A Framework for Creating Cultural Interactive Guides

Antonio Sorgente, Antonio Calabrese, Gianluca Coda, Paolo Vanacore, and Francesco Mele

Institute of Applied Sciences and Intelligent Systems "Eduardo Caianiello" of the National Research Council
Via Campi Flegrei 34, 80078 Pozzuoli (Naples) Italy
{a.sorgente, a.calabrese, g.coda, p.vanacore, f.mele}@isasi.cnr.it

Abstract. The use of mobile technologies and Augmented Reality are necessary in all modern museum applications in which is required an active participation of the visitor. Many systems have been defined for this aim, but each one offers contents that has been strictly selected in the design phase. In this work, we present a platform to define and make programs that can be used for assisting a visitor by providing a particular interaction chosen during the visit.

Keywords: artificial guide, multimedia presentation, multimedia composition, microservices architecture, annotation events interface

1 Introduction

Experts and practitioners of museums, and generally those who work on cultural heritage, make great efforts to improve the experiences of visitors in the cultural spaces and the application of new technologies in such a domain is growing day by day [1, 2]. The challenge in this area consist in/of offering contents to a visitor in a natural and attractive way taking into account the spatial context of the visit, in relation to what the visitor is seeing; his interests and available time to spend on the visit [3, 4].

In addition, the interaction based on speech and natural language processing make the museum visit more attractive and improve the overall experience quality [5]. Finally, to have an immersive and exhaustive interaction new interfaces have to be designed to fit within the museum environment and meet user needs. Such interfaces have to be adaptable to various devices, such as mobile or wearable devices, personal computer, etc.

In recent years, many systems of multimedia museum guides have been defined to improve the dissemination of cultural heritage. They are of different types, traditional audio guides presenting explanations only through audio, mobile applications that show images and video in addition to audio [4, 6], and systems that use Augmented Reality(AR) and/or Virtual Reality(VR) [7–9]. In

these systems, the type of information and its presentation is selected during the design phase.

In this paper, we present a platform deploying various services for the developing of mobile systems for multimedia presentation of cultural content adopting the *visit with augmented knowledge* paradigm [10]. The main aim is to propose services for: loading information contents through an event formalism as reference, using temporal reasoning services; defining services allowing the understanding of users requests, submitting of the best answer and the composition of multimedia response; for reproduction of the multimedia response.

This work is focused on presentations regarding cultural contents. The system has been defined using ARToolkit[1]. ARToolKit is a software library for building Augmented Reality (AR) systems that contains useful algorithms for image recognition both based on markers and marker less.

2 Architecture

The system is designed to define and make services to assist a user during a visit, but also, to assist the museum curator to facilitate the loading of new information that will then be used by the virtual guide. Our attention in this work is focused on the presentation of the information and not on its generation.

The realised system is implemented adopting a Microservice Architecture [11]. We have chosen this architecture as it offers high scalability and reliability characteristics through fault tolerance mechanisms. Considering the real-time nature of interaction systems, these features are crucial. The high degree of decoupling, obtained through event-based communications, and the high services cohesion, allow us to implement customised load balancing and scalability mechanisms. Furthermore, a low impact evolution, through the introduction of new services, duplication and/or modification of existing ones, is possible.

Fig. 1 shows the overall architecture of the system. The services deployed are split with respect to the type of developer. So, there are services for the museum curators allowing them to prepare and organise the multimedia that will be used during the visit. An example of such an application built on these services is described in [12]. There are services used to develop the interactive guides which are useful for interaction in natural language. Some details of such services are described in [13, 14]. Finally, there are common services.

Generally, the museum visit design takes a long time to be prepared, for the definition of contents and presentation modes. Through these services, museum curators can collect data and select the content that they want to convey to visitors as well as defining a catalogue, although they may not also have specific expertise about technologies and methodologies adopted. Also, the platform is independent from a specific museum and it is easily adaptable and reusable in different museums. In addition, other time is spent on the realisation of the interactive guide and related user's interfaces. Each system has its own presentation style defined at the developing time.

