Photogrammetric Survey for a Fast Construction of Synthetic Dataset

Original
Photogrammetric Survey for a Fast Construction of Synthetic Dataset / Tomalini, Andrea; Pristeri, Edoardo; Bergamasco, Letizia. - ELETTRONICO. - REPRESENTATION CHALLENGES - Augmented Reality and Artificial Intelligence in Cultural Heritage and Innovative Design Domain:(2021), pp. 215-219. [10.3280/oa-686.34]

Publisher:
FrancoAngeli

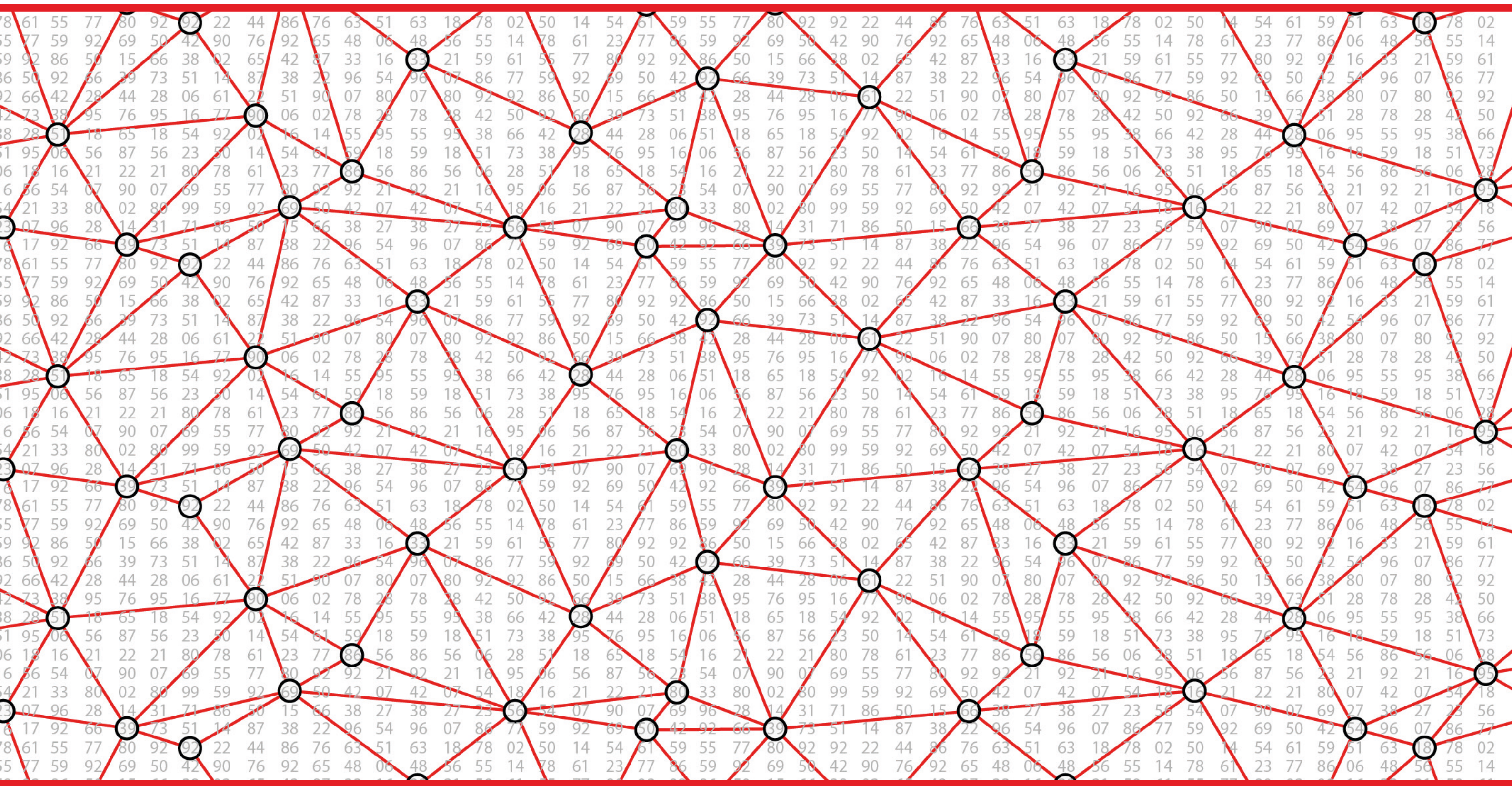
Published
DOI:10.3280/oa-686.34

REPRESENTATION CHALLENGES

Augmented Reality and Artificial Intelligence in Cultural Heritage and Innovative Design Domain

