International Journal of Electronic Government Research
Volume 12 • Issue 2 • April-June 2016

A Location-Aware Architecture for an IoT-Based Smart Museum

Giuseppe Del Fiore, IDALab, Department of Innovation Engineering, University of Salento, Lecce, Italy
Luca Mainetti, IDALab, Department of Innovation Engineering, University of Salento, Lecce, Italy
Vincenzo Mighali, IDALab, Department of Innovation Engineering, University of Salento, Lecce, Italy
Luigi Patrono, IDALab, Department of Innovation Engineering, University of Salento, Lecce, Italy
Stefano Alletto, ImageLab, University of Modena and Reggio Emilia, Modena, Italy
Rita Cucchiara, ImageLab, University of Modena and Reggio Emilia, Modena, Italy
Giuseppe Serra, ImageLab, University of Modena and Reggio Emilia, Modena, Italy

ABSTRACT

The Internet of Things, whose main goal is to automatically predict users' desires, can find very interesting opportunities in the art and culture field, as the tourism is one of the main driving engines of the modern society. Currently, the innovation process in this field is growing at a slower pace, so the cultural heritage is a prerogative of a restricted category of users. To address this issue, a significant technological improvement is necessary in the culture-dedicated locations, which do not usually allow the installation of hardware infrastructures. In this paper, we design and validate a no-invasive indoor location-aware architecture able to enhance the user experience in a museum. The system relies on the user's smartphone and a wearable device (with image recognition and localization capabilities) to automatically deliver personalized cultural contents related to the observed artworks. The proposal was validated in the MUST museum in Lecce (Italy).

KEYWORDS

Artwork Recognition, Cloud, Embedded Systems, IoT, Localization, Location Awareness, Mobile, Smart Museum, Smart Tourism

1. INTRODUCTION

The Internet of Things (IoT) (Atzori, Iera, & Morabito, 2010; Mainetti, Patrono, & Vilei, 2011) is leading to the development of a plethora of smart objects that are intended to sense environmental parameters and human beings' behavior, in order to provide advanced services to the users. This trend aims at realizing smart environments able to capture, in a pervasive way, all useful information from the real world, and to automatically anticipate users' needs (Mainetti, Mighali, Patrono, Rametta & Oliva, 2013; Mainetti, Mighali, Patrono, & Rametta, 2014). The continuous attention towards this new vision puts an extraordinary stress on the so-called smart cities. They aim at increasing the effectiveness and efficiency of tourism, energy and water supply, healthcare, home and building automation, by integrating various systems through Information and Communication Technology (ICT) (Naphade, Banavar, Harrison, Paraszczak, & Morris, 2011). Unfortunately, the industrial

DOI: 10.4018/IJEGR.2016040103

world is still reluctant to invest directly on smart cities and most enterprises seek to secure their entrance with standardization and business models. To encourage the standardization of hardware and software solutions dedicated to smart cities, a smart city architecture should have a well-defined structure. This issue is well treated in (Anthopoulos, 2015), where the author uses literature findings and combines them with data from well-known smart cities to suggest a proper structure for a smart city architecture. The result is a five-layered architecture that could represent the baseline for smart city standardization. Referring to this structure, the IoT has well-defined position and contribution in the smart city architecture, since it belongs to the "ICT-based Hard Infrastructure Layer", which concerns all smart hardware able to provide sustainable smart city (SSC) services.

Among all the possible areas of applicability of ICT technologies, art and culture (identified by "Education & Tourism" in the aforementioned architecture) are becoming more and more interesting since they play an important role in human beings lives. Over the centuries, hundreds of museums and art galleries have preserved our diverse cultural heritage and served as important sources of education and learning. However, visiting a museum could often be quite boring because museum curators are not able to catch the tourists' attention properly. In particular, it is difficult to define in advance a tour for all the visitors, because interests may vary from person to person (e.g., from children to adults, students group from single visitor, casual visitor to fond-visitor). Given these issues, an important research question has to be answered: *can a smart tourism application effectively enhance customer experience within a museum?* To address such question, interactive and personalized museum tours need to be developed. In this perspective, a significant contribution can be given by the IoT, which aims to create a better world for people, where smart objects around us act accordingly like suggestions' generation that are relative to our insights.

