The i-Treasures Intangible Cultural Heritage dataset

Grammalidis N., Dimitropoulos K., Tsalakanidou F., Kitsikidis A., Roussel P., Denby B., Chawah P., Buchman L., Dupont S., Laraba S., Picart B., Tits M., Tilmanne J., Hadjidimitriou S., Hadjileontiadis L., Charisis V., Volioti C., Stergiaki A., Manitsaris A. Bouzos O. and Manitsaris S.

ngramm@iti.gr

ABSTRACT

In this paper, we introduce the i-Treasures Intangible Cultural Heritage (ICH) dataset, a freely available collection of multimodal data captured from different forms of rare ICH. More specifically, the dataset contains video, audio, depth, motion capture data and other modalities, such as EEG or ultrasound data. It also includes (manual) annotations of data, while in some cases additional features and metadata are provided, extracted using algorithms and modules developed within the i-Treasures project. We describe the creation process (sensors, capture setups and modules used), the dataset content and the associated annotations. An attractive feature of this ICH Database is that it's the first of its kind, providing annotated multimodal data for a wide range of rare ICH types. Finally, some conclusions are drawn and the future development of the dataset is discussed.

Author Keywords

Intangible Cultural Heritage; multimodal capture;

ACM Classification Keywords

I.4.m IMAGE PROCESSING AND COMPUTER VISION - Miscellaneous;H.2.4 Systems - Multimedia databases

INTRODUCTION

Cultural heritage is not limited to tangible objects, such as monuments, paintings, sculpture and artifacts. It also includes Intangible Cultural Heritage (ICH), which refers to practices, presentations, expressions, knowledge, skills and associated tangible objects (e.g. places, tools, clothes, etc.) . This intangible heritage passes from generation to generation and gives people a sense of identity and continuity; it is the result of the continuous interaction of communities and groups with their nature and history and it promotes respect for cultural diversity and human creativity. Examples of such expressions include music, dance, singing, theatre, human skills, and craftsmanship. These manifestations of human intelligence and creativeness constitute our Intangible Cultural Heritage (ICH). The importance of ICH is not limited to cultural manifestation; it rather lies in the wealth of knowledge transmitted through it from one generation to the next [1].

In recent years, ICH has received international recognition, and its safeguarding has become one of the priorities of international cooperation mainly thanks to UNESCO's initiatives [1]. Museums can play the role of a mediator to fill the gaps between generations and share community knowledge; however, there are significant limitations in their role mainly due to the technologies employed, which usually aim at simple digitisation of cultural content. ICH, however, is more efficiently preserved 'with' the people or community by protecting the processes that allow traditions and shared knowledge.

Indeed, ICH may also act as a powerful societal and economic driver for sustainable development, at the local, national, European and global level. It offers multiple benefits giving communities and countries an opportunity to identify common elements in their history and culture, as well as local customs, particularities that may be used to achieve "local branding". It can thus provide local economies with stable jobs and revenues through cost-effective investments.

However, since the UNESCO Convention for the Safeguarding of the Intangible Cultural Heritage in 2003, very diverse approaches to ICH inventorying have been adopted. UNESCO only provides few guidelines on ICH inventorying for States Parties as the Convention itself does not prescribe a rigid approach. Although there is a capacity-building workshop run by UNESCO on community-based inventorying, it is not yet available on open access and most of the UNESCO capacity building workshops are held in states outside of Europe.

To this end, modern ICT technologies can be employed to improve the capture, collection, presentation and transmission of ICH, raise public awareness, provide seamless and universal access to cultural resources, support services for research and education and bring hidden intangible treasures to light. In Europe in particular, a growing number of ICH inventories are digital and use digital platforms for management and access.

“Europeana Sounds”(http://www.europeanasounds.eu/), is an EU project focusing on increasing the audio content, representing the Europe’s sound and music heritage, in Europeana portal. Europeana Sounds consortium, composed of 24 organisations coming from 12 EU states, worked on the ethical and legal issues to facilitate the use and dissemination of old and new field recordings.