[1] https://artoolkit.org/

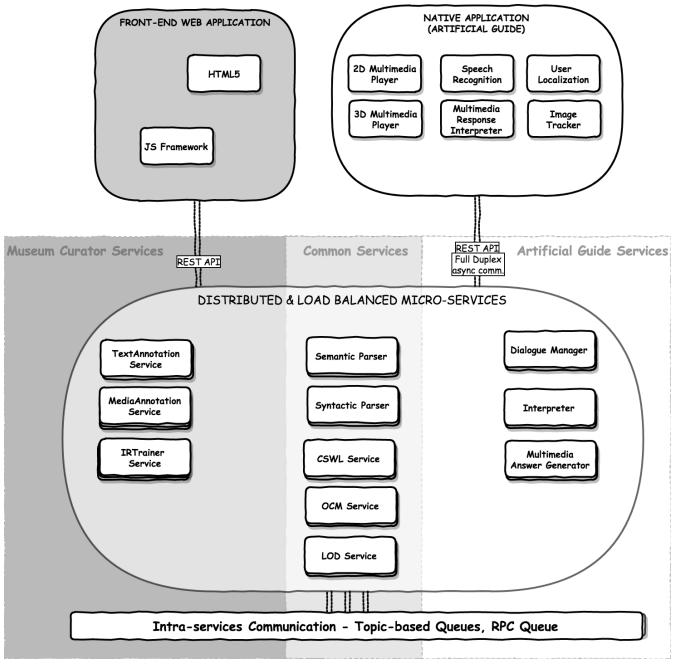

Fig. 1: System Microservices Architecture

3 Application for the Museum Curator

During the definition of an interactive guide, one of the main issue is related to the selection and representation of information to be provided to visitors. Our aim has been one to define an interface that allows a museum curator to load texts and images related an artwork, and make sure that such information is available for the interactive guide to enhance the experience of the visitors.

Fig. 2 shows a screen-shot of Hunter Events Interface (HEI)[12]. The main aim of HEI is to have an environment that allows an user to annotate temporal entities (in automatic and in assisted way), and to interpret such annotations. These latter are useful for services used by the interface for managing the user interaction in natural language. Details of HEI system is presented in [12].

4 Cultural Interactive Guide

The interactive guide, using services deployed by the the micro-services platform, allows us to interact in natural language and receive multimedia response (audio, images and video). A particular implementation in [13, 14] was presented. In this section we report some description modules and the interface that are used to visualise the multimedia response produced by dialogue system.

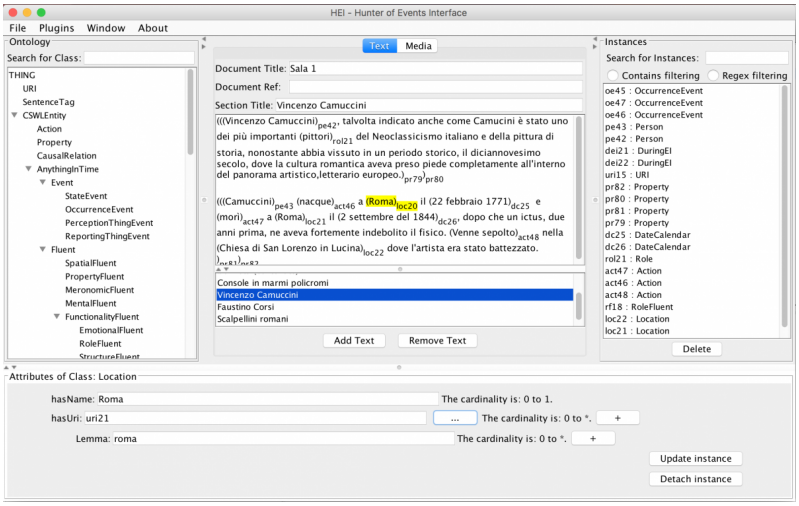

Fig. 2: HEI Annotation GUI

Basic functionalities required by such an application are: to recognise user's requests; to give the response to the user; and to detect the user position. Fig. 3 shows the main modules built in Android environment[2]. The "User Request Module" is responsible for acquiring the user's requests. These can be written in natural language through the keyboard device - "Keyboard Listener" - or through a speech recognition - "Speech Recognition Module" - using the Android Speech Recogniser. "User Positioning" is the module for the localisation of visitor's position in the museum. This can be done through many approaches using different sensors and technologies. Currently, we have implemented the image recognition module ("Image Recognition") based on the Open Source software ARToolKit[3]. The realised framework can be extended easily adding others way to locate the user, and so it is able to detect, in a non-exclusive manner, the user's position through modules that make use of Beacon, Global Positioning System (GPS), Wireless Positioning System (WPS), Depth sensors and so on. Finally, the multimedia responses generated by the dialog system, are interpreted and reproduced through one of the "MMResp Player" modules. In the next section we present how the guide plays the multimedia response.

