Pattern Recognition Letters
journal homepage: www.elsevier.com

Egocentric Visitor Localization and Artwork Detection in Cultural Sites Using Synthetic Data (Supplementary Material)

Santi Andrea Orlando[a,b,2], Antonino Furnari[a,2], Giovanni Maria Farinella[a,c,2,**]

[a] *University of Catania, Department of Mathematics and Computer Science, Via Santa Sofia - 64, Catania 95125, Italy*
[b] *DWORD - Xenia s.r.l., Acicastello, Catania, Italy*
[c] *National Research Council, ICAR-CNR, Palermo,Italy*

ABSTRACT

This supplementary material complements the paper. It reports additional details on the dataset and experiments.

1. Introduction

This document is intended for the convenience of the reader and reports additional information about the proposed dataset and the performed experiments. This supplementary material is related to the following submission:

- Santi A. Orlando, Antonino Furnari, Giovanni M. Farinella, "First Person Vision for visitors behavioural understanding and fruition augmentation in cultural sites", Pattern Recognition Letters, 2020.

The reader is referred to the manuscript and to our web page `http://iplab.dmi.unict.it/SimulatedEgocentricNavigations/` for further information.

2. Datasets

The two datasets have been generated using the tool described in the paper. Each dataset comprises simulated navigations in the virtual version of the real environments. 3D scans of the environments have been generated using a Matterport 3D scanner. The 3D model has been imported in the Unity 3D game engine and the tool has been used to collect both datasets. Frames from the simulated navigations have been automatically labelled with the 6 Degrees of Freedom (6DoF) of the camera and with semantic information on the position of specific artworks. In the following sections, we report additional details on the datasets.

2.1. Galleria Regionale Palazzo Bellomo

This dataset is composed by 4 navigations. Each navigation is performed by a virtual agent with height 1.70 *m*. Table 1 reports the amount of frames for each navigation as well as the total number of frames.

To assess the quality of the collected dataset, we performed a study about the density of the positions sampled by the agent during the navigation. The analysis aims at verifying if the whole environment has been navigated in a uniform manner and if all the areas of the environment have been collected from different points of view. We produced 2D density plots of the collected points according to their orientations. Fig. 1 shows density estimation plots of all the poses which orientations fall in the intervals: $[0°, 45°[$, $[45°, 90°[$, $[90°, 135°[$, $[135°, 180°[$, $[180°, 225°[$, $[225°, 270°[$, $[270°, 360°[$. The figure shows that all parts of the building have been navigated by the virtual agent, which larger areas explored slightly more densely as one would expect.

We labelled the 3D mesh of the Bellomo dataset to specify the presence of the 16 artworks listed in Table 2. To make sure that enough images of the artworks are collected by the virtual agent, for each artwork we specify a set of observation points that the agent will visit during the navigation.

Specifically, we select randomly 5 observation points for each artwork. When the agent reach these points during the navigation a "loot at" behaviour is triggered. Examples of each artwork coupled with the respective RGB frame are shown in Fig. 2.

**Corresponding author: Tel.: +39 095 7337 219; fax: +39 095 330094;
e-mail: `gfarinella@dmi.unict.it` (Giovanni Maria Farinella)
[2] These authors are co-first authors and contributed equally to this work.

