Innovative techniques for the survey of objects no longer accessible and not measurable

Maurizio Perticarini

Abstract

The study focuses on how, through AI and new survey techniques based on radiance fields such as Gaussian splatting, it is possible to recreate lost objects and spaces even without a suitable dataset for their reconstruction. The case study concerns the Huge Wineglass statue by architect Toyo Ito in Pescara, which was removed and replaced with the current The Big Piano fountain due to permanent damage from thermal shock. There are some low-resolution photographs and videos of the Huge Wineglass, which until recently were not sufficient for a virtual reconstruction of the monument. By using artificial intelligence for upscaling and enhancing the quality of the videos, it was possible to create and improve a dataset useful for processing a survey based on new technologies available such as Gaussian splatting and Nerf. Through the online platform Luma AI, the reconstruction of the object and its insertion into the virtual twin of Piazza della Rinascita, the main square of the city of Pescara, in the position where it was originally located, was possible. Studies like this demonstrate how the rapid evolution of image-based acquisition tools, thanks to artificial intelligence, are capable, despite the absence of precise empirical data, of allowing the enjoyment of spaces and architectures no longer present or inaccessible for conducting a traditional survey.

Keywords
artificial intelligence, gaussian splatting, nerf, dataset improvement, upscaling video

Toyo Ito, *Huge Wineglass* in Pescara. Photo by Mauro Vitale.

doi.org/10.3280/oa-1180-c565

Introduction

Technological advancement in the field of artificial intelligence (AI) has opened new prospects in data treatment and analysis, enabling the overcoming of seemingly insurmountable limitations. In this context, the reconstruction of low-resolution images or videos and 3D surveying based on radiance fields, which is gaining more ground in recent years, are useful tools for processing spaces, monuments, or architectures that have disappeared and where it is no longer possible to perform a measurable survey - whether an integrated survey or based on image-based techniques - starting from precise and high-resolution data. This possibility represents a fascinating challenge and is made possible by the integration of innovative methodologies based on AI and advanced image-based survey techniques like Gaussian Splatting. The recovery of information from low-resolution images or videos could be a step forward for various disciplines, from archaeology to forensic analysis, but could also be useful for reconstructing settings or architectures that do not need to be parameterized and measured in detail but can be employed as assets in immersive virtual reality projects or applied games aimed at the dissemination and enhancement of cultural heritage in general. The research focuses on a pioneering approach that starts by increasing the resolution of video frames through algorithms based on Convolutional Neural Networks (CNN) or Generative Adversarial Networks (GAN) - in this case, an online platform called Vmake AI developed by Pixocial in Singapore was used. The post-produced video was then used as a dataset for generating a 3D model through the Gaussian splatting method [Basso 2019] [Karras et al. 2018].

The Huge Wine Glass, a sculpture created by architect Toyo Ito for the city of Pescara in Abruzzo (Italy) and inaugurated on December 14, 2008, constitutes the case study of this research. It was placed in an off-center point of Piazza della Rinascita, the main square of the city; weighing 24 tons and made of acrylic resin, it depicted a large semi-transparent wine glass (fig. 1). In mid-February 2009, there was a sudden failure that led to the securing of the statue with a metal containment structure. The sculpture remained in that state for a long time before being removed because it was considered irrecoverable. Today, in the same spot of the square, a large luminous fountain called The Big Piano has been created (fig. 2).

Toyo Ito's work is still felt as a profound absence by the city of Pescara, a symbol that could never fully consolidate but remained indelible in memory. The project thus aims to revive its memory and does so with the almost dreamlike possibilities that artificial intelligence offers.

Fig. 1. *Huge Wineglass* after the implosion and the current The Big Piano fountain.

Fig. 2. On the left the *Huge Wineglass* by Toyo Ito after the breakage, on the right the new Big Piano fountain. Elaboration by the author.