The digital platform for Intangible Cultural Heritage in Flanders (http://ww.immaterieelerfgoed.be) is a network of heritage workers, heritage organizations, communities and individuals in Flanders (Belgium). They offer a tool for inventorying ICH, but put their priority on the safeguarding of ICH and the exchange of know-how, expertise on ICH, safeguarding measures and best practices between professionals and heritage communities.

The project "E.CH.I. / Etnografie italo-svizzere" (http://www.echi-interreg.eu/) and its continuation E.CH.I.2, funded by the European Union under the Interreg IV Italy-Switzerland border cooperation, was established to document, preserve and make accessible the intangible cultural heritage of the areas of Alps and cross-borders territories, in Italy and Switzerland, through the creation of a Register of ICH of trans-border Italo-Suisse heritage.

A framework for the inventorying, cataloging, searching and browsing of multimedia digital objects related to ICH, was developed to design and implement the AESS archive (Archive of Ethnography and Social History of Lombardy Region, http://www.aess.itc.cnr.it/ricerca). The data concern mainly popular traditions handed down generation by generation, such as traditional fairs, popular songs, and customs.

However, there is still a need for multimodal ICH datasets and digital platforms to manage a wide variety of multimedia digital content, increase public access to inventories and participation in the inventorying process, protect community rights and engagement with their heritage, and link effectively to existing digital databases of other cultural material.

During the course of i-Treasures project, several recordings have been organised by different partners and many data sets have been created. These have been used within the project for data analysis algorithm development and evaluation. In some cases, special hardware and software was used. For instance, for the recordings of rare songs, a prototype light-weight hyper-helmets embedding several sensors for vocal tract capture (i.e., video camera, ultrasound probe, microphone, EGG sensor, piezoelectric accelerometer, and respiratory belt) was designed and produced. Furthermore, since configuring separated sensors and recording their outputs may be a complicated issue if managed individually, a common module was specifically designed, named i-THRec (i-Treasures Helmet Recording software). Due to the use of multiple sensors and specific acquisition needs, recordings were made in a recording studio, instead of more natural environment (e.g. a church in the case of Byzantine). However this also has benefits, since the datasets produced have better quality, so they are more suitable to be used for educational purposes. Also, a novel intangible musical instrument has been developed, that maps in real time information from the performer (both from his/her natural movements and gestures and from his/her brain activity) to music/voice segments. Part of the recorded ICH performances (raw data and annotation data), which are briefly described in the following subsections, have now been made publicly available to other researchers via the project website (http://www.i-treasures.eu – registration to the site is required).

ICH CAPTURE AND DATASET CREATION

Tsamiko dance

Tsamiko is a popular traditional folk dance of Greece, done to music of ¾ meter. It is a masculine (mostly) circular dance with more smoothly steps danced by women. It is danced in an open circle where the first dancer performs variations while the others follow the basic steps. Tsamiko is danced in various areas of Greece such as: Peloponnese, Central Greece, Thessaly, W. Macedonia, with variations in kinesiological structure (10, 12, 8, 16 steps). Recordings of Tsamiko dance performances were made using a multi-Kinect setup (Figure 1).

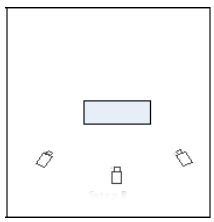


Figure 1. Setup with three Kinect sensors used

The recordings were organised in cooperation with the Department of Physical Education and Sport Science of the Aristotle University of Thessaloniki and Prof. Stella Douka. Time-stamped colour maps, depth maps and skeletal data streams were recorded concurrently by 3-4 Kinect sensors. Both experts and students have been recorded, each person separately, for the duration of a single dance performance (approx. 4 minutes). These sequences (Figure 2) have also been used for the visualisation of the expert avatar movements in the Tsamiko game-like application.

Figure 2. Male expert dancing Tsamiko

Căluş dance

Romanian Căluş dance originated as a healing and fertility ritual performed by groups of an odd number of men, bound together by an oath. By the beginning of the 20th century its ritual form survived mainly in southern Romania and among Romanian minorities in northern Bulgaria, although remnants of this custom could be found in much of the rest of Romania, and throughout the Balkans. The Căluş tradition, especially through the art of the dance, caught the attention of the pilgrims, historians, researchers and has always raised the general admiration. time, the beauty and complexity of this custom. At the end of 2005, the Căluş custom was included by the UNESCO in the list of the immaterial masterpieces of humanity.