Together with the IoT vision, also the role of mobile and wearable devices is increasing. Mobile devices, such as smartphones and tablets, are almost ubiquitous in our society, since they are not only communication means, but also technological tools for controlling other devices and communicating information about users (Alam, Saini, & El Saddik, 2015). For this reason, the trend to use them for interacting with smart environment is increasingly widespread. Finally, another key element of the current digital world is the Cloud, which is becoming the main mean for sharing data, information, and events between services and users.

Taking into account these considerations, it could be possible to design and develop a smart system able to improve the user experience in a museum. The fundamental concept of the system should be the location-awareness, that is, all the provided services have to act according to the users' position and their movements. The capability to detect the location of each user can represent the base of more complex systems that involve several IoT-enabling technologies. For example, by tracking the visitors in the museum, a service could provide external users with real-time information about queues, and the environment itself could modify its status according to specific events (e.g., the number of users in a room). Then, by placing Near Field Communication (NFC) totems in the internal areas, another service could automatically send a notification on visitors' smartphones about the opportunity to incrementally pay tickets to access specific rooms of the museum. Another interesting idea to improve the cultural experience of visitors could be to provide them with a smart wearable device that automatically recognizes the artworks they are admiring. By exploiting this opportunity, each user can receive personalized cultural contents on him/her mobile device and eventually share information in the Cloud, leading to the so-called Social Internet of Things (SIoT) (Militano, Nitti, Atzori, & Iera, 2015): vocal comments about his/her extemporaneous feelings, automatic posts on social networks with information about the particular artwork that is being admired, and so on.

Basing on these considerations and the current state-of-the-art, we have proposed an architecture able to provide the users with a real interactive cultural experience. In our system, which is the enhancement of the work presented in (Mighali et al., 2015), the user is basically equipped with a smartphone and a wearable device able to capture videos and images. The actual business logic is, instead, managed by several location-aware services running on a processing center. The wearable device accomplishes two main tasks: it recognizes the artwork in front of the user by processing several video frames, and continuously localizes the user by leveraging a Bluetooth Low Energy (BLE) infrastructure. The results of this twofold processing activity are sent to the processing center and then used by the location-aware services that are in charge to provide all the other features of the system. More in detail, they use a bidirectional communication channel with the visitor's smartphone to provide personalized cultural contents about his/her specific tour and information about the museum. They are also able to provide information to external users and to interact with other heterogeneous technologies that control the status of the environment (e.g., a building automation system that manages lighting and thermoregulation of the museum). Finally, the users may also exploit their mobile devices to share information and feelings in the Cloud among family and friends.

The painting recognition mechanism was evaluated by exploiting a set of challenging images from the Maramotti Museum of contemporary art of Reggio Emilia, Italy. Then, the whole architecture was functionally tested at MUST museum in Lecce, Italy.

The rest of the paper is organized as follows. Section 2 describes the current state-of-the-art related to smart environments and IoT solutions for museums. Section 3 provides an overview on the architecture design. A more detailed description of the system is presented in Section 4. Section 5 describes a first validation of the proposed system. Conclusions and future steps are summarized in Section 6.

2. BACKGROUND

Traditionally, museums are spaces that provide several information to visitors. In many cases, objects are accompanied by textual descriptions, which are often inadequate to satisfy the cultural interests of all visitors. Because of this, visitors' access to museum collections can be often unsatisfactory or not appealing, especially to new generations. Personalization through multimedia content is one answer to this problem (Karaman et al., 2014; Kovalenko, Mrabet, Schouten, & Sejdovic, 2015), since it offers visitors a customized presentation of appropriate information related to the visitor's tastes and preferences.