Methodology for increasing video resolution

It is important to understand that upscaling algorithms operate on the pixels of existing images and do not add real information to the original resolution. Instead, they attempt to generate estimated details based on patterns learned from a vast dataset during the training process. The algorithms involved are:
Convolutional Neural Networks (CNN), which are a type of deep neural network commonly used in the field of computer vision. These networks are particularly effective at analysing images due to their ability to learn hierarchical features from two-dimensional data. CNNs divide the image into small overlapping fragments called "filters" or "kernels" and apply convolution operations to detect salient features such as edges, textures, and patterns.
Generative Adversarial Networks (GAN), which are a type of neural architecture consisting of two neural networks, the generator and the discriminator, that compete with each other during the training process. The generator tries to create data (e.g., high-resolution images) that are difficult to distinguish from real images, while the discriminator attempts to distinguish between the generated and real images. This competition leads the generator to constantly improve the quality of the generated images.
Super-resolution through deep learning aims to generate high-resolution images from low-resolution images. Super-resolution algorithms use deep neural networks to learn the complex relationships between the pixels of the low-resolution image and their corresponding high-resolution versions. This allows them to infer missing details and enhance the quality of the image.
Upscaling techniques, which use algorithms to increase the resolution of an image or video. These algorithms may be based on simple interpolations, such as linear or bilinear interpolation, which fill in missing details between existing pixels. However, AI-based techniques tend to be more sophisticated, as they consider the spatial and semantic relationships between pixels to generate more realistic details.
In summary, algorithms that enhance the resolution of photographs and videos through artificial intelligence combine advanced deep learning approaches, convolutional neural networks, and GANs to generate high-resolution images that approximate missing details in low-resolution versions.
The video used for conversion dates back to January 27, 2009, and was probably recorded with a smartphone of that era. The resolution is 480 x 360 pixels, and the movement of the camera, being devoid of optical stabilization, is very shaky. The only strength of the footage is that it was taken by circling around the object, framing it from all sides. This type of shooting, although executed for other purposes, was very useful for the second phase of the workflow described here, namely the 3D rendering of the statue. After some video cropping operations to use only the parts that concerned the object to be detected, it was uploaded to the online platform called Vmake AI developed by Pixocial in Singapore: this platform allows - free of charge for small-sized videos - to increase and improve the video resolution

up to 4K format. In the described case, the resolution was increased to HD format, thus reaching 1920 x 1440 pixels. As seen from figure 3, in addition to having increased the size of the format, the video has had a considerable improvement in resolution, thanks to the denoiser that eliminated the background noise caused by the camera lens and the shaking of the footage. The only negative factor is the slight decrease in the level of detail, this because denoisers work by reducing the intensity variations between nearby pixels, which are often interpreted as noise. This can lead to the loss of fine details, making the image less sharp and giving it a "smeared" appearance. However, this was not significant for the purpose of the subsequent 3D rendering, as the surface of the statue was smooth and with almost no roughness.

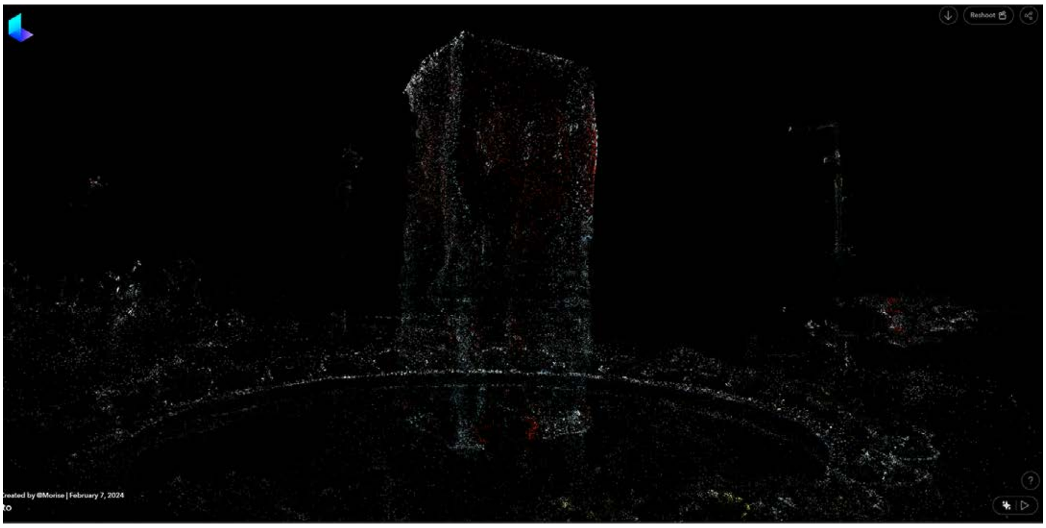

Fig. 3. Sparse point cloud derived from the Gaussian splatting process through the Luma AI platform. Elaboration by the author.