edited by

Andrea Giordano
Michele Russo
Roberta Spallone



FrancoAngeli OPEN ACCESS

diségno

diségno

the Series of the UID – Unione Italiana per il Disegno

director Francesca Fatta

The Series contains volumes of the proceedings of the annual conferences of the Scientific Society UID – Unione Italiana per il Disegno and the results of international meetings, research and symposia organised as part of the activities promoted or patronised by UID. The topics concern the Scientific Disciplinary Sector ICAR/17 Drawing with interdisciplinary research areas. The texts are in Italian or in the author's mother tongue (French, English, Portuguese, Spanish, German) and/or in English. The international Scientific Committee includes members of the UID Scientific Technical Committee and numerous other foreign scholars who are experts in the field of Representation.

The volumes of the series can be published either in print or in open access and all the authors' contributions are subject to double blind peer review according to the currently standard scientific evaluation criteria.

Scientific Committee
Giuseppe Amoruso *Politecnico di Milano*
Paolo Belardi *Università degli Studi di Perugia*
Stefano Bertocci *Università degli Studi di Firenze*
Mario Centofanti *Università degli Studi dell'Aquila*
Enrico Cicalò *Università degli Studi di Sassari*
Antonio Conte *Università degli Studi della Basilicata*
Mario Docci *Sapienza Università di Roma*
Edoardo Dotto *Università degli Studi di Catania*
Maria Linda Falcidieno *Università degli Studi di Genova*
Francesca Fatta *Università degli Studi Mediterranea di Reggio Calabria*
Fabrizio Gay *Università IUAV di Venezia*
Andrea Giordano *Università degli Studi di Padova*
Elena Ippoliti *Sapienza Università di Roma*
Francesco Maggio *Università degli Studi di Palermo*
Anna Osello *Politecnico di Torino*
Caterina Palestini *Università degli Studi "G. d'Annunzio" di Chieti-Pescara*
Lia Maria Papa *Università degli Studi di Napoli "Federico II"*
Rossella Salerno *Politecnico di Milano*
Alberto Sdegno *Università degli Studi di Udine*
Chiara Vernizzi *Università degli Studi di Parma*
Ornella Zerlenga *Università degli Studi della Campania "Luigi Vanvitelli"*

Members of foreign structures
Caroline Astrid Bruzelius *Duke University - USA*
Pilar Chías *Universidad de Alcalá - Spagna*
Frank Ching *University of Washington - USA*
Livio De Luca *UMR CNRS/MCC MAP Marseille - Francia*
Roberto Ferraris *Universidad Nacional de Córdoba - Argentina*
Glaucia Augusto Fonseca *Universidade Federal do Rio de Janeiro - Brasile*
Pedro Antonio Janeiro *Universidade de Lisboa - Portogallo*
Jacques Laubscher *Tshwane University of Technology - Sudafrica*
Cornelie Leopold *Technische Universität Kaiserslautern - Germania*
Juan José Fernández Martín *Universidad de Valladolid - Spagna*
Carlos Montes Serrano *Universidad de Valladolid - Spagna*
César Otero *Universidad de Cantabria - Spagna*
Guillermo Peris Fajarnes *Universitat Politècnica de València - Spagna*
José Antonio Franco Taboada *Universidade da Coruña - Spagna*
Michael John Kirk Walsh *Nanyang Technological University - Singapore*

Graphic design and cover by Enrico Cicalò and Paola Raffa

REPRESENTATION CHALLENGES
Augmented Reality and Artificial Intelligence in Cultural Heritage and Innovative Design Domain

edited by

Andrea Giordano
Michele Russo
Roberta Spallone

FrancoAngeli OPEN ACCESS

REPRESENTATION CHALLENGES
Augmented Reality and Artificial Intelligence in
Cultural Heritage and Innovative Design Domain

