CogWatch – Cognitive Rehabilitation of Apraxia and Action Disorganisation Syndrome

D2.2.2 Report on devices II

Deliverable No.		**D2.2.2**	
Workpackage No.	**WP2**	Workpackage Title	**System Devices and Networks**
Task No.	**T2.2**	Activity Title	**Monitoring and Feedback Devices**
Authors (per company, if more than one company provide it together)		**José M. Cogollor, Javier Rojo, Manuel Ferre, Rafael Aracil and José María Sebastián (UPM-ROMIN), Matteo Pastorino, Alessio Fioravanti and Maria Teresa Arredondo (UPM-LST), Manish Parekh, Martin Russell, Hoi Fei Kwok and Chris Baber (UOB-EECE), Ricardo Ruiz (RGB)**	
Status (F: final; D: draft; RD: revised draft):		**F**	
File Name:		**Cogwatch_D2.2.2_Devices_II**	
Project start date and duration		**01 November 2011, 36 Months**	

EXECUTIVE SUMMARY

This deliverable identifies devices implemented in the second prototype of CogWatch project which focuses on toothbrushing. The second prototype draws on some components used in the first prototype (tea making) including the instrumented coaster for actions involving the glass of water and Kinect^{TM} for face motion tracking. New sensor elements include the Leap Motion System, to detect and track hands/fingers, an inertial sensor (Shimmer 3) and a microphone for tracking movements of the toothbrush in the mouth. Finally a new modular version of the Non Invasive Blood Pressure (NIBP) monitor is developed for heart rate and blood pressure.

The updated architecture for this new prototype is presented and the needed set up for the installation of the devices is suggested considering the limitations and constraints of them.

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF TABLES

REVISION HISTORY

Revision no.	Date of Issue	Author(s)	Brief Description of Change
V0	09/06/2014	UPM	First draft of the contents.
V1	12/06/2014	UPM	Sections: 1 and 3.1.1.
V2	10/06/2014	RGB	Section 3.1.7.
V3	17/06/2014	UPM	Sections: 3.2.1 and 4.
V4	21/06/2014	UoB	Sections: 3.1.3, 3.1.4 and 3.1.5.
V5	21/06/2014	UPM	Sections: 3.1.2, 3.2.2 and 4.
V6	23/06/2014	UoB	Section 3.1.6.
V7	24/06/2014	UPM	Section 2 and Conclusions.
V8	18/07/2014	UPM	Sections: 3 and 4. Final review of content.
Final Draft	21/07/2014	UPM	Final Draft ready for Peer Review
Final	01/10/2014	UPM	Final version including Peer Review comments.

LIST OF ABBREVIATIONS AND DEFINITIONS

Abbreviation	Abbreviation
AADS	Apraxia and Action Disorganization Syndrome
ADL	Activities of Daily Living
API	Application Programming Interface
BBL	Back Bottom Left
DFT	Discrete Fourier Transform
DMO	DirectX Media Object
DoF	Degrees of Freedom
GMM	Gaussian Mixture Model
IPC	Ingress Protection Code
IMU	Inertial Measurement Unit
IR	Infrared
LED	Light - Emitting Diode
MFCC	Mel Frequency Cepstral Coefficients
NIBP	Non Invasive Blood Pressure
PCB	Printed Circuit Board
PHS	Personal Healthcare System
SDK	Software Development Kit
VTE	Virtual Task Execution

1. INTRODUCTION

This section introduces the deliverable "D2.2.2 Report on devices II" for the second prototype of CogWatch project. It provides the main purpose and scope of the system.

1.1 Purpose

This Report on devices II describes the components and functions of the devices considered in the second prototype P2 in order to better know the functionalities to be provided, hardware and software components, and especially, advantages/disadvantages when installing at the set up, regarding the complexity of the scenario.

CogWatch requires the development of novel tools and objects, portable and wearable devices as well as ambient systems to personalise the cognitive rehabilitation carried out at home for stroke patients with Apraxia and Action Disorganization Syndrome (AADS) symptoms. The main specifications of the devices were presented in the previous deliverable "D2.1 Report on system specification".

1.2 Scope

Traditional approaches to rehabilitation for AADS include the therapist providing verbal or visual cues as the patients perform selected tasks (such as preparing a drink or food). Repeated practice with encouragement from the therapist, leads to improving performance and avoiding mistakes that can occur in a trial–and–error approach and at the same time avoiding danger to the patient.

This approach is labour intensive for therapy time and also limits the opportunity for the patient to practice the skill. In addition, the occurrence of mistakes can lead to their perseveration. For these reasons, the CogWatch project proposes a novel solution to AADS cognitive rehabilitation by providing an instrumented environment in the corresponding scenario, including motion tracking with action recognition software.

Moreover, CogWatch can lead to be a Personal Healthcare System (PHS) that delivers personalised, long–term and continuous cognitive rehabilitation of Activities of Daily Living (ADL) for stroke AADS patients at home using portable, wearable and ubiquitous interfaces and virtual reality modules. It is customized to suit the needs of individual patients at the same time as being practical and affordable for home installation so that rehabilitation takes place in familiar environments performing familiar tasks.

The CogWatch system is being developed in relation to a set of scenarios involving ADL tasks described in *"D1.1. Activities of Daily Living*" that comprises tasks of basic self-care such as preparing food and drinks, using the toilet, or washing and grooming in the bathroom. Their performance involves a sequence of component actions on environmental objects directed at some desired end goal. It is thought that successful performance depends on specifying object actions in spatiotemporal terms at a higher cognitive level and then elaborating these into specific movements of limbs, which are monitored as they progress against expected sensory consequences (see “D1.2 Literature review”).

The first prototype P1 focused on the task of tea making and the devices used were described in D 2.2.1. In the second prototype P2, the ADL task selected to develop the system is focused on tooth brushing. This task involves the following components: toothbrush, toothpaste tube (mounted in vertical dispenser), jug of water, mug (for cleaning brush and rinsing mouth), bowl (for spitting water from the mouth), towel (for drying mouth).

2. GENERAL ARCHITECTURE FOR SECOND PROTOTYPE

Figure 1 presents a description of the whole architecture proposed for the second prototype in terms of devices and software modules.

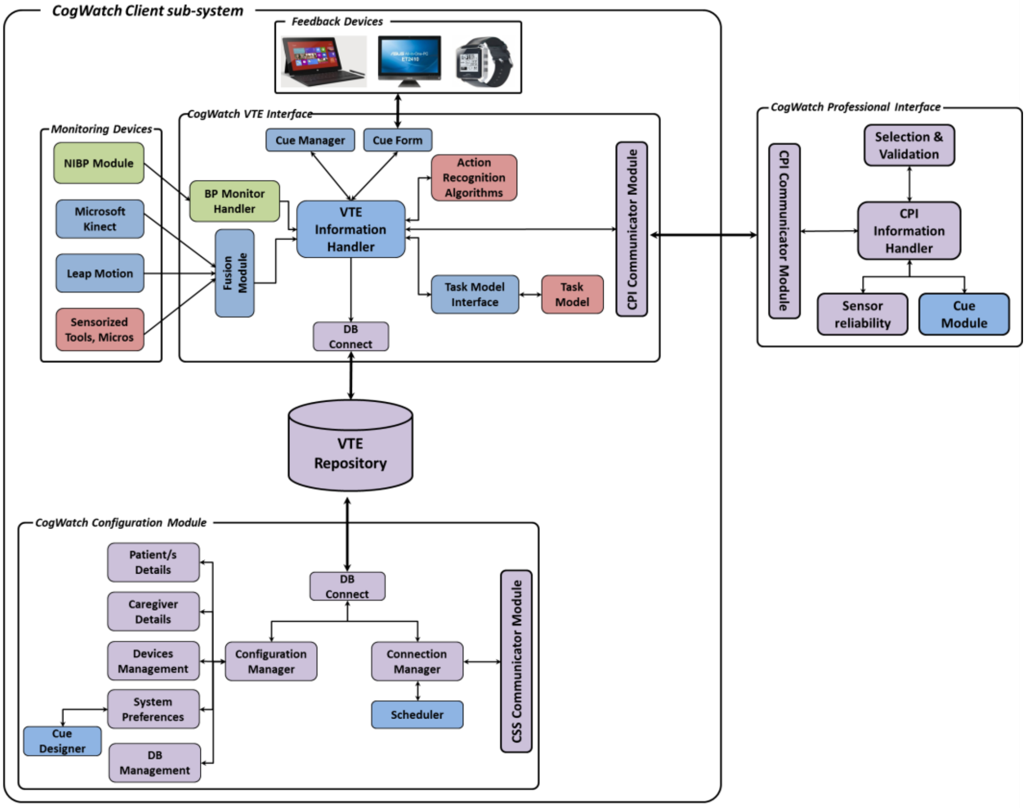


Figure 1. Architecture for second prototype.