Methodology for 3D reconstruction with Gaussian splatting

In recent years, the landscape of applied artificial intelligence has undergone profound transformations, significantly influencing the methodological approach in various sectors, especially in the context of generative representation. Algorithms and complex systems have played a decisive role in improving three-dimensional visual representations and in defining more accurate Digital Twins. These advances have greatly reduced errors in the reproduction of three-dimensional spaces, introducing technological innovations such as the Neural Radiance Field (NeRF) and complex machine learning architectures like Generative Adversarial Networks (GAN), which have played a crucial role in the evolution of spatial representation [Mildenhall et al. 2020] [Balloni et al. 2023] [Croce et al. 2024] [Gao et al. 2022] [Remondino et al. 2023]. The NeRF approach, originally proposed by Soft3D, has been refined over time with increasingly advanced deep learning techniques, embracing volumetric ray-marching and exploiting a continuous differential density field to describe geometry. However, while NeRF introduced positional sampling and encoding to improve quality, the use of the Multi-Layer Perceptron inevitably impacted computational speed negatively. The success of NeRF catalyzed the development of new approaches, such as Nvidia's Instant NeRF or Plenoxels, which aimed to improve rendering quality and speed, often introducing regularization strategies. Despite the significant progress, the challenge related to training and rendering times remained pressing. Currently, the system offering the best visual results is known as Gaussian Splatting, an evolved technique that, without directly utilizing neural networks, leverages photogrammetric algorithms and Structure from Motion (SfM) for three-dimensional rendering [Chen et al. 2024] [Kerbl et al. 2023]. This method presents an innovative approach using three-dimensional Gaussian functions to realistically render scenes, avoiding the resource

expenditure typical of neural network-based models. These processes, although somewhat embryonic from the standpoint of measurement reliability (very established in traditional processes like laser scanning and photogrammetry), prove valid from the perspective of rendering and photorealistic output. Compared to photogrammetry, it is possible to survey reflective and transparent objects, as these methods consist of complex mathematical functions that consider both the chromatic aspect and the density of the detected objects. The type of rendering is no longer based on ray tracing and the processing of mesh surfaces to which a texture is applied; thus, what emerges from these types of processes resembles more a fog of pixels (in the case of NeRF) or, precisely, a set of Gaussians (ellipsoids) that together compose the detected object. In the case of NeRF, the neural network is trained to understand how light interacts with the volume of the scene, including scattering and absorption effects, which can simulate reflections and refractions; in the case of Gaussian Splatting, shaders are used to simulate reflection and refraction by calculating the paths of light through the Gaussian distributions and adjusting the color and intensity of the light based on the optical properties of the objects.
The platform used for this purpose, which offers the possibility to carry out surveys based on both NeRF and the new Gaussian process for free, is Luma AI. It allows users to upload previously made videos or videos made within the App, and through the cloud, it performs preliminary training operations. Finally, it creates the 3D of the depicted object. As seen in figure 4, the algorithm creates a sparse cloud using the same process that occurs in photogrammetry, needing, however, many fewer frames for point alignment. To each point of the sparse cloud, a Gaussian (ellipsoid with RGB chromatic properties and density properties) is matched. It can be noted how the graphic rendering, despite the significant limitations that the survey has from the perspective of the informational database, is quite good: the reflections and refractions of the statue's material recreate the silhouette of the red wine glass, and the latter changes its shape depending on the angle of view (fig. 5). Luma AI allows the export in standard output formats such as .ply or .obj, as it is capable of converting the product of the rendering based on NeRF and Gaussian splatting both in the form of a point cloud and as a mesh. However, exporting in standard formats does not allow viewing the true result of this type of survey in other software. Some software like Blender 3D, Unity, or Unreal Engine have developed add-ons capable of reading these particular formats, which manage to maintain the original visual aspect. Therefore, by exporting in the formats .luma and .ply (Gaussian), it is possible to visualize and especially modify the models within specific software for editing, rendering, or the development of applications.

Fig. 4. Result of the Gaussian splatting relief with the Luma AI platform. Elaboration by the author.