Scientific Committee

Salvatore Barba
Università di Salerno

Marco Giorgio Bevilacqua
Università di Pisa

Stefano Brusaporci
Università dell'Aquila

Francesca Fatta
Università Mediterranea di Reggio Calabria

Andrea Giordano
Università di Padova

Alessandro Luigini
Libera Università di Bolzano

Michele Russo
Sapienza Università di Roma

Cettina Santagati
Università di Catania

Alberto Sdegno
Università di Udine

Roberta Spallone
Politecnico di Torino

Scientific Coordination

Andrea Giordano
Università di Padova

Michele Russo
Sapienza Università di Roma

Roberta Spallone
Politecnico di Torino

Editorial Committee

Isabella Friso
Università IUAV di Venezia

Fabrizio Natta
Politecnico di Torino

Michele Russo
Sapienza Università di Roma

The texts as well as all published images have been provided by the authors for publication with copyright and scientific responsibility towards third parties. The revision and editing is by the editors of the book.

ISBN printed edition: 9788835116875
ISBN digital edition: 9788835125280

Peer Reviewers

Marinella Arena
Università Mediterranea di Reggio Calabria

Salvatore Barba
Università di Salerno

Marco Giorgio Bevilacqua
Università di Pisa

Cecilia Bolognesi
Politecnico di Milano

Stefano Brusaporci
Università dell'Aquila

Francesca Fatta
Università Mediterranea di Reggio Calabria

Andrea Giordano
Università di Padova

Massimo Leserri
Università di Napoli "Federico II"

Stefania Landi
Università di Pisa

Massimiliano Lo Turco
Politecnico di Torino

Alessandro Luigini
Libera Università di Bolzano

Pamela Maiezza
Università dell'Aquila

Domenico Mediati
Università Mediterranea di Reggio Calabria

Cosimo Monteleone
Università di Padova

Michele Russo
Sapienza Università di Roma

Cettina Santagati
Università di Catania

Alberto Sdegno
Università di Udine

Roberta Spallone
Politecnico di Torino

Marco Vitali
Politecnico di Torino

Patronage

Cover image: Michele Russo

Copyright © 2021 by FrancoAngeli s.r.l., Milano, Italy.

This work, and each part thereof, is protected by copyright law and is published in this digital version under the license *Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International* (CC BY-NC-ND 4.0)

By downloading this work, the User accepts all the conditions of the license agreement for the work as stated and set out on the website
https://creativecommons.org/licenses/by-nc-nd/4.0

Index

AR&AI theoretical concepts

AR&AI virtual reconstruction

AR&AI heritage routes

AR&AI classification and 3D analysis

AR&AI urban enhancement

AR&AI museum heritage

AR&AI building information modeling and monitoring

AR&AI education and shape representation

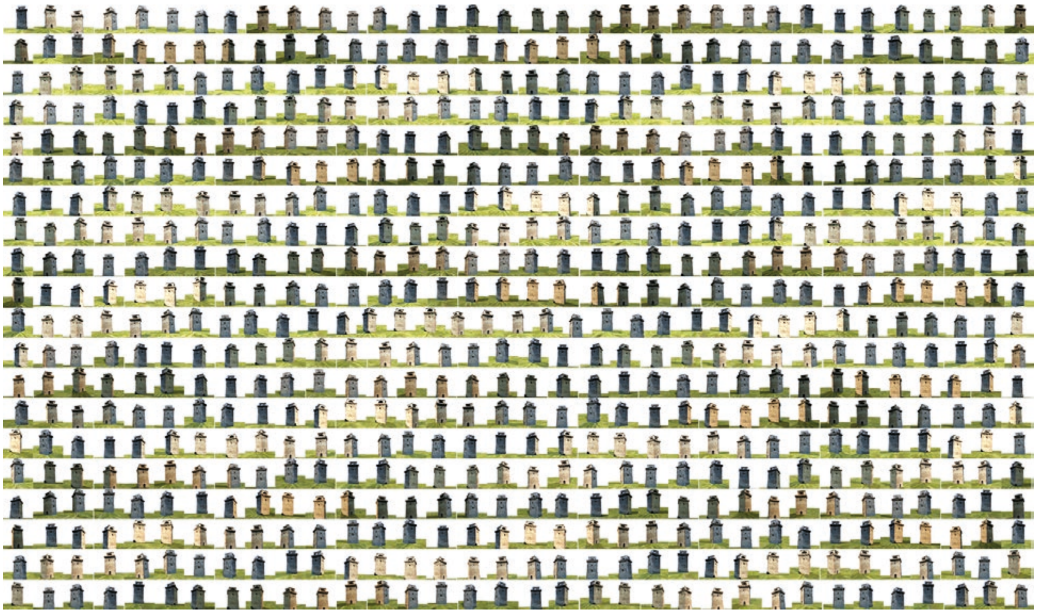
Photogrammetric Survey for a Fast Construction of Synthetic Dataset