4.1 MMResp Player

The task of MMResp consists of analysing the response of dialogue system and visualises it. An example of multimedia response is shown in listing 1. It is represented in JSON format. Such a script contains references and execution times

[2] Minimum API Level 15 (Android Ice Cream Sandwich - ver.4.0.3-4.0.4), Target API Level 24 (Android Nougat - ver.7.0)

[3] https://artoolkit.org/

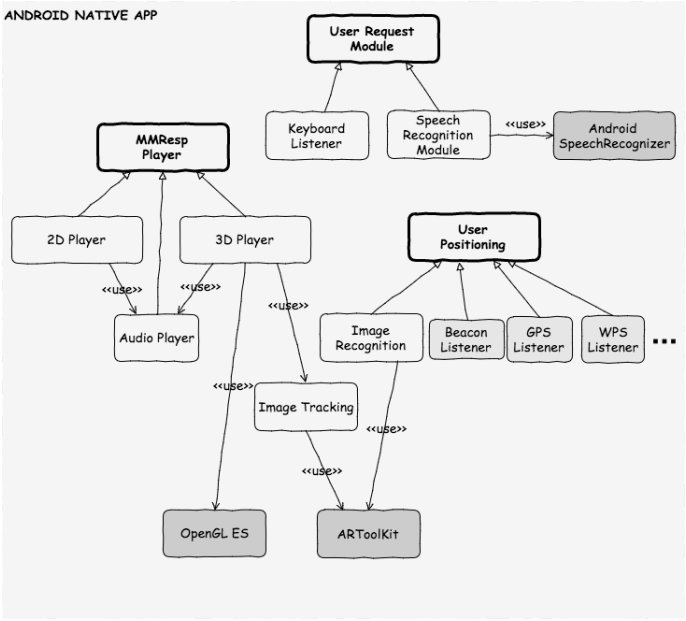

Fig. 3: Android Native Application modules

of media files (images and audio) that the module *MMResp Player* reproduces on the user's device.

Listing 1: json of multimedia response

```
{   "images": [
        {   "regionId": "region1",
            "timing":  {"duration": 2290, "start": 0 },
            "mediaUri": "http://smcm.isasi.cnr.it/SIMArt/media/giulio_cesare.JPG",
            "mediaId": "1559",
            "subimage": {  "size": { "width": -1, "height": -1 },
                           "origin": { "x": 0, "y": 0 }
                        }
         },
            ...
    ],
    "audios": [
        {   "timing": { "duration": 39000, "start": 0 },
            "mediaUri": "http://smcm.isasi.cnr.it/SIMArt/tts/contenuto-12.mp4",
            "mediaId": "1491",
        }
    ]
}
```

The module reproduces the answers in three different way: audio only, when the screen is locked (for example the device is put in a pocket); 2D mode when the user watches the multimedia response without aiming the device at the artwork; finally, in 3D mode when the artwork is framed by the device. Our attention was

focused on 3D mode. Some systems statically position the extra information in the space. So, one has found such information in the environment. Our idea is to fix additional information in the museum space taking into account the users point of view. We have based the 3D media projection on "useful/useless areas" concepts (Fig. 4). When the system recognises an artwork captured by the device camera, the screen area around the artwork is analysed. In this way we can project the 3D multimedia responses on the best free space of the screen without media overlapping the artwork. As shown in Fig. 4, after the device captures the artwork and it is recognised, we detect the green areas that have enough space to contain an image, video or 3D object.

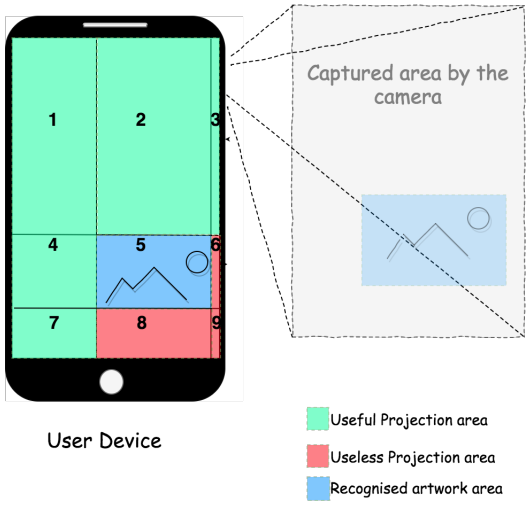

Fig. 4: 3D projection - useful/useless areas

After the detection of an artwork, the screen is partitioned (all rectangles shown in Fig. 4) by lengthening the extremities of the rectangle that contains the recognised artwork. In the first step, all the areas that have a minimum size are labelled as useful (1,2,4). Then, for each area not labelled, if it borders with an Useful Area, then it becomes Useful (3,7). In the latter case, it can not be used alone, but always combined with useful areas detected in the first step. The remaining areas are useless (6,8,9).

5 An example of application

We have implemented an interactive guide based on the modules presented in the previous sections. In this phase of the project, the aim is to evaluate the system from a functional point of view and evaluate the communication among all modules.