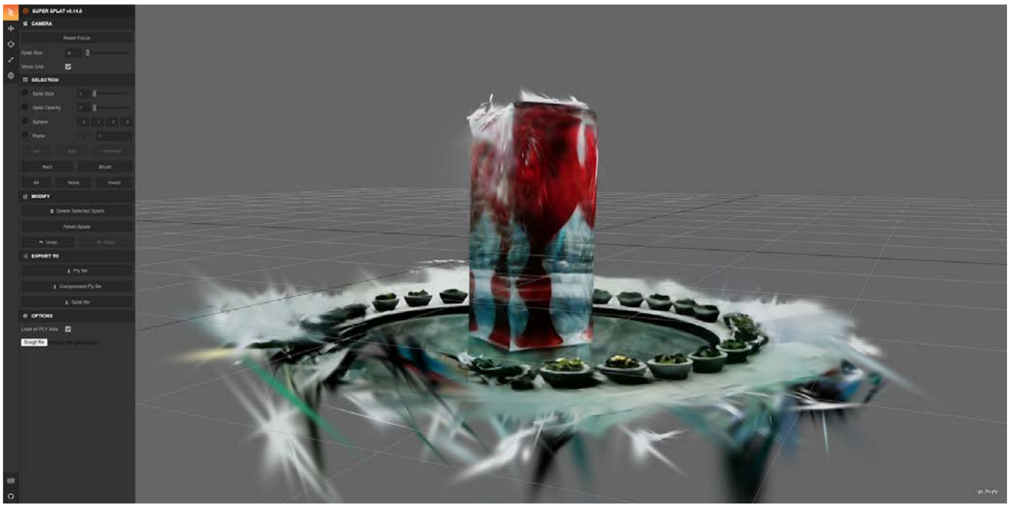

Fig. 5. Gaussian splatting survey imported into the Playcanvas platform in its super splat manager, isolated from context through editing commands. Elaboration by the author.

Results

For the management of the surveys, Playcanvas was used, an online platform founded by Will Eastcott and Dave Evans in 2011, created as a web-based game development and 3D visualization platform, leveraging the power of HTML5 and WebGL to provide an accessible and high-performance development environment for creating games, interactive experiences, and 3D applications directly in the browser. This platform integrated "Super Splat," a space for managing and editing Gaussians that is easy and intuitive to use, but also very accurate in terms of rendering and visualization. Through the platform's commands, it is possible to perform cropping operations and clean up excess points and isolate the survey object from its context; it is also possible to modify its origin point and perform translation, rotation, and scaling operations (fig. 6). Within the platform, there is a space where Gaussian objects can be placed, the points of the generative point cloud can be viewed, and, through sliders, points that are more or less distant from each other can be selected to allow greater control in the editing phases. It is possible to add other Gaussian objects within the scene to compose complex scenarios made up of multiple surveys. Using this feature, it was possible to insert the statue within a specifically made survey of the current state of Piazza della Rinascita. As evident from figure 7, the survey of the context (carried out through a video in the form of a circular panorama, whose focal point is the new Big Piano fountain) was then modified by removing the part pertaining to the new fountain and was used as the context for inserting Toyo Ito's work. From figure 7, it can be noted that in addition to blending in with the context, the statue reflects light respecting the angle of the viewpoint. Indeed, these rendering systems show their strengths in similar situations, where reflections and refractions of light are crucial to understanding the shape of the detected object. Regardless of the quality of the 3D, heavily conditioned by the quality of the informational dataset, this type of survey would not have been returned by traditional range and image-based techniques, precisely because of the sculpture's conformation and its reflective and refractive material to light. The silhouette of the red wine glass, in the best-case scenario, would have been converted into a static texture and baked onto the mesh. In this case, however, there is a different perception of the glass from every point of view.

Conclusions

In conclusion, the use of methodologies based on Neural Radiance Fields (NeRF) and Gaussian Splatting underscores a significant evolution in the digital reconstruction of objects, architectures, and historical sites that no longer exist. These technological advancements offer