Andrea Tomalini
Edoardo Pristeri
Letizia Bergamasco

Abstract

In this work we show how Physically Based Rendering (PBR) tools can be used to extend the training image datasets of Machine Learning (ML) algorithms for the recognition of built heritage. In the field of heritage valorization, the combination of Artificial Intelligence (AI) and Augmented Reality (AR) has allowed to recognize built heritage elements with mobile devices, anchoring digital products to the physical environment in real time, thus making the access to information related to real space more intuitive and effective. However, the availability of training data required for these systems is extremely limited and a large–scale image dataset is required to achieve accurate results in image recognition. Manually collecting and annotating images can be very resource and time–consuming. In this contribution we explore the use of PBR tools as a viable alternative to supplement an otherwise inadequate dataset.

Keywords
synthetic dataset, image recognition, visual programming language, physically based render.



doi.org/10.3280/oa-686.34

Introduction

In the context of built heritage enhancement, the use of mobile computing technologies for its fruition can revolutionize the user experience [Barsanti et al. 2018; Lo Turco et al. 2019]. AR systems, if properly combined with ML algorithms, can expand the level of knowledge that can be accessed while observing the asset [Spallone et al. 2020]. While it is very easy to imagine the database containing the information associated with an architectural asset, it is less immediate to imagine the query needed to access and explore it. Considering that within the same database very different information is stored: referring to the history, the construction technique, the history of the architect, etc.; one can understand how solutions such as audio tours, information panels, or QR codes are not suitable to answer the subjective curiosity of the user [Andrianaivo et al. 2019]. With the help of a mobile device, starting from the recognition of the object itself, one could connect and reorganize all this information according to the user's preferences [Vayanou et al. 2014].

To enable this kind of navigation, one of the first steps is to ensure that the mobile device can recognize the object in the frame. However, while some disciplines already apply Deep Learning for image recognition [Norouzzadeh et al. 2018; Liu et al. 2020], research is not as flourishing for architectural feature recognition.

This research work proposes a methodology to enrich the training dataset needed to build a software capable of recognizing the built heritage from pictures coming from a mobile device with the help of PBRT tools. Once the architectural artifact has been surveyed, its digital twin can be inserted into a modeling environment and used for the creation of possible views, even improbable ones, expeditiously, taking into account different lighting and meteorological settings which could affect the picture taken from the end user.

Case Study Definition and Data Acquisition

Given the still preliminary stage of the work, the Saracen Tower of Spotorno was chosen as a case study by the research team because of its small size and its position visible from different points of the town. For this specific use case, a three–dimensional model has been useful for the creation of photorealistic rendering used to train a ML algorithm. For this reason, and to optimize working time, it was decided to generate a photogrammetric model by carrying out a free–net adjustment with a subsequent assignment of the model scale by applying the method of least squares over 3 known distances, measured using a metric token (fig. 1).

The approach of using elements of known length is a cheap, expeditious and well–established procedure, both in the orientation of the photogrammetric block in industrial applications [Luhmann et al. 2010], and in the survey of archaeological heritage [Nocerino et al. 2013]. The important thing is to correctly size the supports taken as reference – they must be proportionate to the object to be surveyed –, the distance of the images, and the degree of precision and accuracy required by the final model.

The aerial photogrammetric shots were taken using a DJI Mavic Mini drone, equipped with a 1/2.3" CMOS sensor. The dataset was integrated with some images taken from the ground with a Sony Alpha 6000 camera equipped with a 23.5x15.6 mm sensor.

