The Imitation Game to Cultural Heritage: A Human-like Interaction Driven Approach for Supporting Art Recreation

Fiammetta Marulli[1(✉)] and Luca Vallifuoco[2]

[1] DIETI, University of Naples Federico II, Naples, Italy
fiammetta.marulli@unina.it
[2] MoonRiver Graphic Art Studio, Naples, Italy
luca.vallifuoco@live.it

Abstract. Smart IoT technologies set a milestone in supporting new enjoyment models for Art and Cultural Heritage, providing amazing technological experiences. However, users, while interacting almost purely by mediation of smart devices or augmented VR displays, practically keep themselves out from living the fullness of the surrounding cultural sites experience, establishing no direct dialogues or interactions with artworks. It sounds quite like to be in the living-room, looking at very appealing documentaries, equipped with exciting smart technologies. This paper focuses on the importance of "re-humanize" art recreation models, proposing a human-like interaction driven approach. Holographic projections, reproducing human or fantasy characters, play the human presence imitation game, when users are detected close to any artwork, interacting and dialoguing with them in natural language. An IoT infrastructure, an NLP platform and a Holographic Projection Engine implement a system for supporting holographic projections. Preliminary experiments were promising, thus motivating authors to further investigations.

Keywords: Cultural Heritage · Art recreation paradigms and models · Human-centered computing · Natural language interfaces · Natural Language Processing · Speech recognition · Holographic simulations · Internet of Things · Smart environments

1 Introduction and Related Works

Current Internet of Things (IoT) smart technologies set an effective milestone in supporting novel models for enjoying Art and Cultural Heritage (CH). The widespread "talking" museum and "smart sites" [1] or the augmented virtual reality (AVR) systems [2] evidence successful applications. This type of applications implements very exciting and advanced solutions (e.g., multimedia smart guides [3], provided with intelligent profiling systems for multimedia contents tailoring [4–6]), but it turns into a too strongly device-driven approach. Indeed, the experience of users visiting cultural sites or artworks exhibitions, is strongly mediated by the usage of personal smart devices or provided AVR displays (e.g., the Oculus Rift display).

A.L. Brooks and E. Brooks (Eds.): ArtsIT/DLI 2016, LNICST 196, pp. 145–153, 2017.
DOI: 10.1007/978-3-319-55834-9_17

An increasing number of present researches focuses on studying the "appeal" and the appreciation level of such technologies against users' emotions and involvement, when they are applied to CH domain. Some preliminary results from these studies [7, 8] highlighted a still not significant increase or improvement in the cultural attitudes and sensitivity of the users, after a "smart" visit experience.

Most of "common" users (those ones having no particular skills in the art field) are strongly charmed by smart and amazing new technologies but very little engaged with the cultural experience offered by living museum atmospheres. On the other side, art lovers and experts are more interested to establish a kind of deep "sentimental" correspondence and dialogue with artworks and museum atmosphere (by admiring and touching objects, whenever it's allowed), so discarding quite fully the usage of technologies.

This kind of antithetical behaviors evidences a still weakly pervasive and still intrusive approach in creating synergies between smart technologies and cultural sites. Smart cultural approaches should be able to engage users really, without jeopardizing their interest towards the cultural context or the support offered by smart Information Technologies (ITs). Very often, users, while interacting by their smartphones or AVR displays, keep themselves out quite fully from living and experience the surrounding cultural site. No "dialogue" or direct deeply pulsing interactions are established with the artworks. A museum visit would risk to "sound" quite like the same to be in the own living-room, looking at very interesting documentary packed in contents, equipped with amazing technologies, thus putting users fully out of the cultural site atmosphere.

Another not trivial issue concerns the selection and organization of an appropriate knowledge to transfer to users. They, differencing by cultural and social background, by age and sensitivity, have to be approached in different ways, if the goal is an effective engagement with the context they are experimenting. The author of [9] proposes an authoring platform for automatic generation of tailored textual artworks descriptions, basing on users' profiling information (e.g., art biographies in the shape of little fables for schoolchildren).