A recording of the Calus dance expert Mr Florian Teodorescu was organised by CERTH in July 2014 in the Department of Physical Education and Sport Science of the Aristotle University of Thessaloniki. The same multi-Kinect setup (Figure 1) was used, as in the Tsamiko case.

The expert was wearing a traditional costume and he was holding an ornate stick (Figure 3). The stick plays a significant role in the choreography of the dance. He performed four different styles of Calus, which people dance in different geographical regions of Romania.

The basic steps and motifs of Calus were also separately recorded and these videos will be used in the corresponding educational course. Time-stamped colour maps, depth maps and skeletal data streams were concurrently captured by the multi-Kinect setup. The data sequences were manually annotated with the expert's assistance.

Figure 3. Calus expert dancing

Walloon dance data collection

The Walloon traditional dances are peasant dances originated from the 18th, 19th and early 20th centuries and practiced in the Walloon region of Belgium (South). These dances were mostly danced in popular balls in the villages but almost disappeared at the end of the 19th century. Few people interested in preserving and perpetuating this intangible cultural heritage at their own initiative by interviewing older people who used to perform the dance.

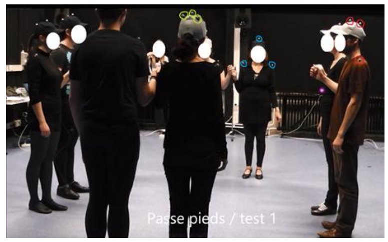
In order to contribute to the preservation and transmission of theses dances, two data collections were recorded during the project at the Numediart Institute from the University of Mons, Belgium, using several motion capture systems. The first data collection was recorded using the first version of the Microsoft Kinect [2] and a high precision motion capture system called

Qualisys [3]. A second data collection was recorded later using the second version of the Microsoft Kinect and the Qualisys also.

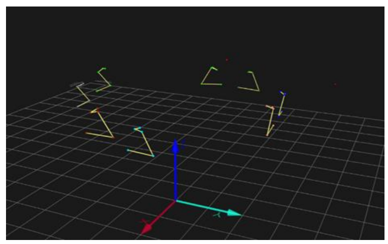
1st data collection:

Two different setups have been considered for the first Walloon dance database collection:

- Recording of one Walloon dance choreography (passepied) performed by eight dancers (Figure 4).
- Recording of one Walloon dancer performing the basic steps of the different styles of Walloon dance (Figure 5).



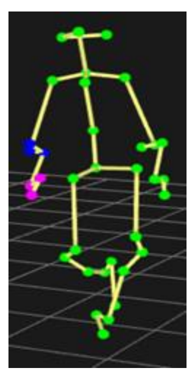
(a)



(b)

Figure 4. Recording of Walloon dance (Passepied) performed by eight dancers: video captured by RGB camera (top), and output of motion capture (head and shoulder positions) (bottom).

The database contains point cloud data (".oni' file format) captured with six Kinect V1 sensors using OpenNi [4] and mocap data captured with the Qualisys optical motion capture system (".c3d", ".tsv" and ".mat" file format). The Qualisys captured locations of markers (5 markers in the case of multiple dancers, placed on the head and the shoulders of each dancer and 68 markers in the case of a single dancer, at a framerate of 179fps).

(a) (b)

Figure 5. Expert of the Walloon dance performing basic steps. (a) RGB camera, (b) output of the motion capture system (Skeleton is reconstructed from 68 markers locations).

2nd data collection

In the second data collection, two dancers were recorded (an expert and a student) performing basic steps in addition to common errors that students make. The database contains data captured with two different motion capture systems:

- Qualisys motion capture system (68 markers) with a framerate of 177fps. Data is extracted in “c3d” file format, “txt” (1 line = 1 frame) and “xml” format defined by i-Treasures.
- The second version of the Kinect. Skeletal data (positions and rotations of 25 joints) is extracted in “txt” and “xml” file formats.

The data collection contains also the manual annotations of these sequences and the music pieces played during the performances.