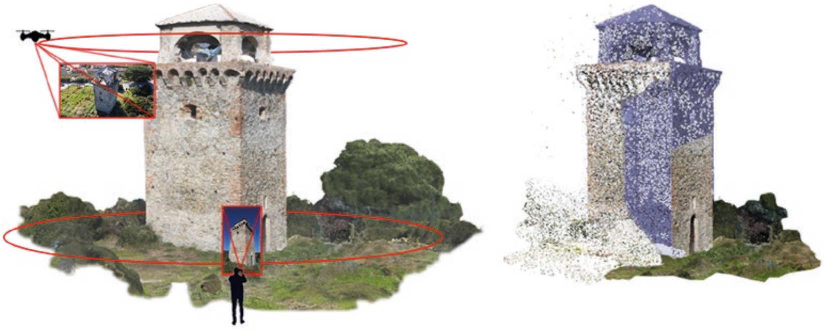


Fig. 1. Conceptual scheme of the survey and construction phases of the textured mesh model.

Despite the non–professional tools, it is now known within the scientific community that the processes of generating point clouds from georeferenced photogrammetric blocks provide excellent results even when the starting data is not a set of images acquired with a calibrated photogrammetric camera. [Cardenal et al. 2004]
The Agisoft Metashape software was used for the 3D model generation operations.
The method used has already been tested and considered appropriate for the survey of elements located in the vicinity of the supports used as references but, despite the cost–effectiveness of the process, even at greater distances the uncertainty is within a few centimeters. [Calantropio et al. 2018]

Dataset Construction Workflow

As previously mentioned, the generation of the 3D model, starting from aerial and ground photos, was carried out using the Agisoft Metashape software. Since the final aim was to produce photorealistic renderings in a fast way, it was fundamental to optimize the mesh, firstly, to reduce the calculation time, secondly, to achieve a representation without gaps, without visible polygon edges and with a more homogeneous appearance. A dense cloud of 5,198,304 points was used as input for the mesh calculation, resulting in a mesh of 345,280 faces. Then, to decrease the complexity of the geometry, the coplanar faces were merged and the areas surrounding the tower were decimated. The output of this process was a mesh of 219,269 faces.
An algorithm was written within Rhino's Visual Programming Language (VPL) environment to automatically generate the useful photorealistic views from this last textured mesh. Grasshopper was chosen as the programming environment because of its ability to naturally manage complex geometries. Moreover, being integrated within Rhino, it was possible to easily connect it to different PBR rendering engines (Rhino Render and V–Ray).
In the algorithm, three working phases can be identified:
1. The identification of useful views around the case study. A hypothetical circular path was drawn around the tower. Along this path 26 chambers were positioned, 13 at one elevation and 13 at a slightly higher elevation, with the centre of the tower as the intake point.
2. Through the analysis of the epw weather file of the location, the solar path was imported. The months between the summer solstice and the winter solstice were selected, and for each month 5 moments in the day were selected in order to have a render for each possible significant position of the sun.
3. The last step was to automate the rendering procedure of all the views for each chosen moment during the selected months. We have 26 chambers, for each of them 5 positions of the sun were selected for the 6 chosen months for a total of 780 images (size of each image 480*480 px). The images were exported in .png format with the contour of the architectural object. It was decided on a first hypothesis to include as little context as possible and to avoid representing the sky, in order to prevent the ML algorithm that would be trained on this dataset from identifying features on objects or landscape components (clouds, bushes or other) other than the tower.
A first dataset built in this way was generated in about 3 hours. Many of these steps could be further automated, thus reducing the required time. (fig. 2)

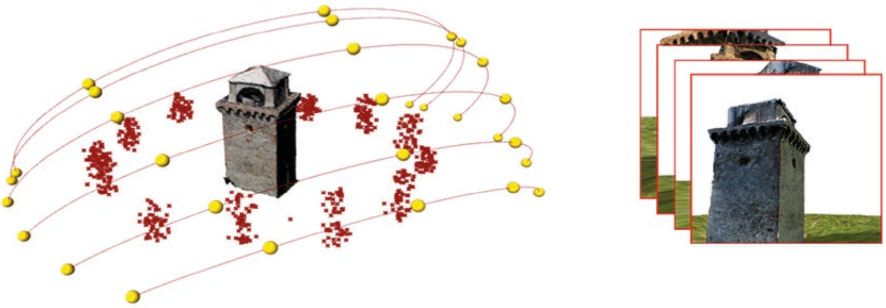


Fig. 2. Selection throughout the year of 30 different lighting conditions and relative camera positions.