For these reasons, many solutions have been recently proposed for interactive user-profile-based guides, but none of them provides a flexible and scalable solution that can solve all the problems in one system. An example is the "SmartMuseum" (Kuusik, Roche, & Weis, 2009): by means of PDAs and RFID technology, a visitor can gather information about what the museum displays, building a customized visit based on his/her interests inserted on the museum website, before the visit. In (Sen, Roy, & Sarkar, 2014), authors proposed a system based on Active RFID that enhances the visiting experience through an audio system and pushbuttons. These approaches brought an interesting novelty when first released, but they have some very limiting flaws. Indeed, being tied to RFIDs does not allow reconfiguring the museum without rethinking the entire structure of the knowledge on which the project is based. In (Kuflik et al., 2011), authors proposed to customize museum experiences with machine learning techniques applied on the answers to questionnaires that the users should compile before the visit. In both proposals, the main flaw is the need to invasive interaction, asking the visitors to do something that probably they would not want to do. One of the valuable attempts for user profiling through wearable sensors was the "Museum Wearable" (Sparacino, 2002), a wearable computer that orchestrates an audiovisual narration as a function of the visitors' interests gathered from his/her physical path in the museum. However, this prototype does not use any visual analysis algorithms for understanding the surrounding environment. For instance, the estimation of the visitor's location is based on infrared sensors distributed in the museum space. This methodology allows only a rough estimation of the position and therefore the content delivery may not be well suited for the visitor. On a different note, the work presented in (Sartori et al., 2015) proposes a framework to automatically classify abstract paintings based on the emotional response they trigger in the visitors. Using statistical analysis and eye tracking devices, they calculate a score of each painting based on the emotions it provokes. Differently, our solution uses computer vision techniques to be adapted to different environments, providing personalized content without the need of explicit interaction with the user. Table 1 summarizes the key characteristics of the aforementioned smart museum projects.

Taking into account, more in general, recent solutions able to anticipate user's needs in a smart environment, authors in (Mainetti, Mighali, & Patrono, 2015a; Mainetti, Mighali, & Patrono, 2015b) propose two architectures able to overcome the heterogeneity of smart devices and to automatically manage the environment basing on user-defined rules and users' movements. The systems exploit an indoor location service based on Bluetooth Low Energy and provides a simplified development tool that allows even common users to develop new services to interact with smart objects.

About the indoor localization feature, in (Subbu & Thomas, 2014), the authors present a personalized smart control system that: (i) localizes the user exploiting the magnetic field of the smartphone, and (ii) controls appliances present at the user's location. In particular, the author employ common smartphones to measure the magnetic field in an indoor environment and effectively localize the user through his magnetic signature; given such localization, the lighting in the room is regulated as the user walks through it. Another example of location-aware services in a smart environment is reported in (Wang, Zixue, Jing, Yota, & Zhou, 2012). Here, the authors propose a system that collects information from the environment and then provides services to improve the lifestyle of the users, mainly from an energetic point of view. The last example is provided in (Moreno, Hernandez, & Skarmeta, 2014). In this case, the authors propose an access control mechanism, which consists of an engine embedded into smart objects able to make authorization decisions by considering both user location data and access credentials. User location data are estimated using magnetic field measured and sent by the phone.

3. SYSTEM ARCHITECTURE DESIGN

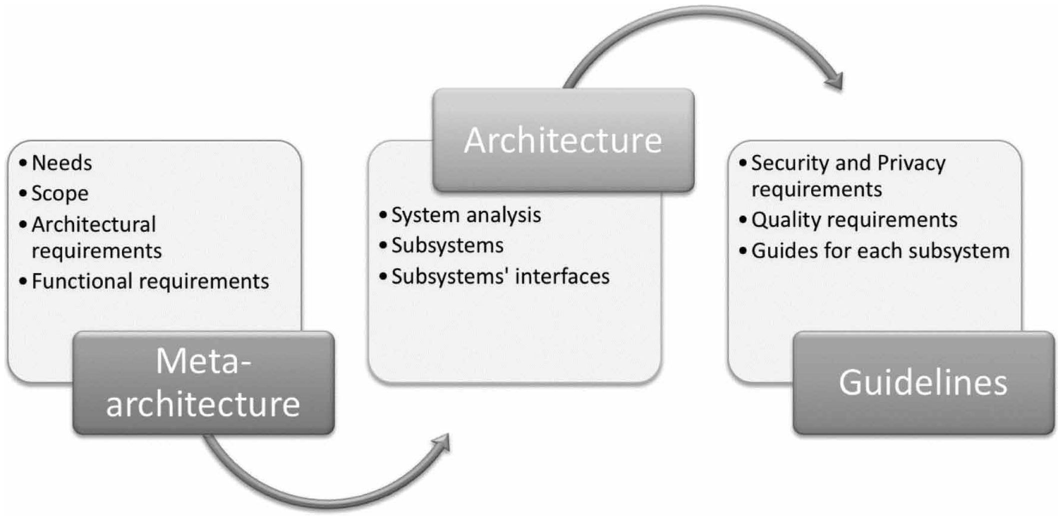
The designed methodology followed to develop the proposed architecture is substantially inspired to the guidelines provided in (Anthopoulos, Gemma, Fathy, & Sang, 2015) and described by Figure 1. The main steps of the process are the following:

- **Identification of needs:** This step concerns the identification of the current weaknesses in the considered city sectors. Specifically, for the cultural field, the status quo of the most city is characterized by a poor interest in the cultural heritage, which is seen as outdated and so suitable only for a restricted category of users. So, the main need is to increase the interest in the cultural heritage through the ICT technologies;
- **Scope definition:** This step concerns the identification of potential city services whose improvement could satisfy the identified needs. According to the need mentioned in the previous point, the museum environment represents the best choice since it is perfectly suitable for a deep technological enhancement;

Table 1. Comparison between the proposed architecture and similar systems presented in literature

Reference Work	Services Provided			
	Audio Information	Personalized Information	Interaction-Less Requests	Location Awareness
Kuusik, Roche, & Weis, 2009	√	√		√
Sparacino, 2002		√	√	√
Kuflik et al., 2011	√	√		√
Proposed	√	√	√	√

Figure 1. The design process

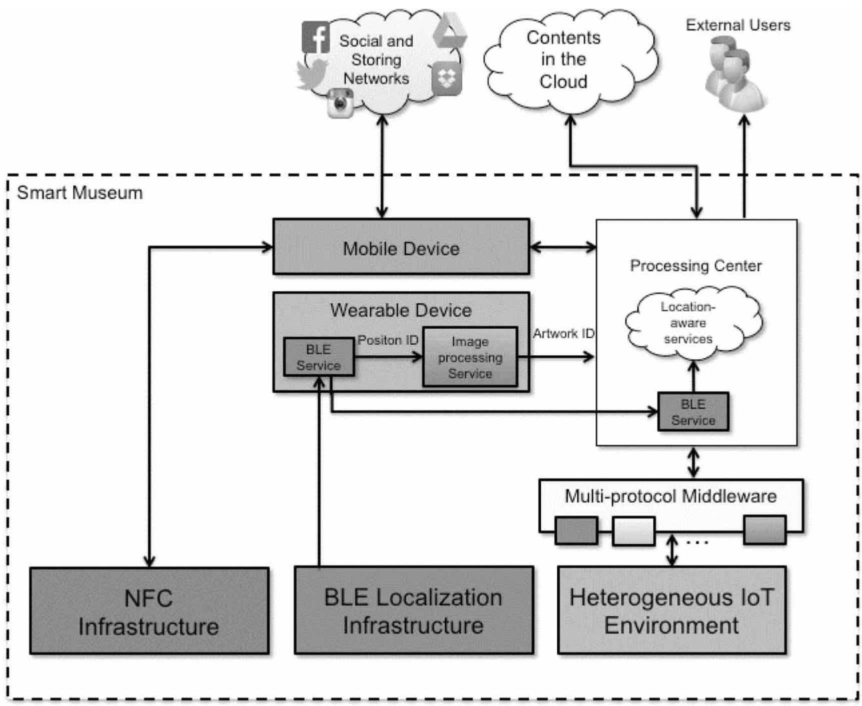


- **Functional requirements' definition:** This step concerns a formal description of the functional requirements of the system. In the specific case of the museum, the system has to improve the satisfaction during the visit by both anticipating user's needs and giving him a personalized cultural experience. Moreover, it has to be no-invasive and as low-cost as possible. In accordance to these requirements, some architectural subsystems have been defined, namely (i) an image processing algorithm running on a wearable device, (ii) a localization service, (iii) a processing center, and (iv) a mobile application. Each of them have to satisfy further requirements. The image processing algorithm has to be as fast and accurate as possible since the real-time content delivery is essential to keep high the user experience. The localization service should be not invasive and it has to identify the exact room in which the user is located. The processing center has to deal with the core business logic, since it has to manage both the museum environment and the interaction with external services. Furthermore, its structure should be as modular as possible to both easily extend it to other technologies and add new functionalities. Finally, the mobile application should be user-friendly in order to be used by any kind of user. Of course, all the subsystems should cooperate autonomously, minimizing the need of explicit user intervention;
- **Subsystem and corresponding interfaces' definition:** This step aims at describing how the subsystems are interconnected. During this phase, all the interfaces among the system services have been defined, above all: (i) the interface between the localization service and the image processing algorithm, (ii) the interfaces between the wearable device and the processing center, (iii) the interface that interconnects the services on the processing center, and (iv) the interface between these services and a multi-protocol middleware that control the museum environment;
- **Information security and privacy requirements' definition:** This step involves security and privacy issues related to the data exchanged among system services, but it is still under development.