Contemporary Dance

Contemporary dance is a dance genre inherited from occidental modern dance choreographs, which appeared during the mid-twentieth century. It is today one of the most dominant genres, though it does not have a standard definition. Contemporary dance style is partly inspired from many dance genres such as modern, classical and jazz dance.

Figure 6. Contemporary dance recording. The dancer wears the 68 retroreflective markers for optical motion capture.

A recording of this dance style was made at the Numediart Institute from the University of Mons. Six professional contemporary dancers (3 males and 3 females), from the dance school P.A.R.T.S[1], have been recorded. For each performer, movements of the whole body were recorded thanks to the Qualisys optical motion captured system [2]. A total of 68 retroreflective markers were placed on the whole body (see Figure 6). The trajectories of these markers were captured at a framerate of 175 FPS, by 11 infrared cameras covering a capture scene of 4m×4m (see Figure 7).

Each dancer individually performed one common choreography, as well as improvisation on five emotionally charged musical pieces, picked in a tagged musical database (including tags on arousal and valence scales) [13]. These five musical pieces were chosen to cover the arousal-valence emotion diagram (see Figure 8).

Additionally, each dancer executed three technical dance gestures: a) Arabesque b) Port de bras c) Triplet.

[1] P.A.R.T.S : www.parts.be

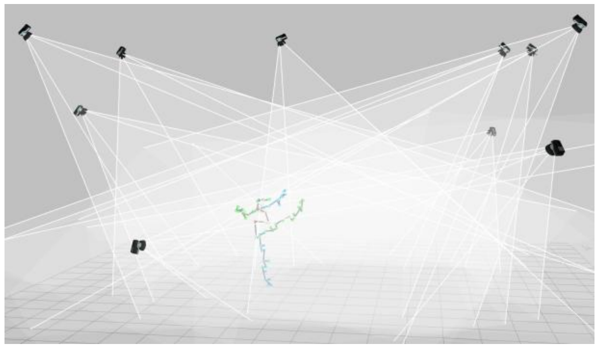


Figure 7. Contemporary dance recording. The capture scene is covered by 11 infrared cameras.

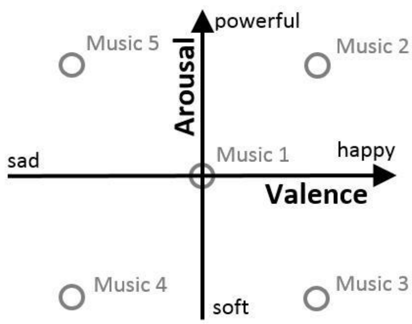


Figure 8. Arousal-Valence emotion diagram. Five musical pieces were chosen at the extrema of the diagram.

Human beat-box data collection

HBB professional performers and teachers, Davox (in October 2013) and Matthieu (in May 2014) were recorded in CNRS labs (Figure 9). Several songs including music instruments' imitation (percussion, wind instruments, string instruments), modern beats, scratch etc. were performed and recorded using high-quality audio equipment. Vocal tract data was also recorded using the hyper-helmet and the i-THRec software [13] for several of these performances. The data set totals more than 9,000 musical events, where each musical event is either a drum sound event or a pitched note covering various imitations. This dataset was used for the development and testing of sound processing algorithms. A more detailed review of the corpus can be found in another paper [12]. Three of Matthieu's performances are concerned by this public dataset. They correspond to two basic rhythm interpretations and one freestyle interpretation. Textgrid and .xml phonetic annotation files are also part of the public dataset.

Figure 9. Recording of Human Beat Box artist Mathieu. The artist wears the hyper-helmet.

Byzantine music

Byzantine music is the music that was cultivated and developed in Byzantium. Today is the music which is chanted in the Greek's Orthodox Church holy Ceremonies. This kind of music has a different sonality (microtonality) which is related with the Greek traditional music. The Byzantine music is a Vocal music without instruments. The purpose of Byzantine music is to praise the God with the help of the orthodox Holy textures to understand the meaning of salvation with ultimate goal the "Theosis" (to reach the god). On February 2016, a recording session organized by UOM in cooperation with CERTH took place in a sound recording studio at UOM's premises (Figure 10). Dr. Dimitris Manousis, a Byzantine music chanter and expert was recorded chanting Hymns, scales and apechema (vocal tunable) of the First mode in rhythmic and melodic parallage (solfège /solmization), in rhythmic Lyrics (rhythmic diction) and melos (melody with lyrics). In addition to sound, vocal tract and facial data were collected using the ultrasound probe attached to the hyper-helmet and a Kinect sensor respectively. Colour maps and depth maps of the face of the chanter were captured by a Kinect sensor.