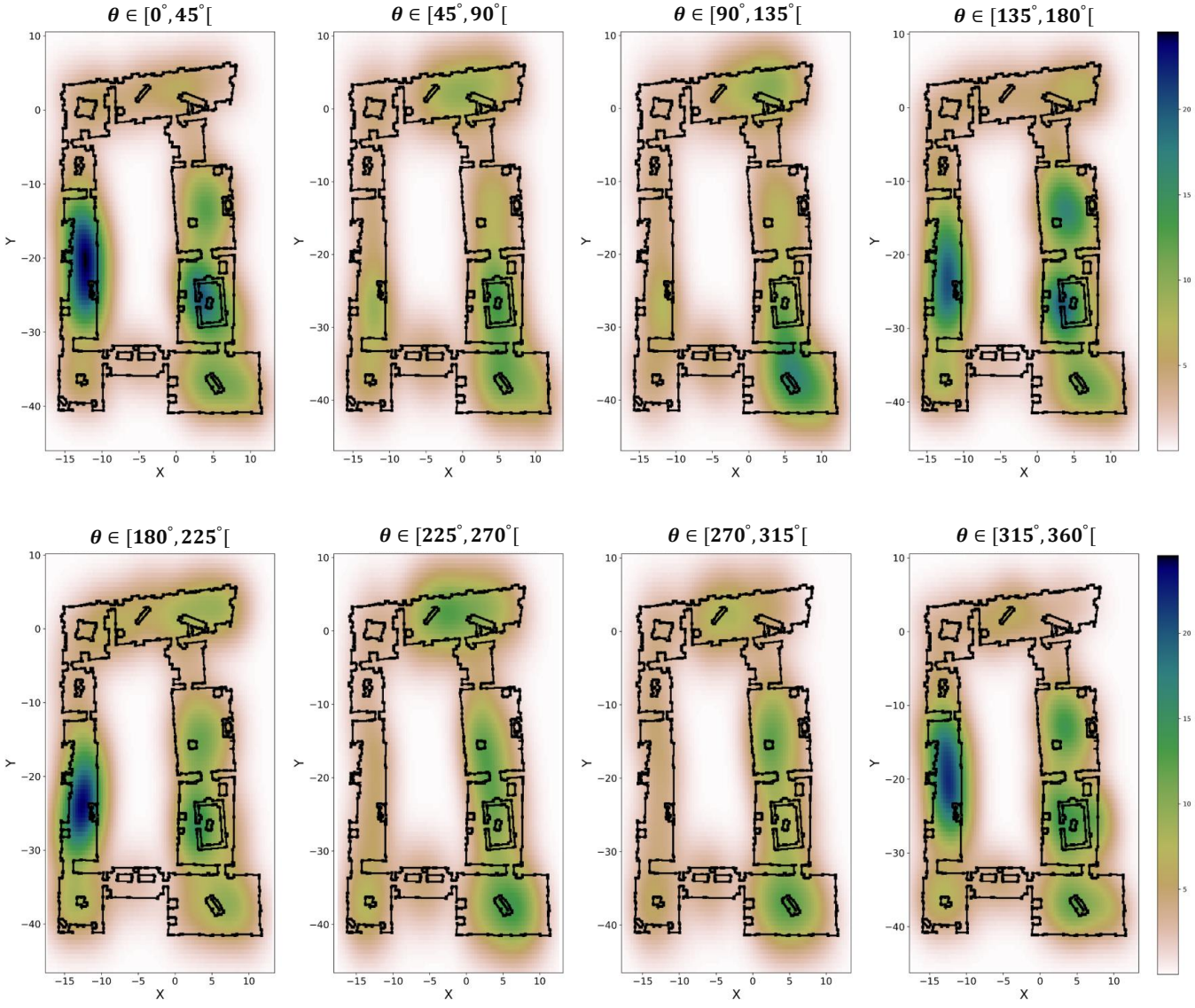

Fig. 1. Examples of 2D density plot of poses with orientations θ falling in intervals showed above each figures.

Table 1. Number of frames contained in each navigation for the Bellomo dataset.

Navigation ID	# frames
1	24, 525
2	25, 003
3	26, 281
4	23, 960
Overall	99, 769

2.2. Area 3 - Stanford Dataset

This environment is an office area from the Stanford "Area 3" dataset used in [1]. This dataset has been collected simulating 30 different navigations for each different height of the virtual agents (1.5 *m*, 1.6 *m*, 1.7 *m*).

Tests are performed on a subset of 10, 000 frames, randomly samples from the navigations of the test set. To assess the influence of the smoothing filters on this dataset, we first apply temporal smoothing to all the test navigations, in a sequential fashion, then report the results obtained on the selected subset of 10, 000 frames.

3. Artwork Detection

Qualitative results highlighting the accuracy of the trained detectors are also reported in Fig. 4.

Acknowledgments

This research is part of the project VALUE - Visual Analysis for Localization and Understanding of Environments (N. 08CT6209090207 - CUP G69J18001060007) supported by PO FESR 2014/2020 - Azione 1.1.5 - "Sostegno allavanzamento tecnologico delle imprese attraverso il finanziamento di linee pilota e azioni di validazione precoce dei prodotti e di dimostrazioni su larga scala" del PO FESR Sicilia 2014/2020,

