PROJECT
Coach assistant via projected and tangible interface
GRANT AGREEMENT
Nr. 769830

D3.2 – Test dataset

SUBMISSION DUE DATE	DELIVERABLE VERSION
Month 18, 31.05.2019	3.0
ACTUAL SUBMISSION DATE	
Month 18, 31.05.2019	

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 769830

DELIVERABLE TITLE	Test dataset
DELIVERABLE No.	D.3.2
Deliverable Version	3.0
Deliverable Filename	Captain_Deliverable_3.2_TestDataset_v3.0.docx
Nature Of Deliverable	DEM = Demonstrator
Dissemination Level	Public
Number Of Pages	10
Work Package	WP3. CAPTAIN appliance hardware prototyping and low level software infrastructure development
Partner Responsible	HoloLamp
Author(s)	Guillaume Chican (HoloLamp), Despoina Petsani (AUTH), Evdokimos Konstantinidis (NIVELY), Giuseppe Conti (NIVELY), Francesco Verrini (NIVELY)
Contributor(s)	
Editor	Guillaume Chican (HoloLamp)
Reviewed by	Michalis Timoleon (AUTH), Andoni Beristain (VICOM)
Approved by	Panos Bamidis, Project Coordinator

PROJECT FULL TITLE	Coach assistant via projected and tangible interface
Type Of Action	Research & Innovation Action (RIA)
Topic	H2020-SC1-PM-15-2017: Personalised coaching for well-being and care of people as they age
Start Of Project	1 December 2017
Duration	36 months
Project URL	www.captain-eu.org

Table of Contents

1. Executive Summary 5
2. Introduction 5
3. Access to the dataset 6
4. CAPTAIN Box 6
4.1. Face dataset 6
4.1.1. Standard motion 6
4.1.2. Appearance in field of view 6
4.1.3. Extreme poses 6
4.1.4. Under exposure 6
4.1.5. Change of scale 7
4.1.6. Fast head motions 7
4.2. Touch dataset 7
4.2.1. Standard tap 7
4.2.2. Drag motion 7
4.2.3. Multiple touchpoints 7
5. CAPTAIN Satellite Detector 7
5.1. Gait, Posture and Fall detection dataset 8
5.2. Authentication issues and emotion detection 10
5.3. Audio for speech recognition and analysis 10
6. Conclusions 10

LIST of Acronyms

Acronym	Description
AR	Augmented Reality
RGB	Red Green Blue
2D/3D	Two/Three Dimensions
ARM	Advanced RISC Machines
API	Application Program Interface
DNN	Deep Neural Networks
JSON	JavaScript Object Notation
IR	Infrared

LIST OF FIGURES

Figure 1 Setting for gait, posture and fall detection recordings 8

1. EXECUTIVE SUMMARY

The CAPTAIN project aims at enhancing the living experience of older adults in their home while supporting everyday life activities. Older adults are usually reluctant with technology usage with Negative attitudes mostly associated with technology creating inconveniences or unhelpful features. Therefore, the smart home solution interaction should be fully natural to meet older adult needs. CAPTAIN pro-poses an unobtrusive solution that does not constrain the user to wear, hold or operate any wearable equipment but the user interface is provided[1] by projections that create a 3D perception and tangible interfaces by detecting finger movements or interacting with everyday objects through body movements and posture. The tangible interface requires some sensing components. The two CAPTAIN de-vices - the CAPTAIN box and CAPTAIN satellite - embed multiple sensors and run the sensing algorithms for this purpose.

In this context, the scope of this deliverable is to present the different datasets collected in order to test the sensing algorithms. Each dataset corresponds to a specific algorithm and sensor. The dataset organized as follows: the dataset called FaceDataset is related to the face tracking algorithms, the dataset called TouchDetectionDataset is related to the touch interface, the Gait Posture and Fall dataset is related to the analysis of the body movement and posture.

2. INTRODUCTION

The aim of this deliverable is to provide the datasets that will be used to test and provide some KPI about the sensing algorithms. The datasets are used by the high-level components of sensing (WP4) and coaching (WP5). The data have been collected by the sensors embedded in the CAPTAIN box and CAPTAIN satellites. The firmware components necessary to collect the data are described in the deliverable D3.4.