This representation is also complemented by that one in the corresponding deliverable *“D2.3.2 Report on networks II”*. The schematic closely follows that implemented in the first prototype but with some differences, especially in the addition of an independent CogWatch Configuration Module which runs continuously in parallel with the CogWatch Client sub-System.

As this document is mainly focused on the description of the devices that interact with the patient, only these will be detailed and explanations about software modules considering the *CogWatch Configuration Module* or the *CogWatch Professional Interface* are presented in *“D2.3.2 Report on networks II”*.

The devices to be considered are also divided into two categories as presented in the first prototype: monitoring devices and feedback devices, which are in charge of monitoring the execution of the task and movements of the patients and providing the corresponding cues and feedback to make the patient aware of the errors committed, respectively.

First of all, to the monitoring devices used for the first prototype are added some new sensors such as Leap Motion or Shimmer3. These are used in order to obtain better data of the hands movements, vital signs from patient and behaviour of the wrist movements of the patient, respectively. Meanwhile, an improved version of the NIBP monitor is also implemented. Finally, in this case, the information provided by Kinect™ is focused on face tracking and head poses.

Considering the feedback devices, the All-In-One computer supporting the VTE and patient interface is replaced in P2 by a tablet PC taking into account that the task will be performed in the bathroom and it offers a more ergonomic solution. The objective of the tablet is the same as for first prototype: Virtual Task Execution (VTE) of the action; verbal and text cues; tactile user interface; and collection of all the corresponding data recorded to process.

The functionality of the devices proposed is detailed in the rest of this document.

3. COGWATCH SECOND PROTOTYPE DEVICES

This section describes in more detail all the devices defined above in terms of hardware, software and limitations in the usage, taking into account the classification proposed and the task to be executed.

3.1 Monitoring Devices for Toothbrushing

As defined in previous deliverables, the monitoring devices are in charge of monitoring and supervising the execution of the task in terms of human movements and object manipulation.

So, reliable and accurate data is mandatory in order to be able to recognize the actions executed by the patient which comprise a wide range of steps involving the manipulation of the toothpaste, glass of water and the approach of the toothbrush to the mouth.

3.1.1 Kinect™ for Face Tracking

In the first prototype, hand motion tracking during tea making was provided by Kinect™ (http://www.microsoft.com/en-us/kinectforwindows/). For this second prototype, the main functionality to be provided by Kinect™ is tracking of the face in order to provide data about the patient’s mouth location as it is approached by the toothbrush.

Regarding the way Kinect™ detects and tracks the body segments, hands tracking while brushing teeth is not reliable since they are directly in contact with the face. The current software provided by default Software Development Kit (SDK) gets confused when tracking hands that are in contact with the own skeleton of the user (this limitation was already presented when dealing with tea preparation in the first prototype).

For that reason, hand tracking is provided by another device, Leap, presented later in the document.

Toothbrushing is a task commonly carried out at the bathroom with the use of a mirror. The second prototype can get benefits from the fact that Kinect™ is able to detect faces via reflection in a mirror so it improves the conditions for installation since the device can be placed behind/over the patient, fixed in the wall or hanging from the ceiling.

Finally, face movements are also detected which can provide useful information about facial gestures in the future.

3.1.1.1 Hardware description

The hardware description is the same as explained in "D2.2.1 Report on devices I". [1]

3.1.1.2 Software description

Although the software description was also detailed in "D2.2.1 Report on devices I", the way both the software libraries and hardware interact with the corresponding application is updated in Figure 2:

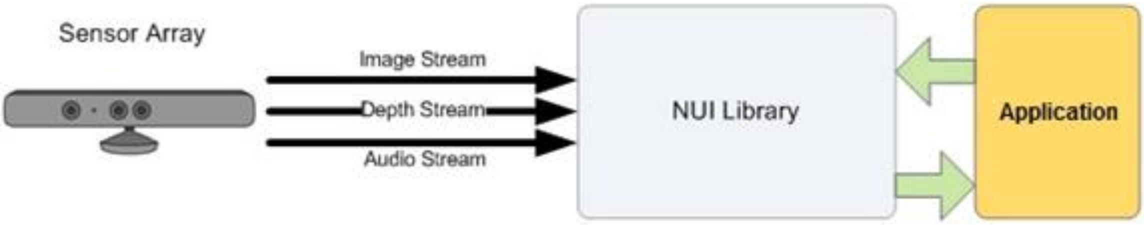


Figure 2. Interaction between HW/SW and application. [2]

The Software Development Kit (SDK) architecture (Figure 3) is composed by:

- The corresponding hardware components, such as the own sensor and the corresponding USB connection devices.
- The corresponding drivers installed when installing the SDK. The drivers make possible the communication with the microphone and audio/video controls.
- Audio/video controls which let the user obtain information from audio and color/depth images.
- DirectX Media Object (DMO) for microphone and audio source localization. [2]
- The corresponding APIs for the development of real time applications.

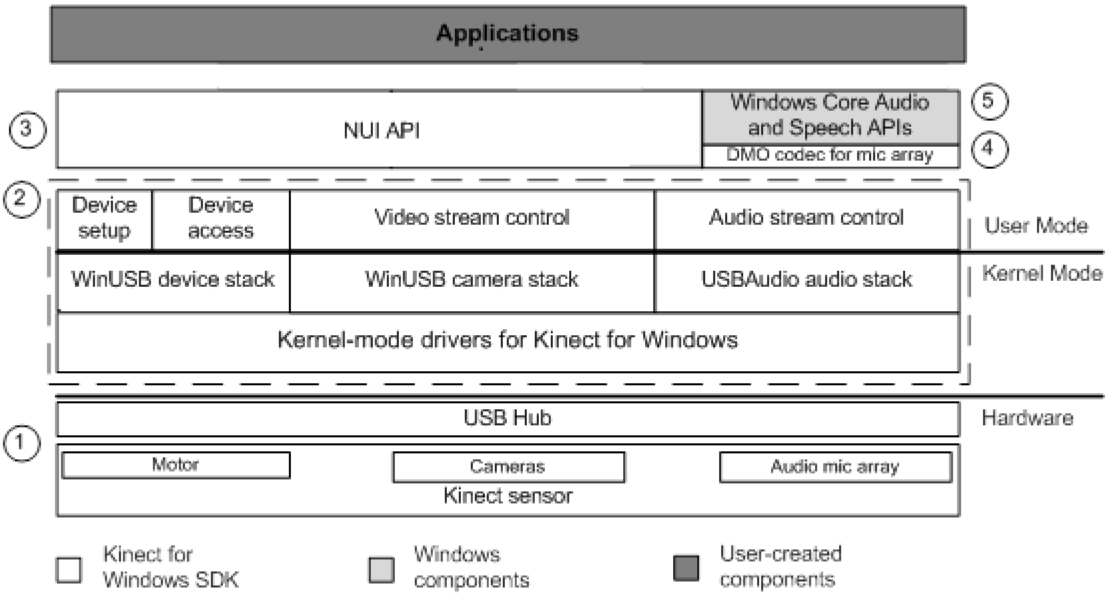


Figure 3. Architecture of the Software Development Kit. [3]

This specific software included the default Windows SDK, makes possible the tracking of human faces in real time applications, such as the CogWatch interface. [3]

The face tracking libraries obtain data from the image provided by the corresponding camera in order to analyze and estimate head poses or gestures. In this case, this is extremely important to provide real time information during the approach of the toothbrush to the mouth and in order to have this step clearly segmented.

From all that information provided and calculated from the image, the most important values are those related to the boundary of the mouth, which is composed by the following points:

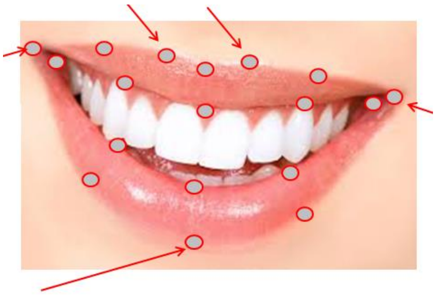


Figure 4. Points detected over the mouth. [4]

A specific 3D workspace around the mouth can be defined by building a rectangle composed by those points highlighted with rows, which are obtained in a specific array. The software provides additional points related to eyes, nose or boundary of head which are not initially considered.

Figure 5 indicates the reference system used to represent the points detected and tracked in the face and especially in mouth. The center of the coordinate system coincides with the optical center of the camera. The distances are expressed in meters by default and the angles for rotation movements in degrees.