Test Classification Pipeline

To test if the data produced by the aforementioned workflow is suitable for training a classification algorithm, a naive image classification pipeline has been created. This has also been useful to identify possible problems with the generated images, such as the issue that will be mentioned in the next paragraph.

The architecture of the pipeline is outlined in Figure 3. The first block is in charge of importing the data as it is. Since in this case we are using a simple binary classifier, we have four different types of data. The two types of data used for training are the pictures generated algorithmically and pictures coming from a public dataset of building facades for the other class. The other two types of pictures used for testing the classifier are real pictures of the tower taken with a camera and other pictures coming from the aforementioned public dataset. The pictures used to train the classifier are then elaborated in the 'Augmentation' and 'Preprocessing' blocks of the script to be ready to be used from the classifier. The test pictures only have to be preprocessed. The classifier is an implementation of a Stochastic Gradient Descent (SGD) [Bottou Leon 2010] classifier which, even if not tuned, allows us to draw some conclusions about whether the data we generated can be useful for image recognition purposes.

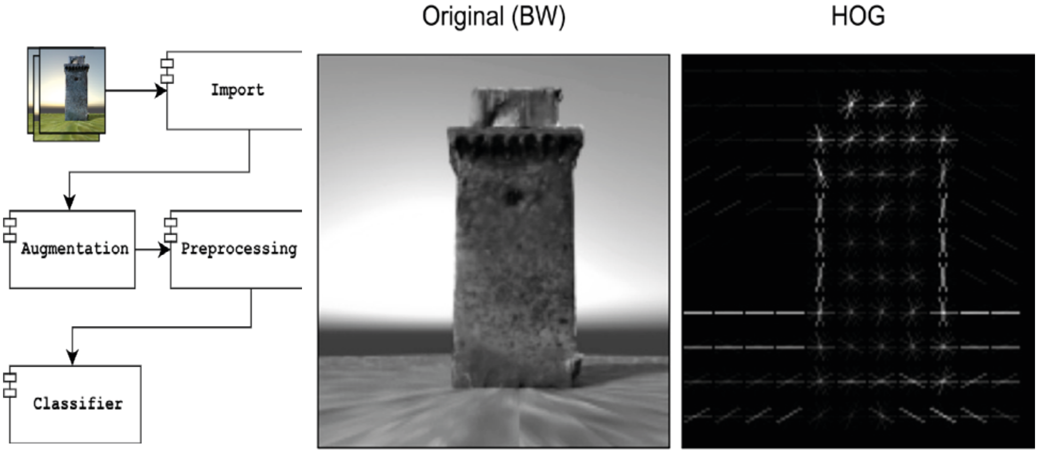


Fig. 3. From the left: Classification pipeline architecture; Two intermediate outputs of the pipeline stages.

Dataset Construction Improvements

After training and testing the classifier on the first dataset, an overfitting problem became immediately apparent. The images, as previously described, were taken with the same camera settings, the same perspective and the tower was always in the center of the picture. As can be seen in Figure 3, to allow the images to be processed by the classifier, they undergo several levels of pre–processing, including the removal of color information. They are then processed by a ML algorithm that calculates the Histogram of Oriented Gradients (HOG), a technique used to select the most interesting features within the image from a software perspective.

If the automatically generated images have a low degree of variability in some regions, which is visible in the HOG data, there is a risk that the classification algorithm will learn to correctly classify only those images with this degree of variability, i.e. the synthetic ones, and misclassify the real images.

To solve this problem, some degree of variability in the relative position of the tower in the image had to be included. The initial VPL algorithm has been modified, spheres have been constructed on the point on which the cameras were initially positioned and for each image the camera has been positioned on a random point belonging to this sphere. A similar solution has also been used for the camera target, with this stratagem the images no longer have the same coordinates and our case study is never in the same position within the image.

Conclusions

Even though it has not been tuned, the classifier scored a ~70% of accuracy in distinguishing images of the tower from images of other buildings.

The benefits of the presented approach versus a more traditional ground survey are clear. In a comparable time we can obtain much more data about the case study allowing to achieve promising results in image recognition with ML algorithms.