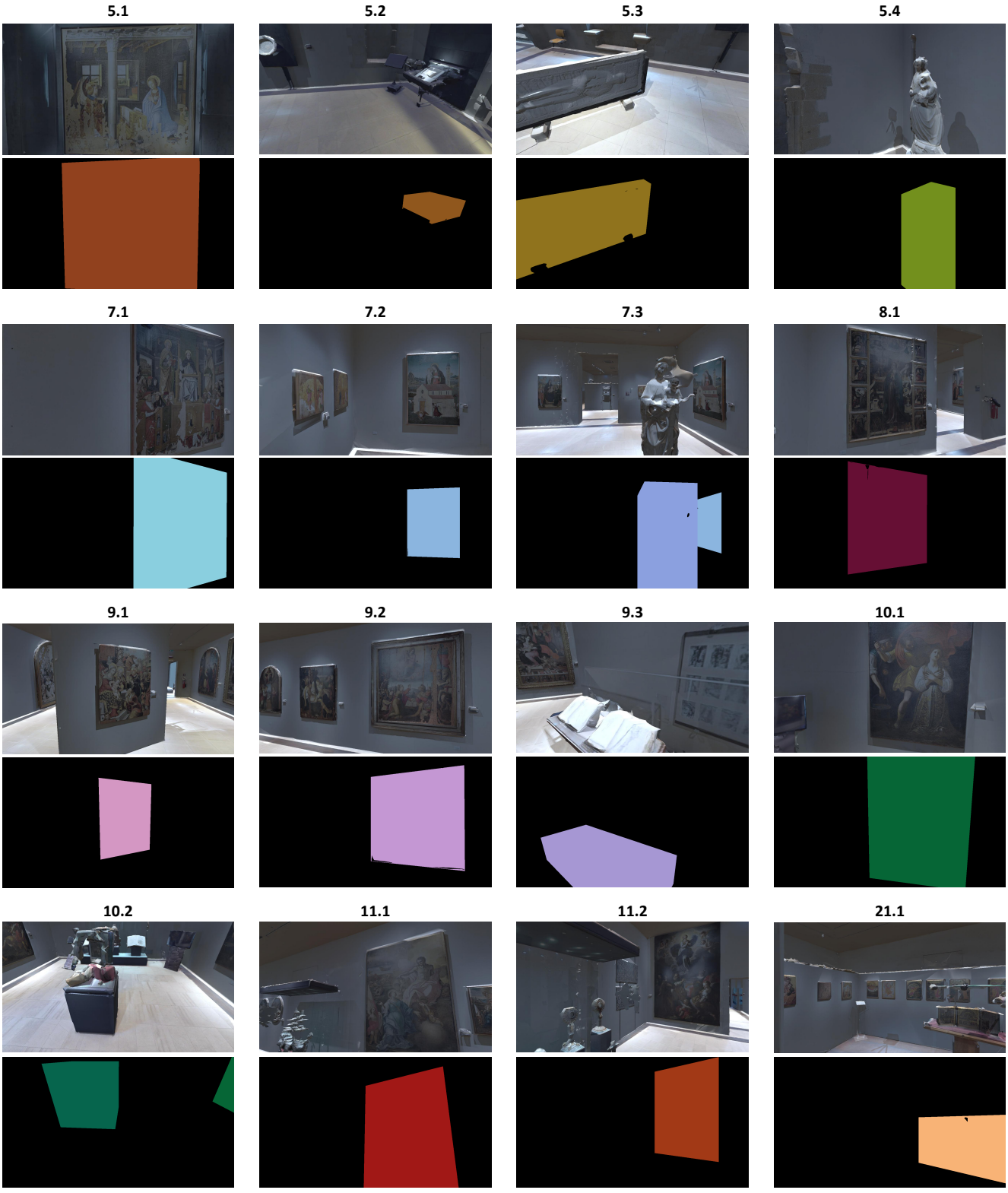

Fig. 2. This figure reports an example for each artwork of RGB frame and respective semantic mask. See Table 2 for further information and colour codes.

Table 2. List of artworks labeled in the Bellomo dataset, with the related coordinates in the 3D model and the RGB color codes used to produce the semantic masks.

ID	Category Name	RGB ID	Coordinates		Semantic Color		
			X	Y	R	G	B
5.1	Annunciazione		2,75	1,23	145	65	30
5.2	Libro d'Ore miniato		6,49	0,53	144	87	29
5.3	Lastra tombale di Giovanni Cabastida		-4,28	2,3	144	115	29
5.4	Madonna del Cardillo		-7,35	-0,2	115	144	29
7.1	Disputa di San Tommaso D'Aquino		0	-14,25	138	207	223
7.2	Traslazione della Santa Casa		0,3	-17,74	139	181	223
7.3	Madonna col Bambino		2,59	-15,48	139	160	223
8.1	L'Immacolata Concezione e Dio Padre in gloria		2,25	-20,58	102	15	52
9.1	Adorazione dei Magi		6,13	-37,67	212	151	195
9.2	Sant'Elena e Costantino e Madonna col Bambino in gloria fra angeli		12,18	-39,84	196	151	211
9.3	Taccuinu di disegni		5,23	-37,27	166	151	211
10.1	Martirio di Santa Lucia		-16	-38,28	6	102	54
10.2	Volto di Cristo		-12,9	-37,27	6	101	77
11.1	Miracolo di Sant'Orsola		-10,5	-27,62	161	24	22
11.2	Immacolata e i santi Chiara, Francesco, Antonio, Abate, Barbara e Maria Maddalena		-13,8	-12,12	162	57	22
21.1	Storia della Genesi		5	-26,75	248	179	118

and Piano della Ricerca 2016-2018 linea di Intervento 2 of DMI, University of Catania. The authors would like to thank Regione Siciliana Assessorato dei Beni Culturali dell'Identità Siciliana - Dipartimento dei Beni Culturali e dell'Identità Siciliana and Polo regionale di Siracusa per i siti culturali - Galleria Regionale di Palazzo Bellomo.