Therefore, main limitations of available smart cultural service infrastructures consist exactly in the way they set the communication paradigm between visitors and cultural objects. Better profiled interactions, exploiting the means of human-like natural interactions and natural language could help in overtaking these problems.

After all, the easiest way for a user to acquire information and to express his/her needs and preferences regarding a desired service or condition, is to use natural language. Incompleteness and ambiguity, typical of natural languages, offer rooms for discussions, but present textual entailment and paraphrases methods [10, 11] provide relevant contributions to limit these problems.

Starting from the discussed standpoint, this paper focuses on the importance of "re-humanize" art recreation models, proposing a human-like interaction driven approach, that could contribute to reinforce the role played by the most advanced smart ITs in reaching the goal of a deep and effective involvement in CH of human visitors.

As detailed in the following, a natural language driven process, exploiting naturally speaking holographic projections, is proposed to make art objects "speaking", thus reestablishing a natural communication with visitors, who are so provided with the opportunity, when in front of an artworks, to be part of new experiences of knowledge.

Holograms, reproducing human figures or fantasy characters (for children audience), play the human presence imitation game, when users are detected nearby an artwork. By exploiting the potentiality of IoT well-assessed infrastructures and applications, an NLP platform and a Holographic Simulation engine, a smart human-like driven model of interaction is implemented in a unique platform. It allows to manage holograms projections and interactions in a humans-like way, dialoguing with human users in natural language. Different appearances for holographic characters can be proposed to support a further customization and variation for the performed interactions. In this perspective, the proposed approach and its implementing system provide support for selecting very popular fables and fantasy cartoons (e.g., Walt Disney Production) as holograms to implement an effective human-like speaking artworks environments. Finally, we illustrate this possibility by showing how an IoT environment can implement this approach into CH.

2 The Imitation Game Approach for Holographic Characters

IoT, coupled with existing capabilities of speech processing and dialogue management in natural language, put the basic prerequisites for an ecosystem of human-like speaking things. This will enable a direct and natural connection between human users and the information cloud, thereby making effectively "smart" not only terminals like computers and mobile phones but also physical environments such as cities, homes, offices, shops and museums. Talking to a disembodied human voice is not like talking to a human figure. Aside the fuller user experience that it provides, human-like figures can be used to interact with users whenever there is need of visual demonstration, as in explaining an artwork.

Holographic human beings bring one step further the capabilities of interaction of "speaking" things, in that they can provide support for full-fledged dialogues in natural language. This basic technology allows capturing, in the form of holographic simulations, sequences of actions performed by human actors that can be matched with requests coming from human users through the integration with technologies for speech processing, natural language understanding, gesture recognition, linked data and knowledge representation. Deployment is through standardized carriers such as ordinary projector optimized for 3D displays with very high resolutions (at least 4 K UHD), always maintaining full size reproduction of the holographic human being so as to make the user experience totally natural and familiar. Since different types of audiences should be enjoyed, diversified holographic resemblances are supported for children and very young people: fantasy characters from fables, cartoons and animation movies (Disney Pictures, e.g.) were selected because of their popularity and the large availability of video materials for graphic elaborations.

Furthermore, holograms of human appearance are obtained by projecting prerecorded video segments, depicting human actors performing a set of facial expressions, mimics and pronouncing a set of answers and questions for users. By the way, the number of different video segments to record, in order to offer a reasonable level of variety for different audiences and coverage for a small museum or exhibition (about 100 artworks),

turns into a bottleneck, demanding significant times and economic resources waste. By adopting animation movies and cartoons characters, the effort due to implement the imitation game for speaking holograms is reduced to dubbing characters only. Such a solution enhances our choice to select cartoons and fables characters to generate holographic figures, beyond the appealing and the popularity of such figures at children and young people eyes.

Figure 1 shows a prototyping experimental installation deployed at the Civic Museum of CastelNuovo, in Naples (Southern Italy), where the fable character "Alice in Wonderland", appears in one of the exhibition rooms, telling the visitors about a set of beautiful marbles and gypsum Italian sculptures.