The CAPTAIN Box (*D2.2 – First Version of system specification, §7.2 Custom devices*) is an AR lamp for natural interactions with the 3D content. It tracks the position of the user's head ensuring that the projection is adjusted to the user perspective in real-time to give the illusion of 3D objects. On top of that, it detects the user's hand touch points to create a tangible interface. The CAPTAIN Box embeds a camera with a fisheye lens for the face tracking and a camera behind an IR optical filter synchronized with a modulated laser line for the touch detection. The CAPTAIN box is standalone and embeds an ARM computer: the Nvidia TX2. Finally it has a sound speaker and a microphone.

The CAPTAIN Satellite Detector (*D2.2 – First Version of system specification, §7.1 Off-the-shelf devices*) is based on Nively's Mentorage device. It embeds a 3D depth sensor to extract and analyses the user's body skeleton and silhouette. The main functionality in the CAPTAIN context is to share the user's posture, gesture and relative indoor position (i.e. close to the fridge, etc.) as well as to perform some risk analysis (i.e. fall detection) and act as the distribute microphones of the system.

Section 3 describes the different datasets related to the CAPTAIN box. Section 4 describes the different datasets related to the CAPTAIN Satellite Detector. Finally, Section 5 concludes the D3.2.

[1] Mitzner, Tracy L et al. "Older Adults Talk Technology: Technology Usage and Attitudes." Computers in human behavior vol. 26,6 (2010): 1710-1721. doi:10.1016/j.chb.2010.06.020

3. ACCESS TO THE DATASET

Based on the information provided on D.1.5. – Data Management plan the data sets collected within the CAPTAIN project are divided based on accessibility, storage location, lifespan and access criteria. The data sets collected for the current deliverable are divided

Based on accessibility

Restricted Data: the recorded mqtt messages that contain RGB images of the participants are restricted and stored locally, encrypted on AUTH premises. These data can be accessible only by the CAPTAIN consortium.

Pseudo-Opened Data: Data collected from the CAPTAIN box, skeleton data and audio voice data are pseudonymised and stored in CAPTAIN Cloud and are accessible for institutional research purposes. For the time being the data sets are accessible upon request send by email to the Coordinator Dr. Panos Bamidis (bamidis@auth.gr, pdbamidis@gmail.com) and the Technical Coordinator Dr Evdokimos Konstantinidis (evdokimosk@gmail.com, konstantinidis@nively.com).

4. CAPTAIN BOX

4.1. FACE DATASET

The face dataset is composed of the 6 videos listed below, recorded by the CAPTAIN box face camera. The face tracking pipeline is a combination of DNN face detection and computer vision tracking algorithms.

4.1.1. Standard motion

The recording shows slow lateral motions of a user in front of the camera. The tracking pipeline is test-ed with this recording. The goal is to detect if the tracking pipeline is running well and if the jitter and latency are small enough. The different parameters of the pipeline are fixed with this recording.

The tracking pipeline should also be tested when running with a live video stream. The quality of the tracking depends on the elapsed time between two consecutive frames, which can be higher in a live stream due to the processing time. The algorithms are optimized for the ARM computer selected for the CAPTAIN box (TX2 computer).

4.1.2. Appearance in field of view

The recording shows a user appearing in the camera field of view. The DNN face detection algorithm is tested with this recording. A first KPI is the ratio of detected appearances. A second KPI is the average processing time, which corresponds to the average latency between the time of the appearance of the entire head in the field of view and the time of the detection.

4.1.3. Extreme poses

The recording captures a side view of the face and is used for testing the face detection algorithm. The extreme pose is a challenge for the face alignment algorithms because some face landmarks are not visible in the camera field of view with these poses.

4.1.4. Under exposure

The recording shows an underexposed user's face. The captured face was backlit in order to achieve the underexposure. Bad light conditions are a main challenge for this task that is tested using this re-cording.

4.1.5. Change of scale

The recording shows a user stepping back and then stepping closer to the camera. The tracking algo-rithm is tested with this recording. Change of scale is a main challenge for this task. Each frame is pro-cessed by the face detection algorithm to get a bounding box used as a ground truth. The KPI is the average ratio of the tracked bounding box width over the detected bounding box width.