References

[1] S. A. Orlando, A. Furnari, S. Battiato, G. M. Farinella, Image-based localization with simulated egocentric navigations, in: International Conference on Computer Vision Theory and Applications (VISAPP), 2019.

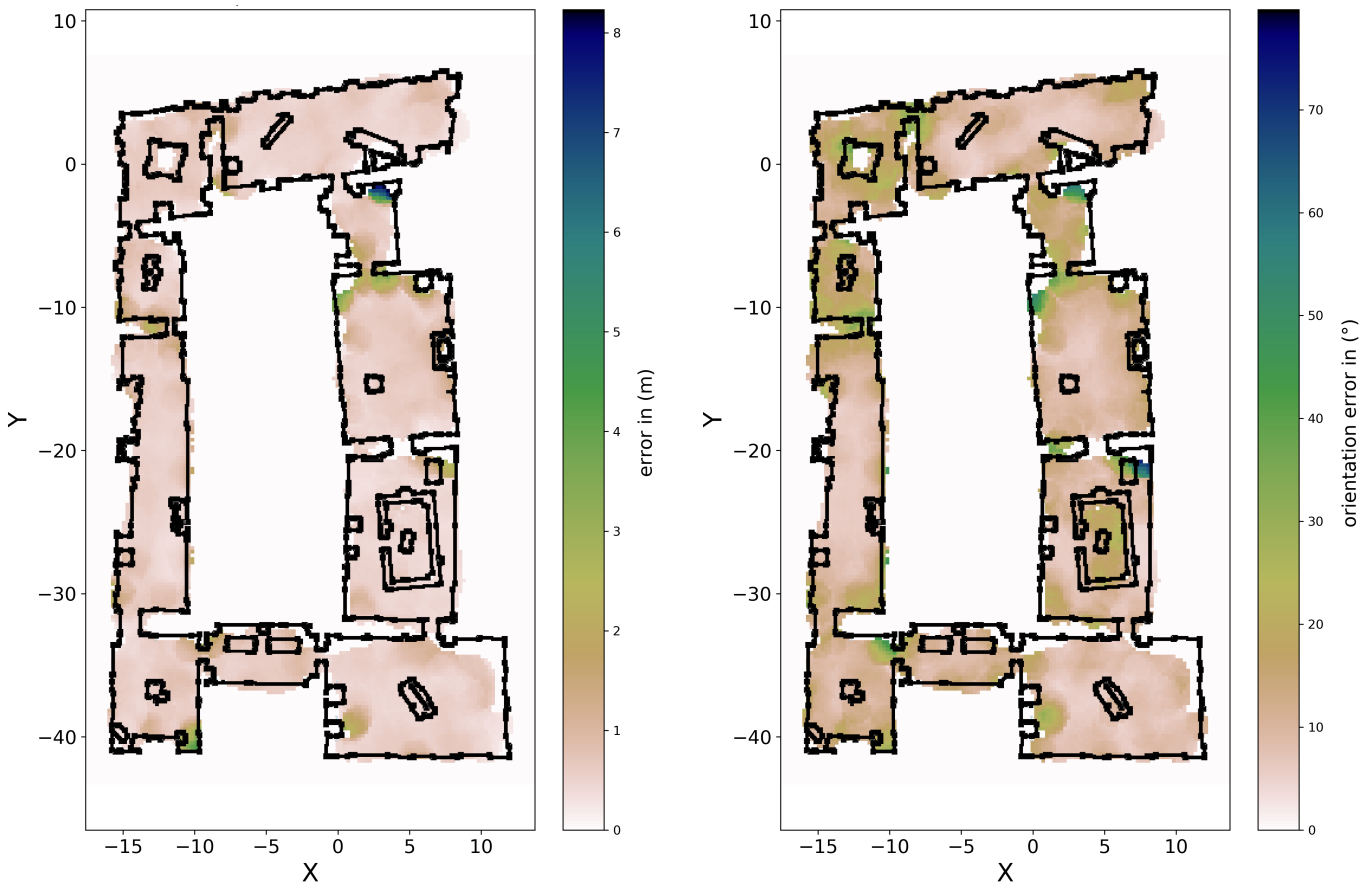

Fig. 3. Average position (left) and orientation (right) errors estimated on the area of the Bellomo environment using a circular Parzen window with radius equal to 1 m.

Fig. 4. Qualitative examples of detectors trained on the Bellomo dataset.