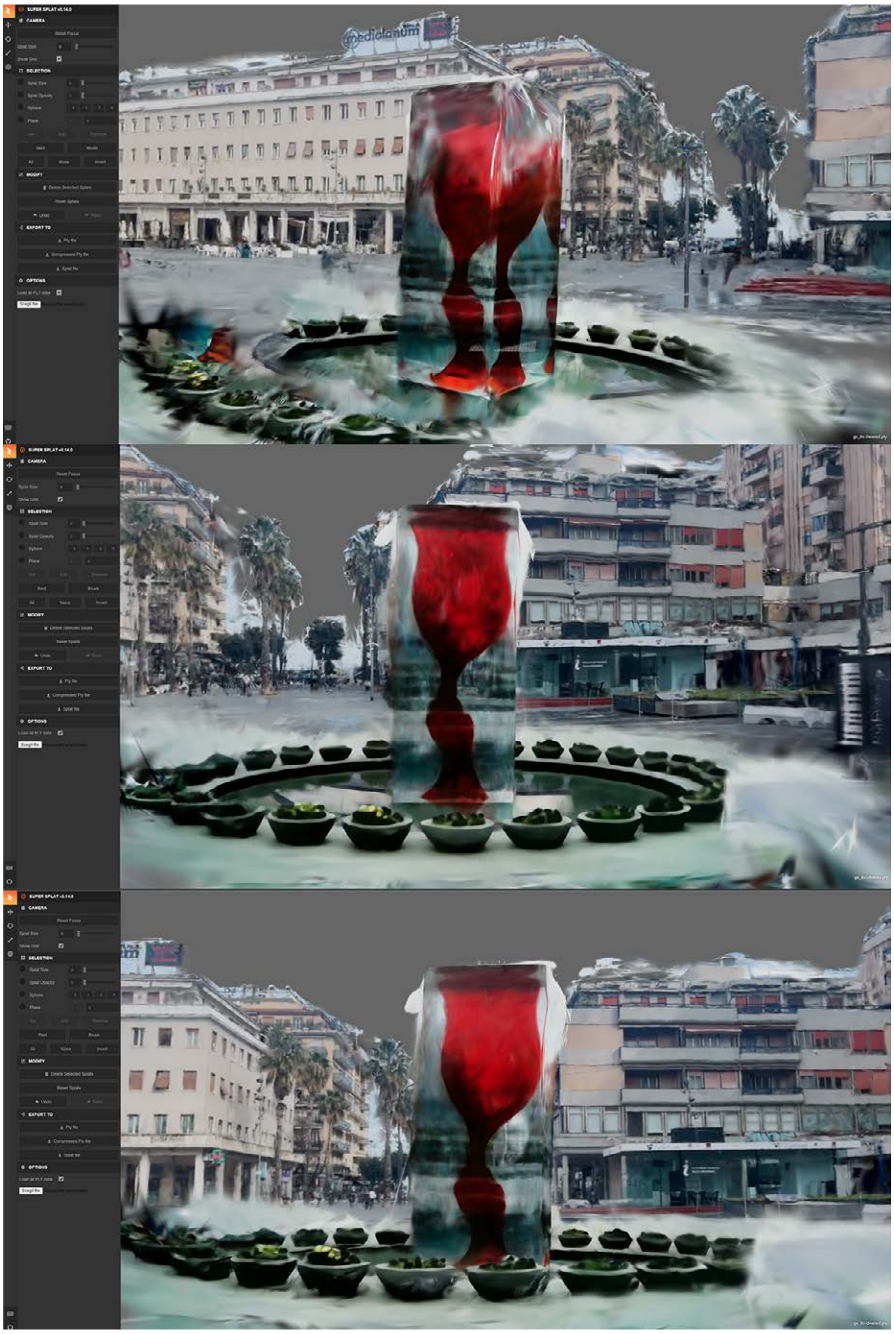

Fig. 6. Result of the union of the two surveys carried out with the same technique through the Playcanvas platform. Elaboration by the author.

an invaluable contribution to the field of digital preservation, allowing scholars and the public to explore and understand lost cultural heritage. Gaussian Splatting proves essential for the effective management of sparse data, enabling the fusion of fragmented information into a single cohesive representation. This technique facilitates the creation of continuous surfaces from point data, significantly improving the visual quality of 3D reconstructions. The importance of these methodologies lies not only in their ability to bring the past to light but also in their potential to democratize access to history. Digital reconstruction based on NeRF and Gaussian Splatting allows overcoming physical and temporal barriers, offering a new dimension for education and research. Through interactive simulations, it is possible to dynamically explore historical and cultural contexts that would otherwise remain inaccessible, broadening our understanding of places.