With the tested classification algorithm, the background of the rendered 3D object plays an important role as it is processed along with the subject during the pre–processing phase and the resulting information is used during the training of the SGD classifier. The output of the render should in fact have in the background different images which are similar to the true background of the real building, thus adding 'noise' behind the subject and reducing the risk of 'overfitting' the classification model, which would decrease its accuracy.

Including the surrounding scenery into the rendering therefore means providing a more complete context to the subject of the survey. This can be achieved by taking some additional pictures with the drone or by using inexpensive hardware to collect spherical images. This information can be augmented using a 3D rendering engine to include night or dusk settings as well. Moreover this technique also provides the operator a mesh with is useful in the mentioned AR applications.

References

Andrianaivo Louis N., D'Autilia Roberto, Palma, Valerio (2019). Architecture recognition by means of convolutional neural networks. In *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, 42 (2/W15), pp. 77-84.

Barsanti Sara Gonizzi, Malatesta Saverio Giulio, Lella Francesco, Fanini Bruno, Sala Francesco, Dodero Eloisa, Petacco Laura (2018). The Winckelmann300 project: dissemination of culture with virtual reality at the Capitoline Museum in Rome. In *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, 42 (2), pp.371-378.

Bottou Leon (2010). Stochastic Gradient Descent. https://leon.bottou.org/projects/sgd (16th February 2021).

Calantropio Alessio, Chiabrando Filiberto, Spanò Antonia (2018). Impiego di tecniche di fotogrammetria digitale speditiva a supporto delle opere provvisionali negli interventi tecnici in emergenza sismica. In *Bollettino SIFET,* 3, pp. 22-31.

Cardenal Escarcena E.J., Mata Emilio, Castro P., Delgado Jorge, Hernandez M. A., Perez Jose Luis, Ramos M., Torres Manuel (2004). Evaluation of a digital non metric camera (Canon D30) for the photogrammetric recording of historical buildings. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 35 (B5), pp. 564-569.

Liu Li, Ouyang Wanli, Wang Xiaogang, Fieguth Paul, Chen Jie, Liu Xinwang, & Pietikäinen Matti (2020). Deep learning for generic object detection: A survey. In *International journal of computer vision*, 128 (2), pp. 261-318.

Lo Turco Massimiliano, Piumatti Paolo, Calvano Michele, Giovannini Elisabetta Caterina, Mafrici Noemi, Tomalini Andrea, Fanini Bruno (2019). Interactive Digital Environments for Cultural Heritage and Museums. Building a digital ecosystem to display hidden collections. In *Disegnarecon*, 12 (23), pp. 7.1-7.11.

Luhmann Thomas (2010). Close range photogrammetry for industrial applications. In *ISPRS journal of photogrammetry and remote sensing*, 65 (6), pp. 558-569.

Nocerino Erica, Menna Fabio, Remondino Fabio, Saleri Renato (2013). Accuracy and block deformation analysis in automatic UAV and terrestrial photogrammetry–Lesson learnt. In *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2 (5–W1), pp. 203-208.

Norouzzadeh Mohammad Sadegh, Nguyen Anh, Kosmala Margaret, Swanson Alexandra, Palmer Meredith S., Packer Craig, Clune Jeff (2018). Automatically identifying, counting, and describing wild animals in camera–trap images with deep learning. In *Proceedings of the National Academy of Sciences*, 115 (25), pp. 5716-5725.

Spallone Roberta, Palma Valerio (2020). Intelligenza artificiale e realtà aumentata per la condivisione del patrimonio culturale. In *Bollettino SIFET*, 2, pp. 1-8.

Vayanou Maria, Karvounis Manos, Katifori Akrivi, Kyriakidi Marialena, Roussou Maria, & Ioannidis Yoannidis (2014). The CHESS Project: Adaptive Personalized Storytelling Experiences in Museums. In *UMAP Workshops*, p. 232-235.

Authors
Andrea Tomalini, Dept. of Architecture and Design, Politecnico di Torino, andrea.tomalini@polito.it
Edoardo Pristeri, LINKS Foundation, Leading Innovation & Knowledge for Society, edoardo.pristeri@linksfoundation.com
Letizia Bergamasco, LINKS Foundation, Leading Innovation & Knowledge for Society, letizia.bergamasco@linksfoundation.com

Isbn 9788835125280