The overall logical structure resulting from this design process is shown in Figure 2. It is composed of the previously mentioned building blocks:

- **The image processing algorithm:** It runs on the wearable device and it is able to detect, in real-time, the artwork the user is observing. It has the capability to quickly analyze the video frames

Figure 2. Overall logical system architecture



captured by the vision device and to identify the target object with high accuracy and reliability. The result of the processing activity is then sent to the processing center;

- **The localization service:** It is distributed between the wearable device and the processing center. The first one detects the current user's position and communicates it to the processing center. Here, the localization information is stored and made available to other services. The information is also used locally (on the wearable device) to speed up the image-processing algorithm;
- **The processing center:** It is the core of the business logic. It allows both the execution of the system services and their mutual interactions. First of all, it stores the current position of the users and provides this information to all interested services. For example, a specific service exploits this feature for providing external users with information about queues in the museum for accessing to specific artworks or sectors. Then, other services use the localization capability to modify the environment accordingly, so immersing the users in a real interactive environment. Finally, another key service receives the information about the artwork the user is observing and accesses, even exploiting the Cloud, the related cultural contents. Then, it provides such contents on the user's smartphone personalized them according to the user's profile;
- **The Mobile Application:** It allows the visitor to receive customized cultural contents about his/her specific tour, and to share multimedia information and personal feelings both in the Cloud and on the social networks. By exploiting recent auto-identification technology, such as NFC, it can also be used to pay "on the fly" incremental tickets during the cultural tour.

On this logical architecture, a physical architecture has been realized. In more detail, the exploited hardware was the following: a wearable device, a server that acts as a processing center, an infrastructure that is able to ensure the indoor localization service, an Android mobile device, and some smart objects to perform environment status changes.

The wearable device is realized through an embedded computer, namely an Odroid-XU (Odroid-XU, 2015). It is a single ARM board measuring just 94x70x18 mm, it has a Samsung Exynos5 Octa Core processor (Cortex-A15) of 1.2 GHz with PowerVR SGX544 graphics card MP3 and it is equipped with 2 GB of DDR3 RAM. The image acquisition is performed through an Odroid USB-cam 720p, which has a resolution of 1280*720 HD, an USB 2.0 plug-and-play interface and supports up to 30 fps. Figure 3 shows the hardware used to realize the wearable device.

The BLE infrastructure was realized through a set of Raspberry Pi boards (Raspberry Pi, 2015). Each of them is equipped with 512 MB of RAM and is based on the Broadcom BCM2835 system with a chip (SoC) that includes an ARM1176JZF-S 700 MHz processor and a VideoCore IV GPU. Furthermore, the system has MicroSD sockets for boot media and persistent storage. To provide each Raspberry Pi with beacon functionalities, each of them was equipped with both a BLE USB dongle and the BlueZ stack (BlueZ, 2014), that is, a software tool that provide the device with the ability to interact with the Bluetooth interface according to the BLE standard.

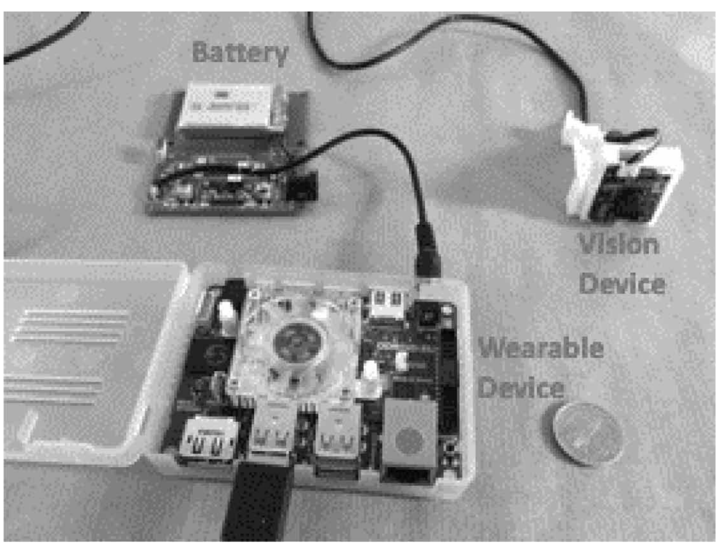
Regarding the Android mobile device, a model equipped with an operating system version greater than 4.3 was chosen to ensure the compatibility with BLE standard. Finally, the status of the environment was controlled through two kind of smart objects, namely KNX-compliant devices and CoAP-based embedded boards.