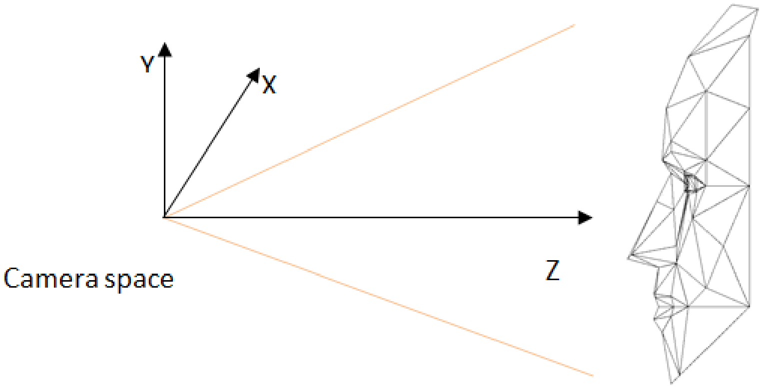


Figure 5. Reference system used by Kinect^{TM} for face tracking. [5]

Although not used directly in the prototype, the software also gives additional information about head poses as indicated in the following figure:

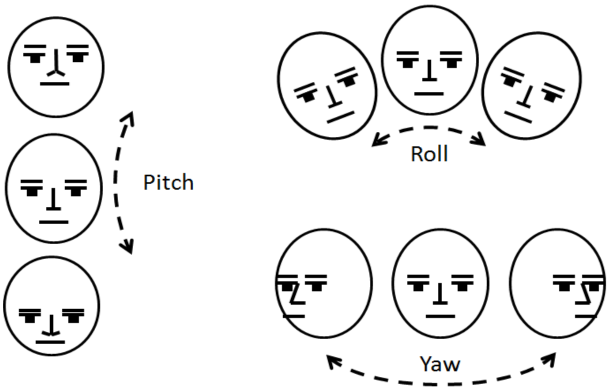


Figure 6. Head poses detected by Kinect^{TM}. [6]

3.1.1.3 Advantages, constraints and technical specifications

First of all, the main advantages Kinect^{TM} provides in this new task of toothbrushing are:

- There are enough points in order to consider a work region around mouth to detect if the toothbrush approaches it or not.
- The current software provides information about the head poses as long as the head is not occluded.
- The head is detected and tracked although the user is manipulating a toothbrush around the mouth.
- Mirror reflected images of the face can be processed (Figure 7); thus the Kinect^{TM} sensor can focus on the user from behind while he or she faces the mirror.
- The sensor can detect up to 2 faces simultaneously.

Figure 7. Face detection in mirror.

Meanwhile, some constraints must be considered when using the sensor for the purpose of face tracking:

- If the toothbrush is bigger than the mouth and occludes it, Kinect^{TM} gets confused by assigning the movements from the toothbrush to mouth.
- The sensor should be just in from of the mirror in order to have a bigger workspace and better detect the head poses. So, the best placement would be behind or over the patient (hanged from the ceiling or fixed on the wall).
- The face tracking is lost when not the whole face is detected. So, when spitting, for example, or when occluding the face with the hand, the sensor is not able to track the face of the patient.
- Z coordinate is modified when detecting reflection in mirror. In this case, Z coordinate is sum of: the distance between Kinect^{TM} and the mirror; and the distance between the mirror and patient. So, small variations in the distance between the patient and the mirror mean bigger separation from the sensor.

Finally, regarding some technical specifications the following points must be highlighted:

- Range of distances for detection: [0.8, 2] m.
- Optimal distance: 1.5 m.
- When working with reflection in mirror: $Z_{final} = Z_{sensor-mirror} + Z_{mirror-patient}$; X and Y coordinates are exchanged.

3.1.2 LEAP Motion for Hand Tracking

For the toothbrushing scenario, the dimensions of the workspace and restrictions of reflecting surfaces lead to search for a different solution for vision recognition. LEAP motion (www.leapmotion.com) affords a solution for the tracking of hands and tools in a reduced workspace such as the proposed set up regarding toothbrush. (See section 4)

The Leap sensor provides finger, hand and tool tracking. The device has an impressive welcome in the market and companies such as HP and ASUS have embedded the device into upcoming/current laptop PCs allowing the option of cursor control in place of the conventional mouse.

What the device provides is an accurate and reliable position of hands and finger joints. Furthermore, the device can provide information about position and orientation of the most salient point of a tool.

Figure 8 shows the possible points provided by hand tracking with LEAP Motion Application Programming Interface (API) 2.0:

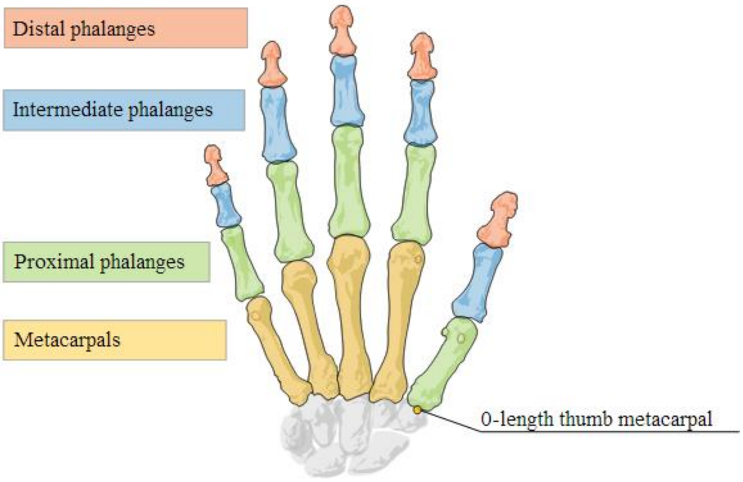


Figure 8. LEAP Motion skeletal tracking. [7]

The last version of SDK provides more information by adding to the position and orientation of hands, tools and fingers the confidence level of the acquired data and the position and orientation of each of the bones of the fingers, as well as the metacarpal bones, which connect the fingers to the base of the hand.

All this data is suitable for recognition of gestures like grasping or pointing which are relevant for the detection of tasks in prototype 2.

3.1.2.1 Hardware description

LEAP Motion is a peripheral device which can be connected to the USB port of a computer. It makes use of two monochromatic IR cameras and three infrared LED's with a workspace which consists of a pyramidal area, to a distance of about 1 meter.

Figure 9 shows the main board of the device:

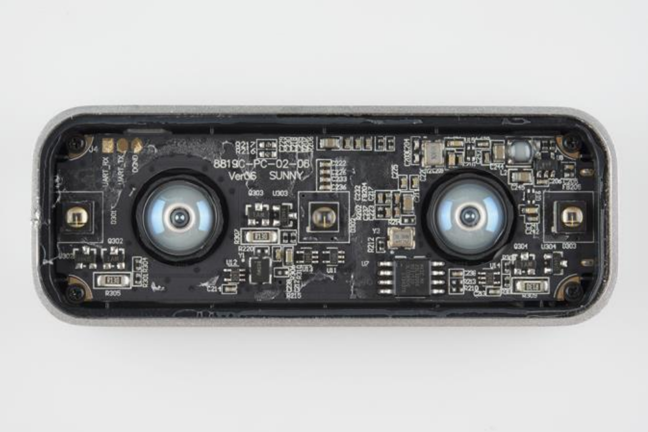
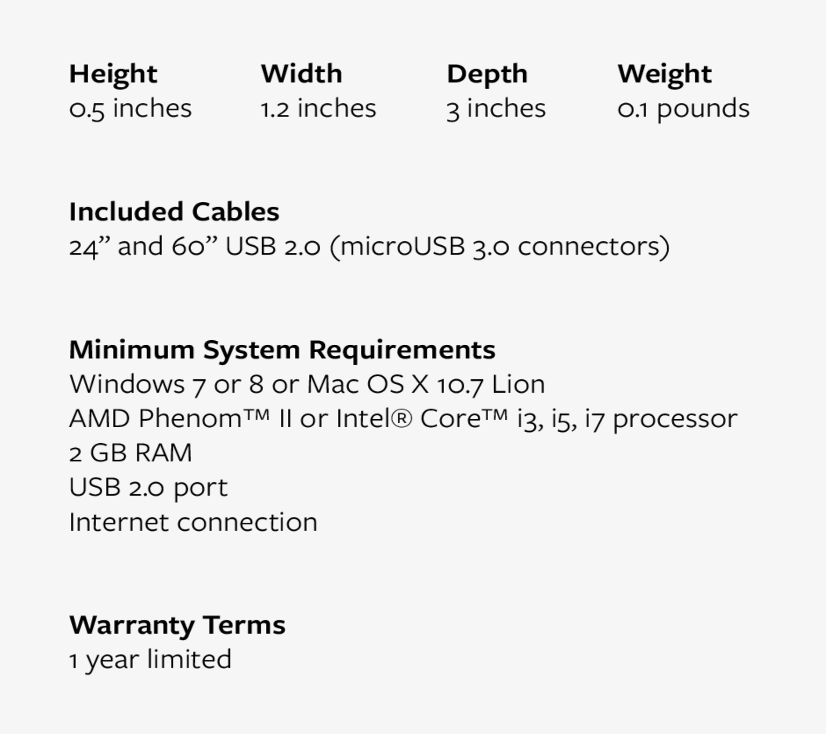