Fig. 1. Alice Hologram waiting for interactions with human visitors.

3 The Natural Language Driven Process for Speaking Holograms

A fundamental step in supporting interactivity is to make the holograms cloud-connected and capable of transferring information back and forth over the Internet. In fact, this is a necessary condition for the holographic figures to be able to answer the requests of the users providing them with the needed information. The IoT brokerage services are essential in this respect, being this a typical case of communication between the machines maintaining the information and the "animated thing", namely the holographic figure, that would provide them to the user. Pro-active and reactive event processing, audio mining, content optimization and context-aware recommendation can be also exploited effectively to turn holograms into revolutionary user experiences and interfaces. A Natural Speech Recognition system and a Questions & Answers (Q&A) engine

are the core components. The Q&A engine is employed to find matching answers for well-formed requests.

At a glance, the processing flow is composed of the following four steps:

1. *Speech to Text conversion.* An instance service is listening on a communication bus, waiting for a user' s request. Speech recognition and translation into textual format was implemented employing Microsoft DotNet 4.0 Framework Solution (ASR and TTS SDKs).
2. *Text Analysis and Request Categorization*. Typical text analysis and text entailment methods are applied to incoming textual requests to identify a correct category. The output is represented by lists of relevant terms and categories useful to match suitable answers and to select video segments for playing holograms.
3. *Question and Answer Matching*. Categories and terms summarizing the request are matched against a Knowledge Base (KB). CH experts provided specific domain ontologies, vocabularies and answers to populate effectively the underlying KB.
4. *Text to Natural Language Speech Generation.* When an answer matches against a submitted request, an event consisting in playing a speaking hologram is created and projected close to the inquired artwork.

4 System Architecture and Components

The capability of involving users into truly engaging interactions derives also from the easiness with which the implemented system for NLP can be integrated within a modular architecture that takes advantage of a variety of technologies from IoT. It is composed, at a glance by:

1. An IoT Application Infrastructure, including:
   - A *Wireless Sensor Network* (WSN), typically made of smart beacons devices, to detect human users' locations and movements, by the exploitation of the most advanced Bluetooth low power consumption technologies;
   - An *Application* (an App) for smart devices, detecting users' presence and exploiting smart devices microphone to collect users' questions.
   - A *Message Delivery Service* (MDS), responsible for receiving vocal inputs incoming from the App for delivering it to the ADM subsystem.
2. An Action and Dialogue Manager (ADM), including:
   - a *Natural Language Processing Engine* (NLPE), engaged with processing vocal signals forwarded by the MDS, and translating them into textual messages. It bases on Voice Recognition and Speech to Text services.
   - A *Questions & Answers* (Q&A) engine, deputed to process the textual request against a KB containing a set of possible answers. This engine is currently implemented as a set of logic and semantic rules to extract the most meaningful terms composing the request and to entail similar ones;
   - A *Video Dispatcher*, receiving the selected answer and using it as a key to query a Multimedia NoSQL Database (MongoDB), to select the corresponding video segment for the holographic figure to be played.

5 A Case of Study: An IoT Environment for Speaking Holograms

For a deeper comprehension of the Human-driven interactions environment, performed by the speaking holograms, we considered the following experiment, a tourist visiting the National Archeological Museum of Sannio Caudino (Benevento, Southern Italy) exhibiting ancient Greek vessels, among which the precious Crater of Assteas. This environment offers a WSN, enabling visitors to interact with the cultural objects exhibited, by exploiting their personal smart devices and an appropriate App. The typical scenario of our case of study consists in a user visiting an exhibition room and walking through several ancient Greek bowls and objects; when he is particularly next to one of them, his presence is detected, because his mobile device (running a provided App) interacts with and a sensor coupled to that object (placed on the basement, for example). This condition triggers the cultural object animation, by projecting a holographic figure; it immediately engages a dialogue, in human natural language with the visitor. The hologram invites its interlocutor to learn more about the place and the object he is close to. At the first interaction, the holographic figure welcomes the user, proposing messages in four different languages, at the current version (Italian, English, French and Spanish).