4. RESEARCH METHODOLOGIES

4.1. Localization

Several system components depend on the localization service. It consists of three main elements: (i) an infrastructure of wireless landmarks, called beacons, that periodically send localization information, (ii) a service installed on the wearable device that collects the information of the beacons to determine its location, and (iii) the service running on the processing center that receives the location of the user and provides it to the other services. More specifically, the network of beacons consists of embedded devices equipped with BLE interface and placed individually in the different rooms of the building. The choice of BLE is mainly due to its low energy consumption in front of a communication range comparable with that of the traditional Bluetooth. In this way, the wireless landmarks can be battery

Figure 3. The components of the wearable device



powered, making the localization mechanism more flexible and less invasive. Each device of the BLE infrastructure sends its location indication together with the Received Signal Strength Indication (RSSI) value. The service running on the user's wearable device collects location information from all the beacons within its listening range and then determines the room in which it is located. To do so, it computes a proximity index *d*, for each beacon, using the corresponding value of the RSSI. The adopted formula is the following:

$$\mathrm{RSSI} = -(10n\log_{10}d + A) \tag{1}$$

where *A* is the received signal strength at 1 m, *n* is a signal propagation constant depending mainly on the environment, and *d* is the distance from the sender (Lau & Chung, 2007). The landmark that has the lowest value of *d* represents the user's current location.

4.2. Artwork Recognition

To recognize the artwork in front of the visitor, an algorithm based on the image retrieval paradigm was designed. One of the main challenges that the method must face is the heavy distortion of artwork that is due to perspective. In fact, this situation prevents the use of template matching techniques that would otherwise be a useful tool to recognize paintings. To overcome this issue and robustly match an artwork captured by a wearable camera during a visit against its counterpart in the museum database, we propose to use local feature descriptors. In particular, we extract Scale Invariant Feature Transform (SIFT) local descriptors from the whole image and do not perform any attempt to extract features only from a detected artwork. SIFT keypoints are extracted from the regions with high discriminative capabilities, and their descriptor is robust to significant geometrical transformations such as rotation, translation and skew. We refer the reader to (Lowe, 1999) for further information on this local descriptor. The reason behind this choice follows the fact that modern museums often display very heterogeneous artworks and a generic detector cannot be employed. Furthermore, the shape of the artwork is a poor cue that can easily lead to false positives, preventing the method from sampling key points from a region of interest. Once key points and their descriptors have been computed, they must be matched against the museum database in order to find the most similar artwork. To achieve the best possible matching performances, after an initial match, the RANSAC (Fischler & Bolles, 1981) algorithm is employed for outlier rejection. In particular, his algorithm, estimates geometrical transformations among two sets of matched keypoints, iteratively generating hypothesis and discarding the keypoints that do not fit the best transformation found. Since due to the peculiarity of the setting, this may not be enough, to further improve the matching quality a threshold θ_S is applied over the distances among SIFT matches. This threshold, which controls the recognition performance, discards matches deemed correct by RANSAC but that have a high distance between the two keypoints in the pair (see Figure 4 for some examples of the matching process). As stated earlier, the method sample keypoints from the entire image instead of using a detector. This introduces the need for a second thresholding step that distinguishes between frames effectively containing an artwork and frames that do not. In fact, if many keypoints are sampled on architectonical details of the museum, the ratio between the amount of keypoints that survived the two previous pruning steps and the original amount of keypoints will be very small. We define the threshold θ_D to recognize this situation and react accordingly, since extracting the most similar painting from the database in such situation would lead to false positives and worsen the method performances. Furthermore, adjusting this threshold allows us to tune the method to the specific museum, being either more robust to noise or having a broader detection range. Algorithm 1, shown in Figure 5, provides a schematization of the proposed painting recognition approach.