Fig. 5 shows a presentation example of a multimedia response. The picture is a self-portrait of Vincenzo Camuccini, and the extra information reported is the painting "La morte di Virgina". In the scene we can see that the image is added to the best area fitting with the media object. In Fig. 5a, the media is shown on left, while in Fig. 5b is shown on right and fits the perspective with respect to the visitor's point of view. The captured screen-shot shows the "3D" switch button to enable or disable the 3-dimensional visualisation of multimedia responses (the example of 2D is shown in Fig. 5c) and the "Speech" button to start a natural language interaction.

(a) (b)

(c)

Fig. 5: Presentation of multimedia response

6 Conclusions and Future Work

In this work, we have presented a micro-service platform that deploys services for preparing and organising information contents about cultural assets, and defining interfaces for the interactive guides building where the interaction is made through a mobile device. The next research steps will be: 1) to adapt these services and evaluate their usability with respect to other typology devices, such as smart-glasses; 2) to evaluate the satisfaction of the visitors and museum curators with respect to the use of such applications.

References

1. De Carolis, B.N., Gena, C., Kuflik, T., Nunnari, F., eds.: Proceedings of the 1st Workshop on Advanced Visual Interfaces for Cultural Heritage (AVI*CH 2016). Volume 1621 of CEUR AI*IA series. (2016)
2. Bordoni, L., Mele, F., Sorgente, A., eds.: Proceedings of the 10th International Workshop on Artificial Intelligence for Cultural Heritage (AI*CH 2016). Volume 1772 of CEUR AI*IA series. (2016)
3. Ardissono, L., Kuflik, T., Petrelli, D.: Personalization in cultural heritage: the road travelled and the one ahead. User Modeling and User-Adapted Interaction **22**(1) (2012) 73–99
4. Eghbal-Azar, K., Merkt, M., Bahnmueller, J., Schwan, S.: Use of digital guides in museum galleries: Determinants of information selection. Computers in Human Behavior **57** (2016) 133 – 142
5. Stock, O.: Language-based interfaces and their application for cultural tourism. AI Magazine **22**(1) (2001) 85
6. Kuflik, T., Stock, O., Zancanaro, M., Gorfinkel, A., Jbara, S., Kats, S., Sheidin, J., Kashtan, N.: A visitor's guide in an active museum: Presentations, communications, and reflection. J. Comput. Cult. Herit. **3**(3) (February 2011) 11:1–11:25
7. Vlahakis, V., Karigiannis, J., Tsotros, M., Gounaris, M., Almeida, L., Stricker, D., Gleue, T., Christou, I.T., Carlucci, R., Ioannidis, N.: Archeoguide: First results of an augmented reality, mobile computing system in cultural heritage sites. In: Proceedings of the 2001 Conference on Virtual Reality, Archeology, and Cultural Heritage. VAST '01, New York, NY, USA, ACM (2001) 131–140
8. Wojciechowski, R., Walczak, K., White, M.: Augmented reality interface for museum artefact visualization. In: Proceedings of the 3rd IASTED International Conference on Visualization, Imaging and Image Processing. (September 2003) 998–1004
9. Chen, C.Y., Chang, B.R., Huang, P.S.: Multimedia augmented reality information system for museum guidance. Personal and Ubiquitous Computing **18**(2) (2014) 315–322
10. Sorgente, A., Calabrese, A., Coda, G., Vanacore, P., Mele, F.: Building multimedial dialogues annotating heterogeneous resources. In Bordoni, L., Mele, F., Sorgente, A., eds.: Artificial Intelligence for Cultural Heritage. Cambridge Scholars Publishing (2016) 49–82
11. Lewis, J., Fowler, M.: Microservices - a definition of this new architectural term (2014) Online: https://martinfowler.com/articles/microservices.html.
12. Sorgente, A., Calabrese, A., Coda, G., Vanacore, P., , Mele, F.: Hei: Hunter events interface. In: Proceedings of 1st Workshop on Temporal Dynamics in Digital Libraries. CEUR Workshop Proceedings, CEUR-WS.org (2017) will be published
13. Sorgente, A., Vanacore, P., Origlia, A., Leone, E., Cutugno, F., Mele, F.: Multimedia responses in natural language dialogues. In: Proceedings of AVI*CH 2016. Volume 1621 of CEUR Workshop Proceedings., CEUR-WS.org (2016) 15–18
14. Origlia, A., Leone, E., Sorgente, A., Vanacore, P., Parascandolo, M., Mele, F., Cutugno, F.: Designing interactive experiences to explore artwork collections: a multimedia dialogue system supporting visits in museum exhibits. In: Proceedings of the 10th AI*IA 2016. Volume 1772 of CEUR Workshop Proceedings., CEUR-WS.org (2016) 26–33