(a) (b)

Figure 10. Byzantine music recordings.

Cantu in Paghjella

Cantu in Paghjella is a Corsican polyphonic singing style engaging three singers: Bassu, Seconda and Terza. Cantu in Paghjella recording sessions were organised by CNRS in Bastia, Corsica in May 2014 (Figure 11a). Two groups of three singers were recorded using the hyper-helmet and the i-THRec software [13], developed within i-Treasures by CNRS. More than 10 songs have been performed and vocal tract data of the three singers have been captured. Since only two hyper-helmets were developed by that time, only two performers per session were integrally recorded, while the third one used only a microphone for acoustic acquisitions. The three singers were separated in three distant rooms and the synchronisation of the recordings was made by clapping once near the systems. Radio-Frequency headsets were made available for the Bassu and for the Terza in order for them to hear the Secunda. One song of each group ("Ecco Bella" and "O Salutaris") are concerned by the public dataset (6 performances, i.e. 2 songs by 3 singers). .textgrid and .xml phonetic annotation files are also part of the public dataset.

(a)

(b)

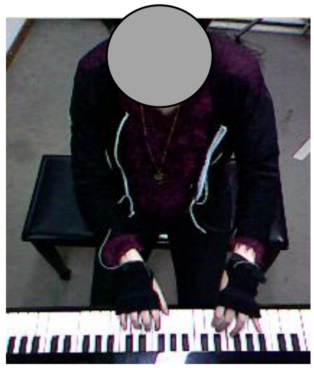
Figure 11. Cantu in Paghjella (top) and Canto a Tenore (bottom) recordings.

Canto a Tenore

Canto a Tenore is a Sardinian polyphonic singing style engaging four singers: Boghe, Bassu, Mezzo and Contra. A group of Canto a Tenore singers was invited to UPMC labs in November 2014 to record five different songs in the two styles: ballu and seria. The three singers were toggling on three available hyper-helmets during this recording (Figure 11b), while the fourth was recorded only by a microphone. Except from vocal tract data, facial data (depth and colour image streams) have also been recorded using Kinect sensors. Three songs (“Amore in seria style”, “Amore in ballu style” and “A Orune in seria style”) are concerned by the public dataset (9 performances, i.e. 3 songs by 3 singers). .textgrid and .xml phonetic annotation files are also part of the public dataset.

Contemporary music composition data

UOM and ARMINES have organized two recording sessions for the contemporary music composition use case. One was organized in Paris, in May 2014, and one in Thessaloniki, in July 2014. Four pianists have been recorded while performing musical ‘piano-like’ gestures with real piano and with Intangible Musical Instrument (IMI) (see Figure 12(a-c)).



(a) (b)

(c)

Figure 12. Recordings for the contemporary music composition use case, while performing musical gestures with real piano (a) and with IMI (b). Their body and finger gestures are captured by the Animazoo suit and the Leap Motion sensors (c). IMI is a construction made of Plexiglas and the whole capture set up consists of the Animazoo suit, Kinect camera, Leap Motion sensors and Emotiv EPOC

In Paris, Animazoo inertial sensors were used (http://synertial.com/, [10]), and more specifically 12 sensors were placed on the upper part of participant's body. These sensors capture rotations (Euler angles). Two pianists, one expert and one learner, were participated while performing the following musical gestures with real piano: a) ascending-descending scale legato, b) ascending-descending scale staccato, and c) ascending-descending arpeggio legato.

In Thessaloniki, Leap Motion Controller (https://www.leapmotion.com/) for finger positions was used. Two expert pianists were participated and they have been recorded while performing musical 'piano-like' gestures with IMI. Four musical gestures were included in the vocabulary: a) ascending arpeggio legato, b) ascending arpeggio staccato, c) ascending-descending arpeggio legato, and d) ascending-descending arpeggio staccato. For arpeggio, experts use the combination of 1,2,3,5 fingerings.