Figure 2 shows the use case for a visitor next to an ancient vessel, thus triggering the projection of the holographic illusion.

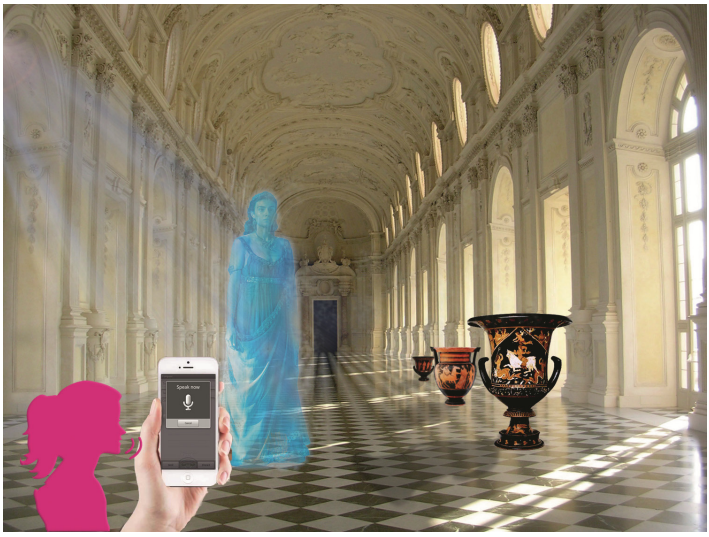

Fig. 2. An example of Speaking Hologram interaction with a user by IoT

Our approach builds on an existing paradigm of pro-active interaction but at the same time forces a significant paradigm shift, where interactions are effectively driven by natural language. Answers are provided as short user categories profiled narrations. When a visitor is detected close to any art object, the hologram figure invites the user to submit a question (suggests some questions and topics) or a request for information (about history, author,

style and techniques, etc.) by using his/her personal smart device. User speaks to the hologram as a human guide, submitting a question among the suggested ones. The visitor can also discard the dialogue request, thus interrupting the conversation (pronouncing some "magic stop words") with the hologram just like in a real human dialogue. The underlying system supporting the holographic projection, automatically recognizes the user's language response at the first time interaction. If a request is correctly recognized by the backend underlying system, a video segment, playing the hologram acting the correct matching answer, is projected. The matching process, as described in the previous section, is based on a NLP and a Q&A engine; it selects the corresponding answer, given a well identified and recognized question or a significant part of it. If any question can't be recognized by the system (too noisy environment, bad spelling, unsupported languages or questions), an idle video segmenti s projected; thus, the hologram, asks the user to submit his/her request once again. A number of trials have been performed to assess users' enjoyment and consequently, the naturalness and the utility of the proposed application. A sample of about 500 visitors were asked to fill in a post-visit questionnaire, during a set of events celebrating the exhibition of the precious Crater of Assteas. Users were stimulated to express their level of agreement with a set of statements, using a 10-point Likert scale, or to make choices between proposed options. We adopted four usability dimensions to have an overall estimation for the proposed approach: simplicity (SIM), usefulness (USN), enjoyment (ENJ) and the naturalness of the interaction (NAT). Table 1 summarizes results extracted from the users' answers, the most relevant questions related to the four dimensions considered and their average ratings. The overall degree of satisfaction manifested by users towards the proposed approach was positive with an average rating of 8.86 (ENJ04).

Table 1. Post-visit questionnaires results.