Figure 9. Leap Motion Teardown. [8]

According to the minimum specifications required, the manufacturer provides the following components and characteristics:



Height
0.5 inches

Width
1.2 inches

Depth
3 inches

Weight
0.1 pounds

Included Cables
24" and 60" USB 2.0 (microUSB 3.0 connectors)

Minimum System Requirements
Windows 7 or 8 or Mac OS X 10.7 Lion
AMD Phenom™ II or Intel® Core™ i3, i5, i7 processor
2 GB RAM
USB 2.0 port
Internet connection

Warranty Terms
1 year limited

Figure 10. LEAP Motion specifications. [9]

It is necessary to point out that internet connection is only necessary in case of using leap as a web service but the device can work without the need of an internet connection.

3.1.2.2 Firmware and drivers description

To run in any system, LEAP Motion requires its own driver to be installed. Drivers are supported and provide by the manufacturer from this link: https://developer.leapmotion.com/.

The device is supported by a community of developers, so it is necessary to log in to access drivers and additional software for developing LEAP applications. Once the driver is installed, the base software allows optimizing settings for the device and recalibrating it.

At the time this document was written, the current version of LEAP motion SDK was 2.0 with the programming libraries included in "LeapDeveloperKit_2.0.3+17004". But further updates are expected as they are continuously done by the update software included in the Leap Motion Control Panel, so improvements are constantly incorporated.

3.1.2.3 Advantages, constraints and technical specifications

First of all, the main advantages LEAP Motion provides in this new task of toothbrushing are:

- ➢ The toothbrush can be detected when static or in movement.
- ➢ A simple coordinate transformation allows comparison with Kinect^{TM} data.
- ➢ It can recognize hands through a transparent (for IR) material (tested with an acrylic layer of 7mm thick) as well as tools (thin cylindrical objects such as toothbrush).

- The accuracy of LEAP in tests can be guaranteed in the order of mm in static detection of the hands. (manufacturer advertises higher accuracy [9])
- The case is water resistant although the manufacturer does not provide an Ingress Protection Code (IPC). Moisture in the case is detected by the device and external IR light compensated, so, these parameters can be supervised to obtain optimal operation conditions.
- It can detect specific movements such as grasping of the toothbrush.

On the contrary, there are some limitations in the use of LEAP:

- Small workspace. It must be located horizontally near the washbasin for initial recognition.
- Toothbrush must not be reflective because the IR won't work properly.
- Toothbrush tracking is lost when it is near/inside the mouth. In that case, the coordinates of the toothbrush can be obtained from the coordinates of the hand (estimation method) or by using a visible tool shaped marker attached in the toothbrush.

Finally, regarding some technical specifications the following points must be highlighted:

- Optimal distance between Leap and hands: [100, 250] mm.
- Ideal workspace is more detailed in Figure 11:

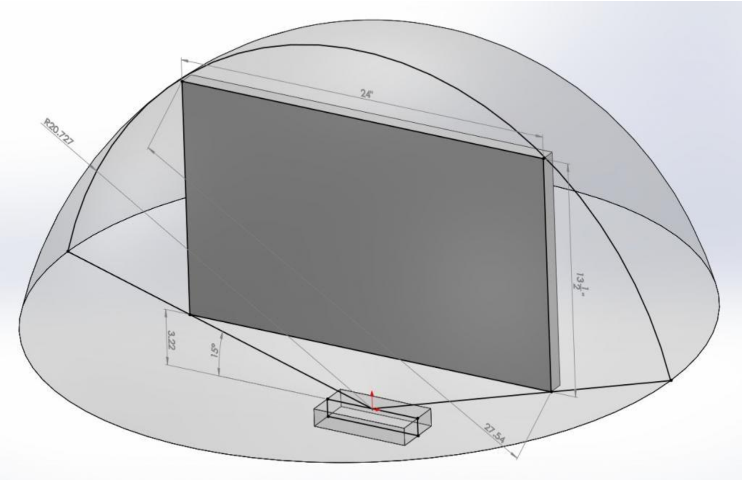


Figure 11. Leap ideal workspace.

3.1.3 Shimmer3 for wrist movement analysis

Shimmer3 (Figure 12) is a small and robust wireless sensor that can be attached to the patient's wrist in order to provide reliable data about movement and behavior while manipulating objects.

Shimmer enterprise has done its best to develop this wireless sensor which ensures quality in the delivery of the data and thanks to its smarter enclosure, it makes easy for the user to attach it on the corresponding area for wearable and remote sensing applications such as human health monitoring and ADL, which toothbrushing is framed in.

In this case, the area to be attached is the wrist of the patient to obtain specific data of the behavior which gives valuable information and support to estimate and classify the movements carried out while manipulating objects, especially, the toothbrush.

Figure 12. Shimmer3 wrist sensor. [10]

3.1.3.1 Hardware description

Shimmer3 sensor provides with an integrated 10 DoF IMU which the main components (Figure 13) to take advantage of are: an accelerometer, gyroscope, magnemoter (each with selectable range) and an on board computer to process 3D orientation data. Its dimensions are 51mm x 34mm x 14mm. [10]

Figure 13. On board components of Shimmer3 sensor. [10]

Accelerometer (Kionix KXRB5-2042) and Gyroscope (Invesense MPU9150)

The primary pair of sensors included in the sensor is an accelerometer and a gyroscope. These sensors measure acceleration and angular velocity. The data from these two sensors combined can be used to model the orientation, direction of movement and speed of the hand/wrist while toothbrushing. This is useful for looking for errors in the brushing action itself, and also recording how much time is spent brushing in various positions.

Magnetometer (STMicro LSM303DLHC)

A magnetometer is a sensor that can measure the surrounding magnetic field. From this, the angle between the sensor's axis and the earth's magnetic field can be derived. Therefore, unlike the accelerometer and gyroscope, which provide information about the direction and rotation of the brushing in relation to the axis of the movement, magnetometer measures the tilt of the hand, and consequently, of the tool grasped (toothbrush) in relation to the earth's axes. It is demonstrated that using data from an accelerometer and magnetometer, one can classify position of toothbrushing correctly in most cases.

A proof of concept preliminary study using Opal APMD movement monitor has been conducted, which incorporated accelerometer, gyroscope and magnetometer. Participants were instructed to brush one of the 12 different locations (top/bottom x inside/outside x right/left/middle) in each trial while data was recorded. Data from four participants were processed and probabilistic neural network was used as pattern classifier. As shown in Figure 14, performances of classifiers using magnetometer data alone, using accelerometer and magnetometer data and using all data were comparable. When classifiers did not use magnetometer, data performed worse than the other classifiers.

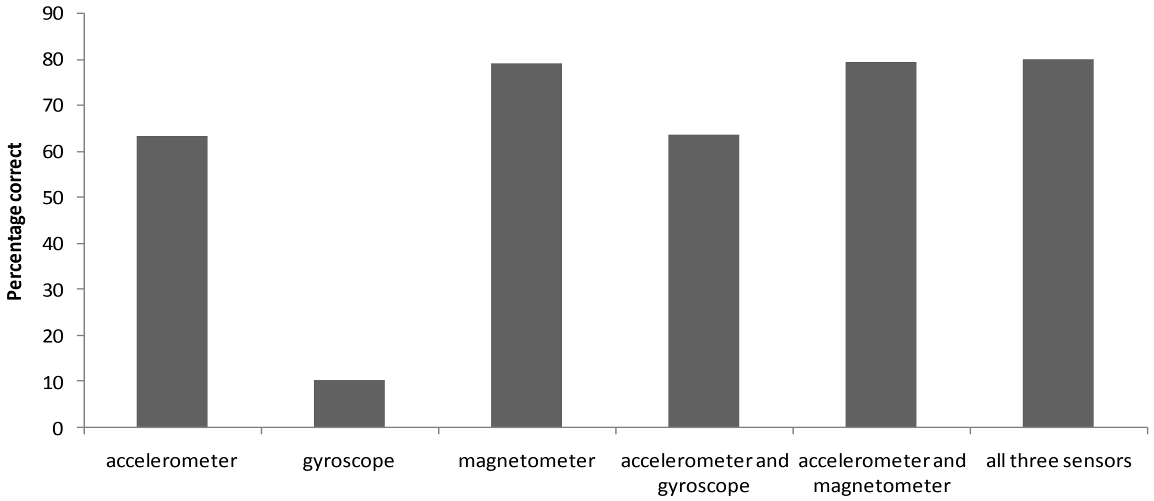


Figure 14. Percentage correct classification of testing data (N= 276) for toothbrushing location (12 locations) using different sensor(s) by probabilistic neural network.