EEG data collection

An EEG dataset from an experiment that targeted the elicitation of affective states (of positive to negative valence and of high to low arousal), by employing music, was acquired and made public. The recordings took place at the Department of Electrical & Computer Engineering of the Aristotle University of Thessaloniki, Greece.

In particular, the Emotiv EPOC [5] was used to record 14-channel brain activity under a 128 Hz-sampling frequency. A 12 subjects (7 females and 5 males; age 24 +/- 2.7 years) × 11 excerpts × various stimulus duration-dataset was acquired from a Beethoven music-based experiment. Each subject gave signed consent prior to the experiment. The general protocol followed is presented in Figure 13, single session and it was implemented using the ExperimentWizard software tool. Excerpts of Beethoven's Tempest and Waldstein piano sonatas were used as stimuli (also included in the public dataset) and subjects assessed the affective states that they experienced due to musical stimulation.

Emotion assessment was based on the Self-Assessment Manikin (SAM) nine point-scales for valence and arousal [5] (Figure 14(a)) and the subject-produced ratings are also included in the dataset. Experiment configuration is shown in Figure 14(b). Acoustic stimuli were provided to subjects through in-ear headphones.

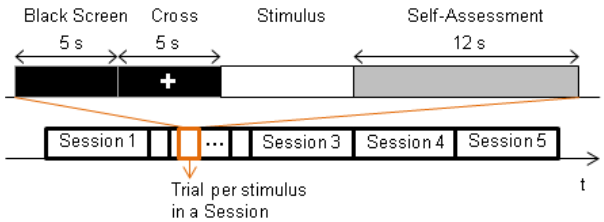
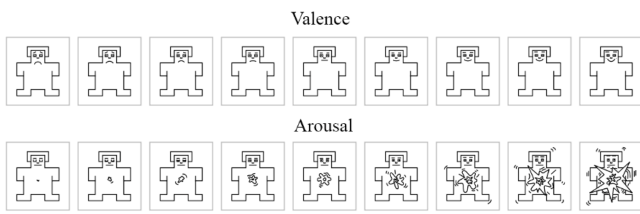
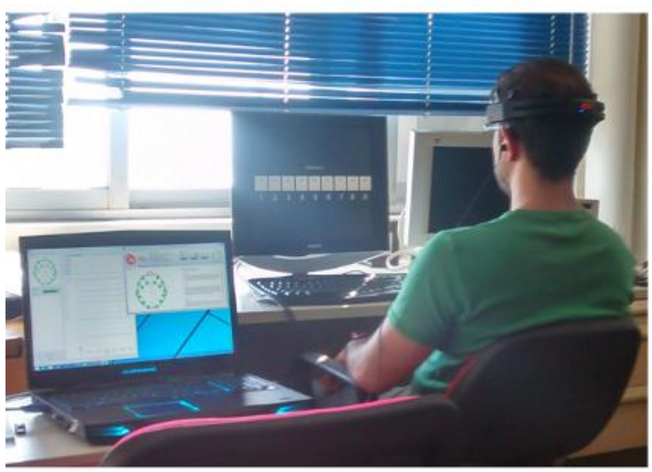


Figure 13. Protocol of music-based (stimuli) emotion elicitation experiment.



(a)



(b)

Figure 14. a) The SAM scales for valence and arousal. b) Experiment set-up.

Pottery data collection

Pottery [14] is the ceramic material which makes up pottery wares, of which major types include earthenware, stoneware and porcelain as well as the place that they are made. Pottery also refers to the art or craft of the potter or the manufacture of pottery and it is considered as one of the more ancient arts. Pottery is made by forming a clay body into objects of a required shape (most commonly vessels) and heating them to high temperatures in a kiln (high temperature chamber), which removes all the water from the clay and induces reactions that lead to permanent changes including increasing their strength and hardening and setting their shape. A clay body can be decorated before or after firing.

