jouan, Pierre, and Pierre Hallot. 2020. "Digital Twin: Research Framework to Support Preventive Conservation Policies." *ISPRS International Journal of Geo-Information* 9, no. 4: 228. <https://doi.org/10.3390/ijgi9040228>.

Kane, Thomas J. 1977. "The Use of Reverse Perspective in the Deduction of Plans and Elevations from Photographs." *Bulletin of the Association for Preservation Technology* 9, no. 3: 30–38.

Klodt, Maria, and Andrea Vedaldi. 2018. "Supervising the New with the Old: Learning SFM from SFM." In *Proceedings of the European Conference on Computer Vision*, edited by V. Ferrari, M. Hebert, C. Sminchisescu, and Y. Weiss, 698–713. Cham: Springer. [https://doi.org/10.1007/978-3-030-01249-6\\_43](https://doi.org/10.1007/978-3-030-01249-6_43).

Leung, Anthony K. H., Stephen Davies, and Steve H. Ching. 2018. "When New Technology Joins Old Documents and East Meets West: Virtually Reconstructing the Fisher Island Pagoda Lighthouse (China)." *Virtual Archaeology Review* 9, no. 18: 12-27. <https://doi.org/10.4995/var.2018.7982>.

Liebowitz, David, Antonio Criminisi, and Andrew Zisserman. 1999. "Creating Architectural Models from Images." *Computer Graphics Forum* 18, no. 3: 39–50. <https://doi.org/10.1111/1467-8659.00326>.

Lopez-Antequera, Manuel, Roger Mari, Pau Gargallo, Yubin Kuang, Javier Gonzalez- Jimenez, and Gloria Haro. 2019. "Deep Single Image Camera Calibration with Radial Distortion." In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 11817–11825. IEEE. <https://doi.org/10.1109/cvpr.2019.01209>.

- Mertan, Alican, Damien Jade Duff, and Gozde Unal. 2022. "Single Image Depth Estimation: An Overview." *Digital Signal Processing* 123: 103441. <https://doi.org/10.1016/j.dsp.2022.103441>.
- Nishida, Gen, Adrien Bousseau, and Daniel G. Aliaga. 2018. "Procedural Modeling of a Building from a Single Image." *Computer Graphics Forum* 37, no. 2: 415–429. <https://doi.org/10.1111/cgf.13372>.
- Parodi, Pietro, and Giulia Piccioli. 1996. "3D Shape Reconstruction by Using Vanishing Points." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 18, no. 2: 211–217. <https://doi.org/10.1109/34.481545>.
- Pérez-Gómez, Alberto. 2005. "Questions of Representation: The Poetic Origin of Architecture." *arq: Architectural Research Quarterly* 9, no. 3-4: 217–225. <https://doi.org/10.1017/S1359135505000278>.
- Piotrowski, Andrzej. 2015. "Heresy, Hybrid Buildings, and a Geography of Architectural Traditions." *Traditional Dwellings and Settlements Review* 27, no. 1: 7–19. <https://www.jstor.org/stable/24719997>.
- Schonberger, Johannes L., Filip Radenovic, Ondrej Chum, and Jan-Michael Frahm. 2015. "From Single Image Query to Detailed 3D Reconstruction." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 5126–5134. Los Alamitos (USA): IEEE. <https://doi.org/10.1109/cvpr.2015.7299148>.
- Son, Hyojoo, Frédéric Bosché, and Changwan Kim. 2015. "As-Built Data Acquisition and its Use in Production Monitoring and Automated Layout of Civil Infrastructure: A Survey." *Advanced Engineering Informatics* 29, no. 2: 172–183. <https://doi.org/10.1016/j.aei.2015.01.009>.
- Sturm, Peter, and Steve Maybank. 1999. "A Method for Interactive 3D Reconstruction of Piecewise Planar Objects from Single Images." In *Proceedings, The 10th British Machine Vision Conference*, edited by T. Pridmore and D. Elliman, 265–274. Nottingham (UK): The British Machine Vision Association. <https://doi.org/10.5244/c.13.27>.
- Sullivan, Elaine A., and Lisa M. Snyder. 2017. "Digital Karnak: An Experiment in Publication and Peer Review of Interactive, Three-Dimensional Content." *Journal of the Society of Architectural Historians* 76, no. 4: 464–482. <https://doi.org/10.1525/jsah.2017.76.4.464>.
- Van den Heuvel, Frank A. 1998. "3D Reconstruction from a Single Image Using Geometric Constraints." *ISPRS Journal of Photogrammetry and Remote Sensing* 53, no. 6: 354–368. [https://doi.org/10.1016/s0924-2716\(98\)00019-7](https://doi.org/10.1016/s0924-2716(98)00019-7).
- Wong, Johnny Kwok Wai, Janet Ge, and Sean Xiangjian He. 2018. "Digitisation in Facilities Management: A Literature Review and Future Research Directions." *Automation in Construction* 92: 312–326. <https://doi.org/10.1016/j.autcon.2018.04.006>.