Question ID	Description	Score
SIM01	It was easy to interact with the exhibit artworks	8.56
NAT01	I appreciate the clearness of the spoken dialogue	8.32
NAT02	The waiting time in the performing interaction was in my expectations	7.89
NAT03	I appreciate the naturalness of the interaction with the environment	8.45
ENJ01	I appreciate the artworks detection metaphor	8.45
ENJ02	Using the infrastructure contributed to increase my will to visit other art exhibitions and my knowledge	8.09
ENJ03	System positively contributed to the enjoyment of my visit	8.87
ENJ04	I overall appreciated the proposed approach and the infrastructure	8.86

6 Conclusions and Future Work

This paper presents a human-like interaction driven approach, in the advanced IoT application context, supporting the Cultural Heritage environment, also suitable for Tourism and Smart Urban Environments, by advancing the available user-experience based on smart devices via the interaction with really and human-like speaking things and holograms. We illustrate this possibility through the description of the dialogue management system and the system architecture supporting speaking things and holograms. Holographic figures represent a variation on the semblance of the intelligent pre-technological interface par excellence, that is, the human figure. In this way, a "static" art exhibition, supported by a fully device-driven interaction, turns into a "living" one and the speaking hologram interact with the visitor, by exchanging questions and answers, in a human natural way. The adopted communication strategy bases on selecting different holographic figures (real human actors performing videos or fantasy characters) to display, according to the audience. Particular attention has been paid to the children audience, in order to make culture and art environment more charming for their learning. The adopted approach in the platform design promises to be scalable and flexible enough to support extensions. As future work, more sophisticated text analysis and semantic based strategies will be exploited, to support more dynamic and complex dialogues. Preliminary experiments evidenced curiosity and appreciation by visitors, so motivating the authors to further improvement.

References

1. Chianese, A., Marulli, F., Moscato, V., Piccialli, F.: A "smart" multimedia guide for indoor contextual navigation in Cultural Heritage applications. In: IEEE IPIN 2013, pp. 1–6 (2013)
2. Brancati, N., Caggianese, G., Frucci, M., Gallo, L., Neroni, P.: Touchless target selection techniques for wearable augmented reality systems. In: Damiani E., Howlett R., Jain L., Gallo L., De Pietro G. (eds.) Intelligent Interactive Multimedia Systems and Services, pp. 1–9. Springer. Cham (2015)
3. Chianese, A., Marulli, F., Moscato, V., Piccialli, F.: SmARTweet: a location-based smart application for exhibits and museums. In: IEEE SITIS 2013, pp. 408–415, December 2013
4. Chianese, A., Marulli, F., Piccialli, F., Valente, I.: A novel challenge into Multimedia Cultural Heritage: an integrated approach to support cultural information enrichment. In: IEEE SITIS 2013, pp. 217–224, December 2013
5. Bordoni, L., Ardissono, L., Barceló, J.A., Chella, A., de Gemmis, M., Gena, C., Iaquinta, L., Lops, P., Mele, F., Musto, C., Narducci, F., Semeraro, G., Sorgente, A.: The contribution of AI to enhance understanding of Cultural Heritage. Intelligenza Artificiale **7**(2), 101–112 (2013)
6. Ardissono, L., Kuflik, T., Petrelli, D.: Personalization in cultural heritage: the road travelled and the one ahead. User Model. User-Adap. Inter. J. Personalization Res. **22**(1), 73–99 (2011)
7. Benedusi, P., Chianese, A., Marulli F., Piccialli, F.: An associative engines based approach supporting collaborative analytics in the internet of cultural things. In: Proceedings of the 10th International 3PGCIC 2015 Conference, November 2015

8. Benedusi, P., Marulli, F., Racioppi, A., Ungaro, L.: What's the matter with cultural heritage tweets? an ontology–based approach for CH sensitivity estimation in social network activities. In: IEEE SITIS 2015, November 2015
9. Marulli, F.: IoT to enhance understanding of cultural heritage: fedro authoring platform, artworks telling their fables. In: Atanasovski, V., Leon-Garcia, A. (eds.) FABULOUS 2015. LNICST, vol. 159, pp. 270–276. Springer, Heidelberg (2015). doi: 10.1007/978-3-319-27072-2_35
10. Androutsopoulos, I., Prodromos, M.: A survey of paraphrasing and textual entailment methods. J. Artif. Intell. Res. **30**, 135–187 (2010)
11. Piccialli, F., Marulli, F., Chianese, A.: A novel approach for automatic text analysis and generation for the Cultural Heritage domain. In: Multimedia Tools and Applications (2016)