3.1.3.2 Firmware and drivers description

Shimmer3 provides with a wide range of packages which are compatible with different programming languages such as Java/Android, C#, LabView or Matlab, so it ensures the flexibility in implementation.

3.1.3.3 Advantages, constraints and technical specifications

Regarding the main advantages this sensor offers in ergonomics, the following can be highlighted:

- Efficient straps that ensure full comfort and mobility to the patient.
- Low consumption of power and light weight.

- Possibility of attachment of a watch to the same strap of Shimmer3.
- Easy adaptation to different user requirements depending on the application to use.

Meanwhile, considering possible constraints when processing, although it offers high accuracy data which gives the opportunity of real time implementation, due to the rapid motion of the tooth brushing action some filtering of the data may be required in order to calculate accurate orientation.

The accelerometer and gyro data is prone to drift if used for dead reckoning the exact position of the hand + toothbrush, so the support from the Kinect^{TM} and Leap sensors remains useful for finding the position of the toothbrush relative to the user's face.

Attempts to measure orientation on the vertical axis is also prone to drift, support from the other sensors may be helpful for increasing accuracy.

Finally, Figure 15 presents relevant technical specifications:

TECHNICAL SPECIFICATIONS	
Microcontroller	**TI MSP430**
A/D Resolution	11 Channels of 12-bit A/D: 7 free for expansion
RAM	16KB
Flash	256KB
Frequency	24MHz
Wide Range Accel.	**STMicro LSM303DLHC**
Range	±2g, ±4g, ±8g, ±16g
Sensitivity	1000 LSB/g at +/-2g
Numeric Resolution	16-bit
Typical Operating Current	110 µA (Running Mag @ 7.5 Hz & Accel @ 50 Hz)
RMS Noise*	27.5 x 10^{-3} m/s^2
Low Noise Accel.	**Kionix KXRB5-2042**
Range	±2g
Sensitivity	600 ±18 mV/g
Typical Operating Current	500 µA
RMS Noise*	5.09 x 10^{-3} m/s^2
Digital Mag	**STMicro LSM303DLHC**
Range	±1.3; ±1.9; ±2.5; ±4.0; ±4.7; ±5.6; ±8.1 Ga
Sensitivity	1100 LSB/Ga at ±1.3
Numeric Resolution	16-bit
RMS Noise*	0.0081 normalised local flux
Gyro	**Invensense MPU9150**
Range	±250; ±500; ±1000; ±2000 dps
Sensitivity	131 LSB/dps at ±250
Numeric Resolution	16-bit
Typical Operating Current	3.5 mA
RMS Noise*	0.0481 dps
Pressure Sensor	**Bosch BMP180**
Range	300 - 1100 hPa
Numeric Resolution	16-bit
Typical Operating Current	1 µA at 1 Hz
RMS Noise (Standard mode)	0.4 m (from Datasheet)
*@ 100Hz Bandwidth	

Figure 15. Technical specifications of Shimmer3 sensor. [10]

3.1.4 Integrated coaster for Rinsing Cup and Jug

The rinsing cup and jug will utilise the existing CogWatch integrated coaster device, already used for prototype 1 and detailed in “D2.2.1 Report on devices I”. This sensor includes a three axis accelerometer for motion data and force sensitive resistors, which are used to measure changes of weight of the objects. This should be sufficient to detect the use of the jug and cup; this device has already been used in a similar fashion in the tea making task.

3.1.5 Toothpaste Dispenser

The toothpaste dispenser’s dispensing button will be instrumented with a switch to detect when it has been pressed.

3.1.6 Microphones for Audio Recognition

3.1.6.1 Motivation

Audio-based identification of the position of the toothbrush in the mouth is motivated by the hypothesis that changes in the shape of the mouth cavity, for example due to changes in the position of the tongue, during tooth-cleaning will alter the resonance properties of the vocal tract and hence the spectral shape of the resulting sound. This, coupled with the different positions of the tooth-brush head in the mouth, will result in audio cues to estimate location in the mouth where brushing is taking place.

A pilot study was conducted, described in “D3.2.2 Report on data analysis for action recognition II”, involving a single individual brushing his teeth. The results showed that if the mouth was divided into 8 regions (bottom-back-left, bottom-back-right, top-back-left, top-back-right, bottom-front-inside, bottom-front-outside, top-front-inside and top-front-outside) then the brushing position could be identified for these regions with an accuracy of 87% using a simple Gaussian Mixture Model (GMM) based classifier.

This experiment used a head-mounted microphone with the microphone element placed close to the corner of the mouth and was conducted in a quiet environment.

3.1.6.2 Hardware description

There are three possible choices of microphone (Figure 16).

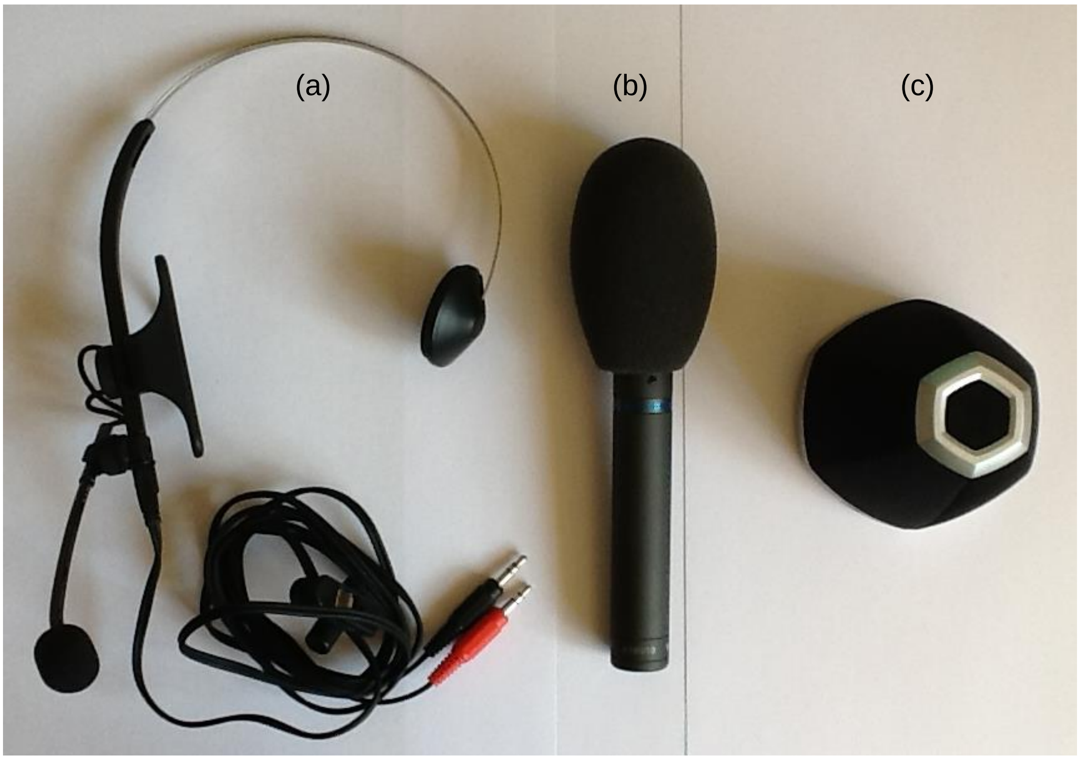


Figure 16. Possible microphones for use in tooth-brushing analysis. From left to right: (a) EmKay head-mounted microphone, (b) Audio-technica omnidirectional condenser microphone, and (c) DEV-AUDIO 'micro-cone' microphone array with seven microphone elements.

A head-mounted microphone (Figure 16(a)) has the advantages that the distance from the microphone to the mouth is small and constant. Because the microphone element is close to the mouth the system gain can be low and the effect of noise is also minimized. However, locating the microphone element close to the mouth may interfere with brushing and the microphone may get knocked. It is also not clear that a head mounted microphone would be acceptable to patients and clinicians.