Figure 4. Examples of artwork recognition

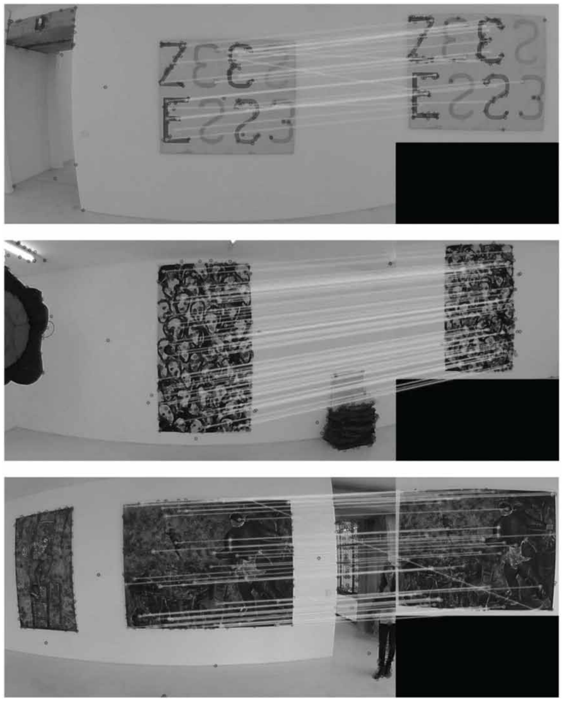


Figure 5. Painting recognition algorithm

Algorithm 1: Painting recognition

input : Current frame, template database
output: Painting identifier

Compute current frame keypoints and local descriptors;;
for *each painting template* **do**
  read SIFT descriptors;
  calculate matching keypoints;
  apply RANSAC algorithm to discard outliers;
  remove matches with distance greater than θ_S;
  compute ratio between remaining matches and total keypoints of the current frame;

Extract the painting with the highest $\frac{matches}{keypoints}$ ratio;
if $\frac{matches_{max}}{keypoints_{max}} > \theta_d$ **then**
  return
Recognized painting identifier;
else
  return
No painting detected

4.3. Mobile Services

Mobile devices, such as smartphones and tablets, represent key elements of the proposed system. More in detail, a specific application interacts with both the services running on the processing center and the user's world in the Cloud. By exploiting the first kind of interaction, the mobile device is able to display the cultural contents related to the specific artwork the user is observing. These contents could be textual information or multimedia data (e.g., audio, video, images) and are customized according to the user's profile. Moreover, exploiting the information about the current user's position, additional services are provided on the mobile device. In more detail, a notification is sent to the user every time a NFC totem is available in that location to pay an additional fee for visiting other rooms of the museum. Of course, the application provides this kind of notification only if the mobile device is equipped with the NFC technology.

Furthermore, since mobile devices are nowadays the main means to connect people, the application is also able to share data and events related to the cultural experience of the user. In particular, it allows to post updates on the user's social networks, and to store multimedia contents in the Cloud. In this way, each user could live again his/her extemporaneous feelings at a later stage. In order to benefit from social sharing, the user has to login on the mobile application by exploiting the OAuth 2.0 authentication standard (Hardt, 2012). It allows the user to access our system through his/her social profiles (Google, Facebook, Twitter). In this way, when a meaningful event is detected at the processing center, it is notified on the user's smartphone, so that the user can decide whether to share it or not.

Regarding the Cloud storing, it is realized exploiting the Amazon Simple Storage Service (Amazon S3) (Amazon S3, 2015), which provides proper APIs for several client platforms. So, the mobile device becomes an Amazon S3 client that stores the multimedia data in the Cloud and generates the REST URIs to get these files back from the Web. Each user automatically creates his/her storing space immediately after the login procedure.

In order to provide all personalized services to users (i.e., customized cultural contents, social sharing, Cloud storing), a preliminary profiling procedure is performed. In more detail, the mobile application asks the user to compile a short questionnaire upon installation. The results arising from this questionnaire are combined with the user's social profile retrieved through OAuth 2.0 authentication, in order to create a user's *cultural profile*. Finally, when the wearable device is provided to the user, s/he has to scan a QR code attached to the device through his/her smartphone, in order to associate the wearable device to his/her cultural profile. The connection between wearable device's features (i.e., localization and image recognition capabilities) and smartphone services (i.e., interaction with social networks and Cloud storing) allows to provide the personalized services.