The invention of the pottery's wheel (2600-2300 BC) facilitated the production of pottery and lessened the construction time, as a result ceramic art was highly developed. The basic phases and gestures of wheel-throwing pottery vary based on the actual object shape to be formed (e.g. bowl, cylinder or sphere), but are more or less the same in most geographical regions, except some specialized pottery techniques used in specific geographic regions (e.g. India) that are of special interest.

For the pottery use case, a craftsman was recorded under real conditions, with the use of wheel and clay, performing real gestures and creating a bowl. In total 5 bowl have been created and 5 repetitions of each gesture have been captured thus. The process of creation of one bowl contains 4 gestural phases. The recording was performed, using a Time-of-flight sensor, (Camboard Nano PMD camera [9]) and inertial sensors (Animazoo [10]) to capture the gestures of the craftsman. A single depth camera capturing the hands of the potter from top view is used providing depth images with resolution of 165x120 pixels. From the other hand the potter wears a suit covering the upper part of his body with 11 inertial sensors providing the rotations of his body segments. In Figure 15 we can see the potter performing a wheel throwing gesture, while wearing the suit and being captured by the PMD camera.

Figure 15. The potter performs a wheel throwing pottery gesture and is captured by the depth camera and the inertial sensors.

CONCLUSIONS AND FUTURE WORK

In this paper, the i-Treasures Intangible Cultural Heritage (ICH) dataset, a freely available collection of multimodal data captured from different forms of rare ICH, was described. More specifically, the sensors, capture setups and capture software used for dataset were described as well as the (manual) annotations provided. An attractive feature of this ICH Database is that it's the first of its kind, providing annotated multimodal data for a wide range of rare ICH types.

In the future, we aim to enhance this dataset and provide a second version with additional recordings that will be obtained during the demonstration and evaluation phase of i-Treasures project. Additional results of analysis and modeling modules (feature extraction and classification) developed within i-Treasures project will also be provided.

ACKNOWLEDGMENTS

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7-ICT-2011-9) under grant agreement no FP7-ICT-600676 "i-Treasures: Intangible Treasures - Capturing the Intangible Cultural Heritage and Learning the Rare Know-How of Living Human Treasures".

REFERENCES

1. UNESCO, "What is Intangible Cultural Heritage?", Retrieved May 13, 2016 from http://www.unesco.org/culture/ich/index.php?lg=en&pg=00002
2. Meet Kinect for Windows. Retrieved May 13, 2016 from https://dev.windows.com/en-us/kinect
3. Qualisys motion capture system. Retrieved May 13, 2016 from http://www.qualisys.com/
4. OpenNI, the standard framework for 3D sensing. Retrieved March 10, 2016 from http://www.openni.ru/
5. Emotiv EPOC neuroheadset. Retrieved May 13, 2016 from http://emotiv.com/epoc/
6. P. J. Lang, M. M. Bradley, and B. N. Cuthbert. 2008. *International affective picture system (IAPS): Affective ratings of pictures and instruction manual*, University of Florida, Gainesville, FL, Technical Report A-8..
7. M. M. Bradley and P. J. Lang. 1999. *International affective digitized sounds (IADS): Stimuli, instruction manual and affective ratings*, The Center for Research in Psychophysiology, University of Florida, Gainesville, FL, Technical Report No. B-2.
8. M. M. Bradley. March 1994. Measuring emotion: the self-assessment manikin and the semantic differential, *Journal of Behavior Therapy and Experimental Psychiatry*, Vol. 25, No. 1, 49-59.
9. pmd[vision]® CamBoard nano: http://www.pmdtec.com/products_services/reference_design.php
10. Animazoo motion capture systems: http://www.animazoo.com/products
11. Leo Breiman. 2001. Random forests. Machine learning 45, 1: 5-32.
12. B. Picart, S. Brognaux and S. Dupont, "Analysis and automatic recognition of Human BeatBox sounds: A comparative study," Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on, South Brisbane, QLD, 2015, pp. 4255-4259.

13. Patrick CHAWAH et al., “An educational platform to capture, visualize and analyze rare singing”, Interspeech 2014 conference, Singapore, September 2014

14. Wikipedia Pottery page, Retrieved May 13, 2016 from http://en.wikipedia.org/wiki/Pottery

The original publication is available at http://dl.acm.org/citation.cfm?id=2948944.