References

Balloni E., et al. (2023). Few Shot Photogrammetry: A Comparison Between Nerf and Mvs-Sfm for the Documentation of Cultural Heritage. In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 48M2, pp. 155–162.

Basso A. (2019). New interpretation tools and metamorphosis of the image. In E. Cicalò (Ed.),,*Advances in Intelligent Systems and Computing 1140*. Proceedings of the 2nd International and Interdisciplinary Conference on Image and Imagination, IMG 2019, pp. 923-935. Basel: Springer Nature Switzerland.

Chen G., Wang W. (2024). A Survey on 3D Gaussian Splatting. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1–15. <https://doi.org/10.48550/arXiv.2401.03890>

Croce V., et al. (2024). Comparative Assessment of Neural Radiance Fields and Photogrammetry in Digital Heritage: Impact of Varying Image Conditions on 3D Reconstruction. In *Remote Sensing*, vol. 16, n. 2. <https://doi.org/10.3390/rs16020301>

Gao K., et al. (2022). NeRF: Neural Radiance Field in 3D Vision, a Comprehensive Review. In *ArXiv*, abs/2210.00379 <https://doi.org/10.48550/arXiv.2210.00379>

Karras T., (2018). Progressive Growing of GANs for Improved Quality, Stability, and Variation. In *Proceedings of International Conference on Learning Representations (ICLR)*. <https://openreview.net/forum?id=Hk99zCeAb>

Kerbl B., et al. (2023). 3D Gaussian Splatting for Real-Time Radiance Field Rendering. In *ACM Transactions on Graphics*, vol. 42, n. 4, pp. 1–14. <https://doi.org/10.1145/3592433>

Mazzacca G., et al. (2023). NeRF for Heritage 3D Reconstruction. In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLVIII-M-2-2023, pp. 1051–1058. <https://doi.org/10.5194/isprs-archives-XLVIII-M-2-2023-1051-2023>

Mildenhall B., et al. (2020). NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis. In *Computer Vision – ECCV 2020: 16th European Conference*. Glasgow, UK, August 23–28, 2020, Proceedings, Part I, pp. 405–421. Cham: Springer Verlag. <https://doi.org/10.1007/978-3-030-58452-8_24>

Murtiyoso A., Grussenmeyer P. (2023). Initial Assessment on the Use of State-Of-The-Art NeRF Neural Network 3D Reconstruction for Heritage Documentation. In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLVIII-M-2-2023, pp. 1113–1118. <https://doi.org/10.5194/isprs-archives-XLVIII-M-2-2023-1113-2023>

Parrinello S., Sanseverino A., Fu H. (2023). HBIM Modelling for the Architectural Valorisation Via a Maintenance Digital Eco-System. In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLVIII-M-2-2023, pp. 1157–1164. <https://doi.org/10.5194/isprs-archives-XLVIII-M-2-2023-1157-2023>

Palestini C., Basso A., Perticarini M. (2022). Machine Learning as an Alternative to 3D Photomodeling Employed in Architectural Survey and Automatic Design Modelling. In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLVIII, pp. 191–197. <https://10.5194/isprs-archives-XLVIII-2-W1-2022-191-2022>

Remondino F., et al. (2023). A Critical Analysis of NeRF-Based 3D Reconstruction. In Remote Sensing, vol.15, n. 14 <https://doi.org/10.3390/rs15143585>

Tancik M., et al. (2023). Nerfstudio: A Modular Framework for Neural Radiance Field Development. ACM SIGGRAPH 2023 Conference Proceedings, pp. 1–12. <https://doi.org/10.1145/3588432.3591516>

Xu Q., et al. (2022). Point-Nerf: Point-Based Neural Radiance Fields. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 5438–5448.

Author
Maurizio Perticarini, University of Padua, maurizio.perticarini@unipd.it

To cite this chapter: Maurizio Perticarini (2024). Innovative techniques for the survey of objects no longer accessible and not measurable. In Bergamo F., Calandriello A., Ciammaichella M., Friso I., Gay F., Liva G., Monteleone C. (Eds.). *Misura / Dismisura. Atti del 45° Convegno Internazionale dei Docenti delle Discipline della Rappresentazione/Measure / Out of Measure. Transitions. Proceedings of the 45th International Conference of Representation Disciplines Teachers.* Milano: FrancoAngeli, pp. 1895-1902.

Isbn 9788835166948