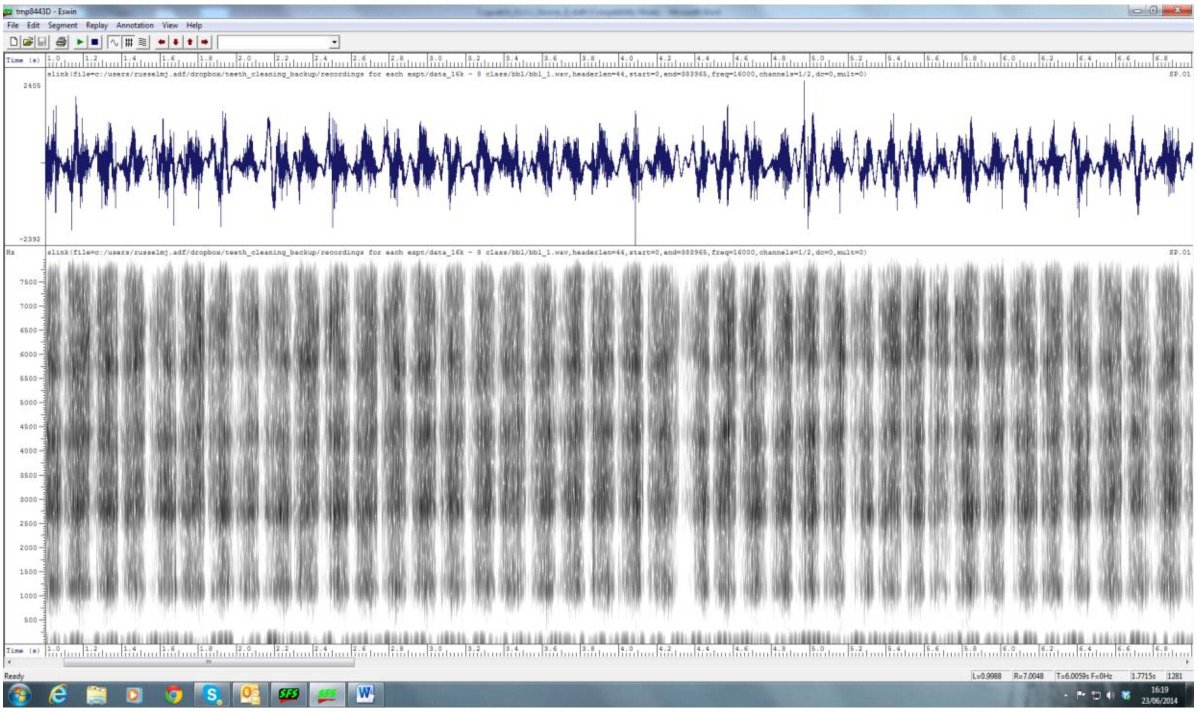
An omnidirectional microphone such as that shown in Figure 16(b) can be unobtrusive, and although a large high-quality microphone might be used for development a miniature microphone (similar to the type used in mobile phones) could be used in the actual system, perhaps embedded in the mirror. The disadvantages are that the distance from the microphone to the mouth will vary, and the system will be sensitive to noise.

The final solution is a microphone array, such as the DEV-AUDIO micro-cone system shown in Figure 16(c). This system allows beam-forming to be used to focus on the most important sound source and therefore offers both resistance to noise and convenience. In a real application an array of microphones could potentially be embedded in the mirror or placed round its edge, depending on the optimal configuration.

3.1.6.3 Signal Processing

Figure 17 shows a speech spectrogram (a time-frequency plot) of a 5s recording of tooth-cleaning for the position BBL (back bottom left) from the pilot study. Time is shown on the horizontal axis and frequency on the vertical axis. The power at a particular time and frequency are displayed on a grey scale, with power proportional to darkness. The sampling frequency was 16kHz and the maximum frequency shown is 8kHz.

The horizontal bands correspond to resonances of the mouth cavity when it is configured for brushing in the BBL position. The vertical striping corresponds to the periodic structure of brushing and indicates a brushing frequency of approximately 4Hz.



```
slink(file=c:/users/russelmj.adf/dropbox/teeth_cleaning_backup/recordings for each expt/data_16k - 8 class/bbl/bbl_1.wav,headerlen=44,start=0,end=383965,freq=16000,channels=1/2,dc=0,mult=0)
```

```
L=0.9988   R=7.0048   T=6.0059s F=0Hz   1.7715s   1281
```

Figure 17. Spectrogram of a 5s section of a recording of tooth-brushing for position BBL (back-bottom-left).

The mouth cavity resonances, indicated by the dark horizontal bars occur at approximately 1200Hz, 2800Hz, 3300Hz (possibly), 4250Hz and 5550Hz.

We believe that this structure is being exploited in our automatic tooth-brush head position estimation pilot experiments. The experiments are described in more detail in "D3.2.2 Report on data analysis for action recognition II".

3.1.6.4 Feature Extraction

In order to be used for action recognition, the audio signal needs to be converted into a sequence of feature vectors. Ideally these vectors should accentuate aspects of the signal that are important for recognition and ignore noise. The procedure used in the pilot study is identical to that used in a typical speech recognition system (for example, see [11]). In the latter, the details of the signal processing are inspired by knowledge of the human auditory and speech production systems. The motivation is that human speech has evolved in the context of human hearing. Of course, the same motivation does not apply to general audio action recognition and therefore the proposed approach should be regarded as a baseline.

The acoustic signal is transformed into a sequence of Mel Frequency Cepstral Coefficients (MFCCs). The details of this process are as follows:

1. The acoustic signal is partitioned into segments, each of length approximately 20-30ms and with adjacent segments offset by approximately 10-20ms. Suppose that the number of samples in a window is $N = 2^M$.
2. A Hamming window is applied to the segment, followed by a Discrete Fourier Transform (DFT). The DFT coefficients are replaced by their module, and coefficients $\frac{N}{2}+1$ to N are discarded (without loss of information because of the conjugate symmetry of the DFT).
3. The DFT spectrum is mapped from a linear to a non-linear, perceptually motivated frequency scale (called the mel scale) and the frequency axis is quantized into critical bands. For a sample rate of 16kHz the number of bands is 27. This is normally achieved using a set of triangular mel-scale filters.
4. A cosine transform is applied to create the DFT cepstrum, and typically only the first 13 cepstral coefficients are retained.
5. Finally, each of these 13 static cepstral coefficients is supplemented with its velocity and acceleration, resulting in a 39 dimensional vector.

In Prototype 2 this feature extraction stage will be implemented in C#.

3.1.6.5 Constraints

Although the sample rate of the original audio data is 16kHz, sample the MFCC vectors at between 50 and 100Hz should suffice. The pilot experiments used a 100Hz sample rate.

The pilot experiment on audio-based toothbrush location used recordings made in a quiet environment. The extent to which the results will be compromised by noise is as yet unknown. However, a combination of the use of a head mounted microphone or a microphone array, plus existing techniques for noise robustness that have been developed for automatic speech recognition, may provide some degree of noise tolerance.

3.1.7 Non Invasive Blood Pressure (NIBP) Module

The contribution of the NIBP module in this second prototype is mainly focused on measuring vital signs such as blood pressure and heart rate in order to monitor and control the patient before and after the corresponding rehabilitation session since some patients may suffer from stress when carrying it out.

These measurements can provide useful information about the level of stress and ability of the patients to concentrate during the task.

3.1.7.1 Hardware description

The improvements in the design of the NIBP module have been focused on two major areas:

1) Ergonomy.
2) Aesthetics.

3) Power Supply.

The module has been redesigned by adopting a smaller size based on shape (Figure 18) as compared to previous versions and the appearance has been also improved to be more attractive.

Figure 18. New design of NIBP module.

As shown in Figure 18, the module can be provided with three types of front connections. In case of micro USB connector, it is directly soldered on the PCB.

The main components (Figure 19) that compose the module are:

- Printed label.
- Upper and Lower Case.
- Frontal Label.
- Front connectors.
- Printed Circuit Board (PCB).
- Fixation Means.

Figure 19. Components of NIBP module.

Regarding battery charge, the device now incorporates the mobile concept in which a battery is not easily removable, but instead it is essentially fixed as one more device component. Battery charge is obtained by means of a similar technique as mobile devices, as well as similar charging principles.

NIBP module has an isolated type of sensor that minimizes security risks to the patient due to current leakage. In addition, there are different versions depending of the consumer's requisites:

a)

b)

Figure 20. Different design styles for NIBP module: a) White plastic; b) Metallic plastic.