4.1.6. Fast head motions

The recording shows fast lateral motions of a user in front of the camera. The tracking algorithm is tested with this recording. Fast motions are a challenge for the algorithm. Again, each frame is pro-cessed by the face detection algorithm to get a bounding box used as a ground truth. The KPI is the average ratio of the tracked bounding box position over the detected bounding box position at each frame.

4.2. TOUCH DATASET

The touch dataset is composed of the 3 videos below shot by the CAPTAIN box camera dedicated to the tangible interface. The infrared filter of the camera has been removed in order to see the infrared illumination of the laser line. On top of that an optical bandpass filter for the range corresponding to the laser line illumination has been placed in front of the camera lens in order to only see the laser line.

4.2.1. Standard tap

The video shows some hand gestures, including simple tap gestures. The touch detection pipeline is tested with this recording. Different parameters of the pipeline are fixed. The KPIs tested are

- the average ratio of detected touch points
- the accuracy of the touch point coordinates
- the average processing time, which corresponds to the average latency.

The pipeline should also be tested when running a live video stream. The feeling of quality and low latency in the touch response cannot be easily estimated in a recording.

4.2.2. Drag motion

The video shows a user's finger dragging across the projection surface. The ability of the pipeline to detect drag motion is tested with this recording.

4.2.3. Multiple touchpoints

The video shows a user's finger tapping the projection surface with several fingers at the same time. The ability of the pipeline to detect multiple points at the same time is tested with this recording. A first KPI is the average ratio of detected touch points. A second KPI is the accuracy of the touch point coordinates. A third KPI is the average processing time.

5. CAPTAIN SATELLITE DETECTOR

CAPTAIN satellite device is based on Nively's Mentorage 3D depth sensor and is the main device responsible for streaming images for different features, including:

- gait and posture analysis that enables both risk detection and avoidance and interaction with CAPTAIN system through body movement,

- authentication issues and emotion detection,
- audio for speech recognition and analysis

To this end different datasets are needed in order to test the higher level CAPTAIN algorithms. The datasets collected are described below.

5.1. Gait, Posture and Fall detection dataset

The gait and posture dataset was recorded using the Mentorage device which was placed at 1m height above a TV screen as shown in the figure 1. A mattress (10cm thick) was placed on the floor to ensure the participants' safety. XX participants aged XX took part at the recordings. The recordings experiment received ethical approval from AUTH's ethical committee and all the participants participated voluntarily and have provided their consent.

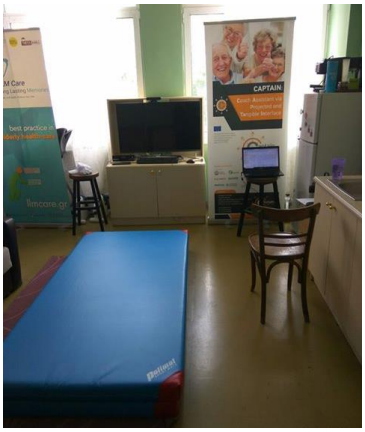

Figure 1 Setting for gait, posture and fall detection recordings

Category	Code	Name	Kind
Backward fall	BF01	Ending sitting	Positive
	BF02	Ending lying	Positive
	BF03	Ending in lateral position	Positive
	BF04	With recovery (stumbling)	Negative
	BF05	Ending sitting and rise up immediately	Positive
Forward fall	FF01	On the knees	Positive
	FF02	With forward arm protection	Positive
	FF03	Ending lying flat	Positive
	FF04	With recovery (stumbling)	Negative
	FF05	Ending sitting and rise up immediately	Positive

Category	Code	Name	Kind
Lateral fall to the right	LFR01	Ending lying flat	Positive
Lateral fall to the left	LFL01	Ending lying flat	Positive
Neutral	N01	Sitting down on a chair then standing up	Negative
	N02	Lying down on the bed then rising up	Negative
	N03	Walking a few meters	Negative
	N04	Bending down, picking something up on the floor, then rising up	Negative
	N05	Coughing or sneezing	Negative
Collapse seated	C01	Collapse forward when seated in a chair	Positive
	C02	Collapse left when seated in a chair	Positive
	C03	Collapse right when seated in a chair	Positive

This dataset is exploited from CAPTAIN technical partners to test the falling detection and body pos-ture recognition. To collect this dataset the CAPTAIN MQTT broker was exploited (*D2.2 – First Version of system specification, §4.3 Protocol*). CAPTAIN Satellite Detector devices captured the RGB images and also recognized and post 16 skeleton point coordinates (X,Y,Z axis) in the mentorage/rgb and mentorage/skeleton mqtt topics respectively. A record and playback tool were created in order to record the mqtt messages in a txt file and the playback can recreate the messages sequences in the same order and timestamps as the original dataset.