Finally, the following figure shows a general representation of the assembly of the module and dimensions:

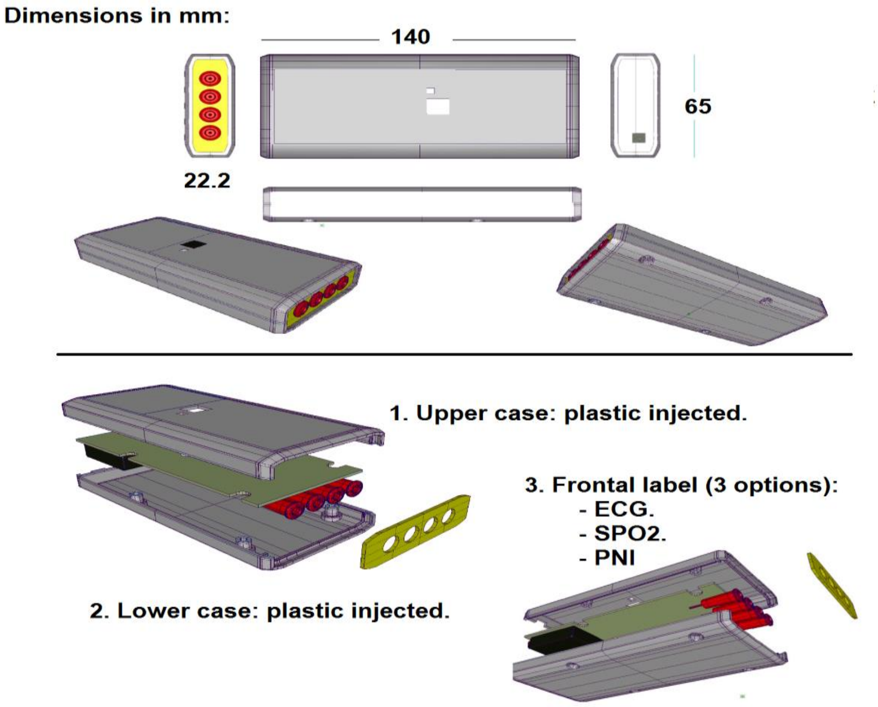


Figure 21. Assembly and dimensions of NIBP module.

3.1.7.2 Constraints

The main constraint has been to make it extremely simple to use. High ergonomy has been achieved by simply pressing the on/off switch. Then measurement is carried out following the instructions from the console, through remote commands to the device.

3.2 Feedback Devices for Toothbrushing

3.2.1 Metawatch

The smartwatch provides the same functionality as for the first prototype. It vibrates in case of error defined by the corresponding error table. For future versions, relevant data from the accelerometer may be obtained in order to have redundant information about the wrist movements of the patient for the action recognition. Some improvements have been done regarding the connection with the VTE and occasional delays in communication port and problems with the pairing port required have been solved.

3.2.1.1 Hardware description

The appearance of the watch has been updated, with different designs more attractive but the main hardware components and architecture are still the same as explained in "D2.2.1 Report on devices I". [1]

3.2.1.2 Operating system, firmware and programming design

As indicated in the previous section, the main software architecture of the watch is still the same. Just few updates have come out such as the firmware version, appearance of the screen interface, new interactions with the user or better and more stable communication with the accelerometer.

3.2.1.3 Constrains

The only constraint to be mentioned for the watch is the need for a degree of water resistance given the new task to be executed in this second prototype is toothbrushing. This is met by the watch specification as waterproof to 3 atm of pressure.

3.2.2 VTE Monitor

VTE Monitor, as described in previous deliverables (see [1]), has been thought to be the main device for interaction with patients. It is necessary in order to provide her/him a relevant feedback during the rehabilitation session. Moreover, the device which supports the VTE Monitor is also the central unit of the CogWatch System and holds all the connection with peripheral devices at the time it controls all the process involving the CogWatch Interface.

3.2.2.1 Hardware description

For this second prototype two different devices are envisaged given the necessity of testing the prototype in both experimental lab and home environments.

For the experimental lab setup, see section 4.2, the same All-in-One Computer as used in previous prototypes has been proposed. So that, hardware specifications are still the same as explained in "D2.2.1 Report on devices I". [1]

To allow use throughout the home environment of the patient, the CogWatch System is to be migrated to a mobile platform. To guarantee compatibility with all the software requirements, a range of different Windows® based tablets have been selected as suitable.

In the following table, Table 1, some of this possible candidates are shown:

Table 1. Candidate tablets for CogWatch P2 System.

List of available tablets which supports CogWatch System
Asus Transformer Book T100TA
Microsoft Surface Pro 2 & 3
Panasonic Toughpad FZ-G1
Dell Venue 11 Pro
Razer Edge Pro
Lenovo IdeaPad Yoga 2 Pro
Sony VAIO Tap 11

Next figure shows Microsoft Surface Pro 2, which is the model selected for implementing the home setup for prototype 2. All its hardware specifications can be found in [12].

Figure 22. Microsoft Surface Tablet. [12]

3.2.2.2 Software description

VTE device must support minimum requirements for the following software:

Table 2. Software requirements for VTE.

Element	Description
Operating system	Microsoft Windows® 7 or 8 64-bit.
Network server	Ethernet, Fast Ethernet, Gigabit Ethernet, IEEE 802.11a/b/g/n. Wireless 10/100/1000 Mbps.
Memory	At least 2 GB.
Data base management	MySQL.
Task Model	Python 2.7.
Microsoft Kinect ™	KinectSDK-v1.8

Leap Motion	LEAP SDK-2.0.3+17004

3.2.2.3 Constraints

The main constraints regarding the visualization of cues are related to the ease of use of the system and accessibility issues.

In support of ease-of-use, the CogWatch System has been designed to be scheduled supervised. The system will inform the user of the start of a programmed session and start up the interface which can also be easily activated at any time from a wake up system based on their monitoring devices.

Secondly, accessibility is obtained by the inner tools from the operating system. Also, some software has been developed to provide:

- Multiple languages.
- Custom Appearance.
- Verbal and text cues with adjustable font size.
- Vibration and sound alarms to advice the user.

Additionally, tablet devices could experience some problems regarding the required number of USB ports or the necessity of a wired internet connexion for improving security. To that end there are additional docks in the market (Figure 23) which allows to provide a ready-to-use platform with all devices plugged in the dock as a main station where the user can easily connect her/his tablet and move freely around the different scenarios proposed by CogWatch.

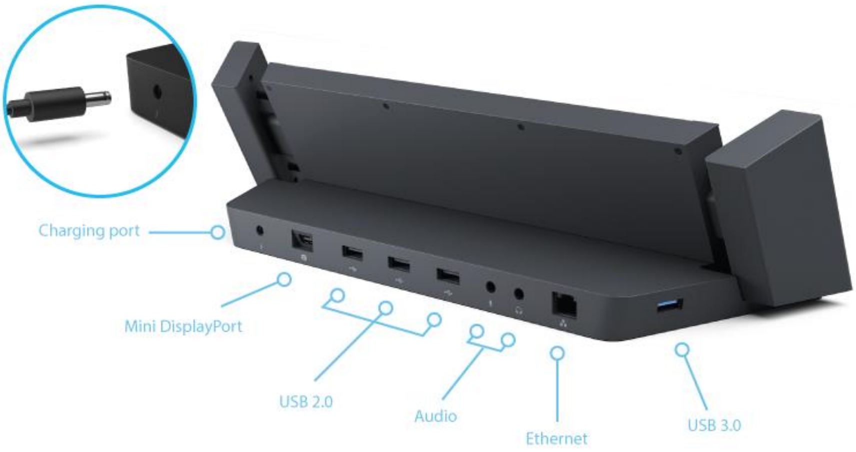


Figure 23. Dock station for Microsoft tablets.[13]

4. SET UP FOR SECOND PROTOTYPE

In this section all the technology and the concept idea for the prototype will be summarized.

4.1 Devices table

The following table summarizes the monitoring devices to be finally implemented in the prototype with relevant information related to the information provided to have a more clear idea of each contribution. Note that only those devices whose data in real time is obtained are considered.

Table 3. Devices implemented in Prototype 2.

Device	Data provided	In support of
Kinect™	Coordinates of mouth points	Detection of toothbrush approach to mouth
Leap	Coordinates from tip of toothbrush and hands/fingers	Continuous tracking of hands while approaching objects and toothbrush when grasped
Shimmer3	Accelerations, orientations and angular velocities of wrist	Detection of specific actions, especially, toothbrushing
Coasters	Accelerations and forces in sensorized objects	Detection of motion. Recognition of actions
Microphones	Featured sound in mouth and toothbrush	Behaviour during toothbrushing

So, during the execution of the whole task, the objective is to monitor and track the behavior of the patient's hands and objects thanks to Leap and strategically located coasters. This makes possible the recognition of the sub-actions focused on manipulating the objects involved in the task over the table.

Once the toothbrush is in the air, coated with toothpaste, by analyzing the coordinates of the tip of the toothbrush given by Leap and the position of mouth given by Kinect™, the approach of the toothbrush to the mouth is detected and recognized.

Once the patient is just executing the toothbrushing action, more deep information can be achieved by analyzing the sounds and noise from the microphones.

Finally, with the implementation of the wrist strap from Shimmer3, a continuous monitoring of the movements and behavior of the patient's hand is obtained to support the rest of the sensors, especially, when grasping the toothbrush and while toothbrushing.

4.2 Experimental Layout for Toothbrushing

For the experimental set-up the following layout is proposed:

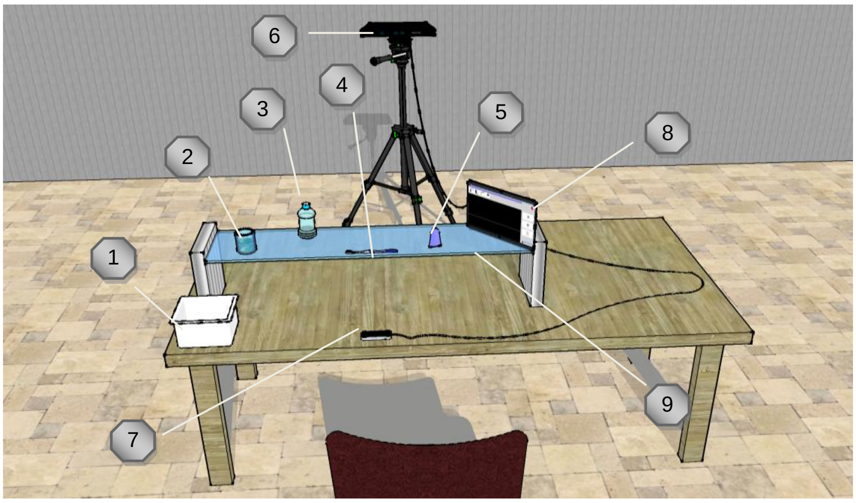


Figure 24. Experimental layout.

where:

Table 4. List of devices.

Element #	Description
1	Spittoon
2	Glass of water
3	Jar/Bottle of water
4	Toothbrush
5	Toothpaste dispenser
6	Kinect™
7	Leap Motion
8	Microsoft® Surface Pro 2
9	Acrylic table

In this case, as this set up is focused on researching and testing in lab, the VTE monitor can be used in the current All-In-One computer since the table solution is more useful for the bathroom scenario.

The acrylic surface is used in order to make easier the tracking of the hand while manipulating the objects. So, the objects must be inside the field of view of Leap. However, this could be changed depending on the final user requirements. If there are some difficulties for patients to reach the objects on the acrylic surface, this will be removed and the objects will be placed directly on the table. In this case, Leap will be more focused on giving support to detect approach to mouth with the toothbrush than delivering tracking while manipulation.

4.3 Home-based Layout for Toothbrushing

In this case, a graphical layout with a more realistic home-based set up is shown:

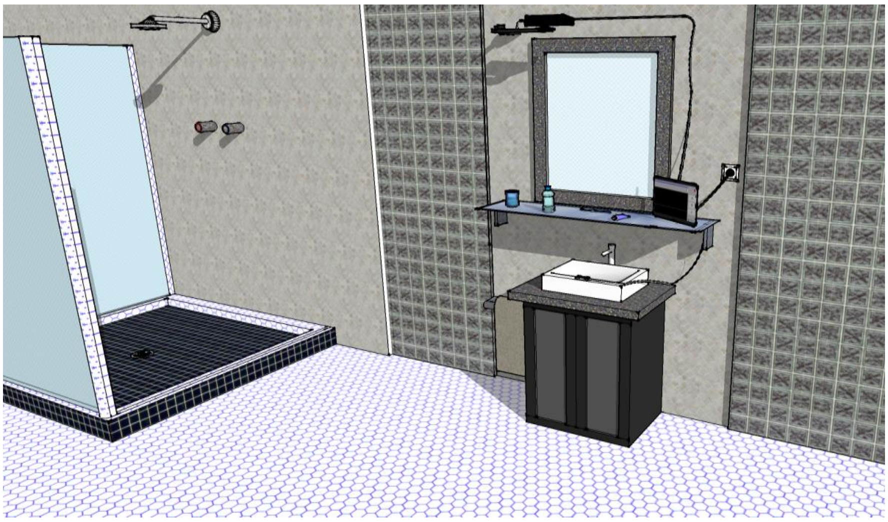


Figure 25. Home-based layout.

For this set-up all the devices remain the same as in the previous one. In order to take advantages of the detection of mirror reflection, Kinect^{TM} could be placed facing the mirror in order to change as less as possible the domestic environment, however this is something to fix according with patient requirements and physical constraints of the real bathroom used.

Initially, the evaluation of the system will be carried out by using the experimental set up in order to move to the home-based layout for more advanced use.

CONCLUSIONS

This deliverable has described the features of the devices being considered in the second prototype of CogWatch in terms of components, HW/SW, implementation and constraints when installation. Detailed technical descriptions have been provided for those components not used in the first prototype for their novelty in the new platform.

The document has settled down the new architecture proposed for the prototype in terms of the components which will provide data of the execution of the task and feedback to the patient. In this case, there are new sensors to be used such as Leap Motion, to obtain information of hand/toothbrush movements or Shimmer3 and microphones to better estimate the motion of the toothbrush outside and inside mouth. Kinect^{TM}, which is also used in the first prototype, now is used to obtain data from the mouth position and head poses.

Additionally, an easy-to-use small module is considered to measure some vital signs such as blood pressure and heart rate of the patient to monitor his/her level of stress before and after the rehabilitation session.

With the architecture and components proposed, it is expected that the system will work properly and effectively in order to provide rehabilitation in the new task considered and focused on tooth brushing.

The components described in the document comprise the subsystem called CogWatch client, which corresponds to the patient side and installed at his/her house. Considering the subsystem focused on the server, it is described in the corresponding deliverable *“D2.3.2 Report on networks II”* and complements to this current deliverable *“D2.2.2 Report on devices II”*.

REFERENCES

[1].CogWatch deliverable D2.2.1 Report on devices I.

[2].«Microsoft - Home». [Online]. Available: http://i.msdn.microsoft.com/dynimg/IC568988.png. [Accessed: 17-jun-2014].

[3].«Microsoft - Home». [Online]. Available: http://i.msdn.microsoft.com/dynimg/IC584339.png. [Accessed: 17-jun-2014].

[4].Mouth image. [Online]. Available: http://www.123rf.com/photo_7251819_mouth-and-teeth-before-and-after-whitening.html. [Accessed: 17-jun-2014].

[5].«Microsoft - Home». [Online]. Available: http://i.msdn.microsoft.com/dynimg/IC584329.png. [Accessed: 17-jun-2014].

[6].«Microsoft - Home». [Online]. Available: http://i.msdn.microsoft.com/dynimg/IC584331.png. [Accessed: 17-jun-2014].

[7].«Introducing the Skeletal Tracking Model — Leap Motion C# and Unity SDK v2.0 Beta documentation». [Online]. Available: https://developer.leapmotion.com/documentation/skeletal/csharp/devguide/Intro_Skeleton_API.html. [Accessed: 21-jun-2014].

[8].«Leap Motion Teardown - Learn.SFE». [Online]. Available: https://learn.sparkfun.com/tutorials/leap-motion-teardown. [Accessed: 21-jun-2014].

[9].«Leap Motion | 3D Motion and Gesture Control for PC & Mac». [Online]. Available: https://www.leapmotion.com/product. [Accessed: 21-jun-2014].

[10]. «Shimmer - Home». [Online]. Available: http://www.shimmersensing.com/shop/shimmer3. [Accessed: 18-jul-2014].

[11]. S. J. Young, G. Evermann, M. J. F. Gales, T. Hain, D. Kershaw, G. Moore, J. Odell, D. Ollason, D. Povey, V. Valtchev, P. C. Woodland, “The HTK Book, Version 3.4”, University of Cambridge, 2006.

[12]. «Microsoft Surface Pro 2 specifications | Surface Pro 2 specs». [Online]. Available: http://www.microsoft.com/surface/en-us/support/surface-pro-2-specs. [Accessed: 21-jun-2014].

[13]. «Docking Station for Microsoft Surface Pro and Surface Pro 2». [Online]. Available: http://www.microsoft.com/surface/en-us/support/hardware-and-drivers/docking-station. [Accessed: 21-jun-2014].