The messages captured from the mqtt topics and consist the collected dataset have the following form.

TOPIC: mentorage/skeleton

DATE:
1558705269490{"Device":{"ID":2,"DeviceID":"","PublicIP":null,"LanIP":null,"GUID":null,"DeviceType":1,"
HardwareVer-si-
on":"mentoragev2.0","LastUpdateDateTime":"","SkeletonNumber":"106436","SocketID":null,"SessionID
":"2","DeviceExceptionCmd":0,"ObligatoryTransmission":false},"Frame":{"ms":1455279527074},"Frame-
Num-
ber":0,"Timestamp":0,"TrackingMode":0,"Skeletons":[{"TrackingId":2,"Position":{"X":0.0,"Y":0.0,"Z":0.0}
,"Joints":[{"JointType":0,"Position": {"X": 0,"Y": 0,"Z": 0},"TrackingState":0},{"JointType":1,"Position":
{"X": 130.523,"Y": 25.1497,"Z": 3024.52},"TrackingState":0} […] }

TOPIC: mentorage/rgb

DATE:
1558705270997{"rgbImage":"\/9j\/4AAQSkZJRgABAQAAAQABAAD\/2wBDAAIBAQEBAQIBAQECAgICAgQ
DAgICAgUEBAMEBgUGBgYFBgYGBwkIBgcJBwYGCAsICQoKCgoKBg-
gLDAsKDAkKCgr\/2wBDAQICAgICAgUDAwUKBwYHCgoKCgoKCgoKCgoKCgoKCgoKCgoKCgoKCgoKCg
oKCgoKCgoKCgoKCgoKCgr\/wAARCADwAUADASIAAhEBAxEB\/ […] }

The RGB image is in base64 format and the skeleton is represented by 3D points.

5.2. Authentication issues and emotion detection

Regarding the authentication of the user the same dataset as posture and fall detection is used. As there are X participants taking part in this experiment and perform different movements this dataset is also considered appropriate for testing user authentication.

Concerning emotion detection, the RGB images taken are not yet in high resolution. For the time being external datasets are used to test this component. The CAPTAIN technical team works towards improving the image resolution in order to produce a high resolution dataset from the CAPTAIN satellite.

5.3. Audio for speech recognition and analysis

The CAPTAIN satellite device has been configured to record audio when the level of sound is above a defined value. Once the level of sound is getting above the threshold level and as long it remains above this level (with maximum time of recording 60 mins), an audio file is generated in wav format and stored locally for 5 minutes. Whenever a new audio is available the device will be able to serve it. For this reason, a message is published on the topic events/s05/audio with the address of the file in the payload. This file will be usable, for the next 5 minutes from its creation, from all the other devices connected to the same network of the CAPTAIN satellite device.

The dataset recorded for testing speech recognition and analysis module collected using the CAPTAIN satellite device. The Audio file can be taken directly from the device in the next 5 minutes from his creation or downloaded directly from any internet browser. 6 participants took part in the experiment in order to collect the audio files. Each participant where asked to say the following commands speaking normally in 3 different distances from CAPTAIN satellite device: 1 m, 3 m. The participants are not native English speakers, but the English language was used in order to be easily exploited by all the CAPTAIN partners.

Commands

- “CAPTAIN can you suggest some recipes for lunch?”
- ”CAPTAIN, remind me at ten minutes to add spaghetti in the water”
- “CAPTAIN show me my friends phone list”
- “CAPTAIN show me the shopping list”

6. Conclusions

The datasets have been presented in this report as well as the corresponding tests. The datasets are used to test and validate the different sensing algorithms that are embed in the two CAPTAIN hard-ware devices, CAPTAIN box and CAPTAIN satellite: the algorithms for the analysis of the body move-ment and posture, the face image detection for head tracking and the tangible interface. The extract-ed datasets are used in order to check low level sensing functionalities and the quality of interaction with the CAPTAIN user.