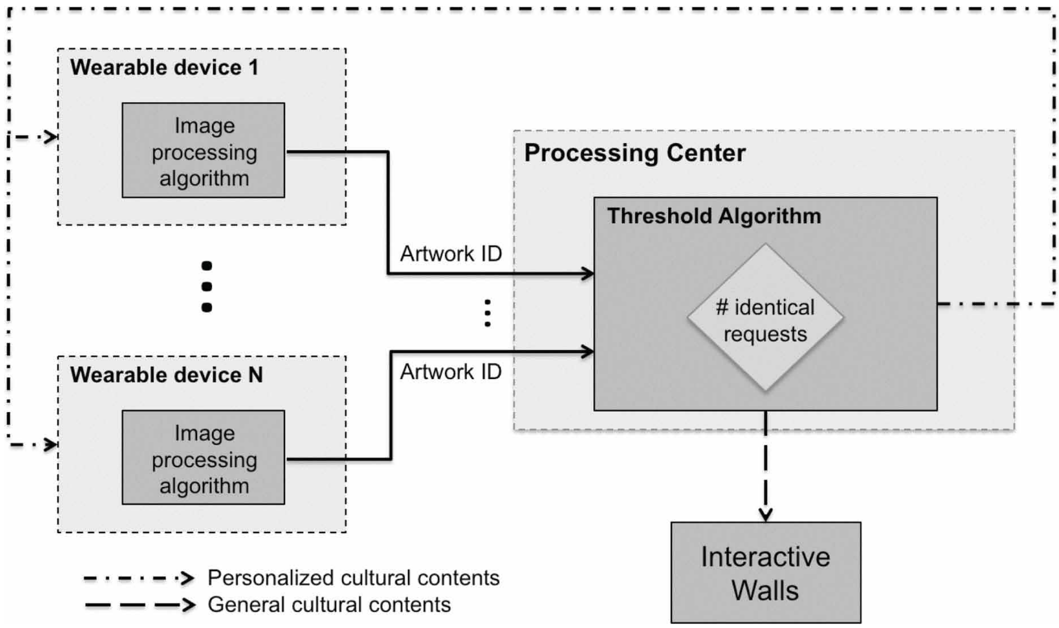
4.4. Cultural Contents Delivery and Cloud Services

Storing, organizing, and retrieving information and multimedia contents for each user is an expensive process from both the computational and memory point of view. For this reason, the Cloud seems to represent the solution that best suits this kind of needs, as its storing and computing capabilities allow to process data more efficiently. In particular, in the proposed system, the Cloud is accessed by the processing center whenever the running services need to retrieve cultural contents destined to the users.

The output obtained by the image processing algorithm (Figure 6), i.e. the unique identifier of the observed artwork, represents the key information for accessing the desired cultural contents. The wearable device sends this information to the processing center through the local WiFi network. There, a specific service is in charge of receiving all requests coming from users and analyzing them to start the proper procedure. More in detail, the interpretation of the artwork identifier can lead to two possible results:

1. Personalized multimedia cultural contents (Audio, Video and Text) related to the artwork on the user's mobile device;
2. Multimedia cultural contents on interactive walls of the museum.

Figure 6. Cultural contents delivery process



In the first case, the processing center provides the interested clients with the multimedia contents related to the observed artwork. Since these contents are sent to the user's personal mobile device, they can be customized according to his/her profile. In the second case, the processing center exploits the wired local networks to send multimedia contents to interactive displays and totems in the museum, so that the same cultural information is simultaneously available to all the involved users. The processing center takes the decision depending on a threshold algorithm that constantly monitors the number of simultaneous visitors looking the same artwork. More in detail, when the number of visitors is below a predetermined threshold, the algorithm converges to the first alternative. On the contrary, if there is a large group of visitors sharing the same "visiting experience", the algorithm converges to the second alternative.

Finally, the architecture exposes in the Cloud another useful service that provides statistical information about the busyness of the museum. Indeed, by exploiting the localization information, this service always knows how many visitors are moving in the museum and where they are. Therefore, this service can be used by external users to know in advance the length of queues in specific areas of the museum or which are the most admired artworks. Moreover, the information provided by this service could also be exploited by the museum supervisors to schedule partial maintenance works or to reorganize the internal spaces.

4.5. Interactions with the IoT

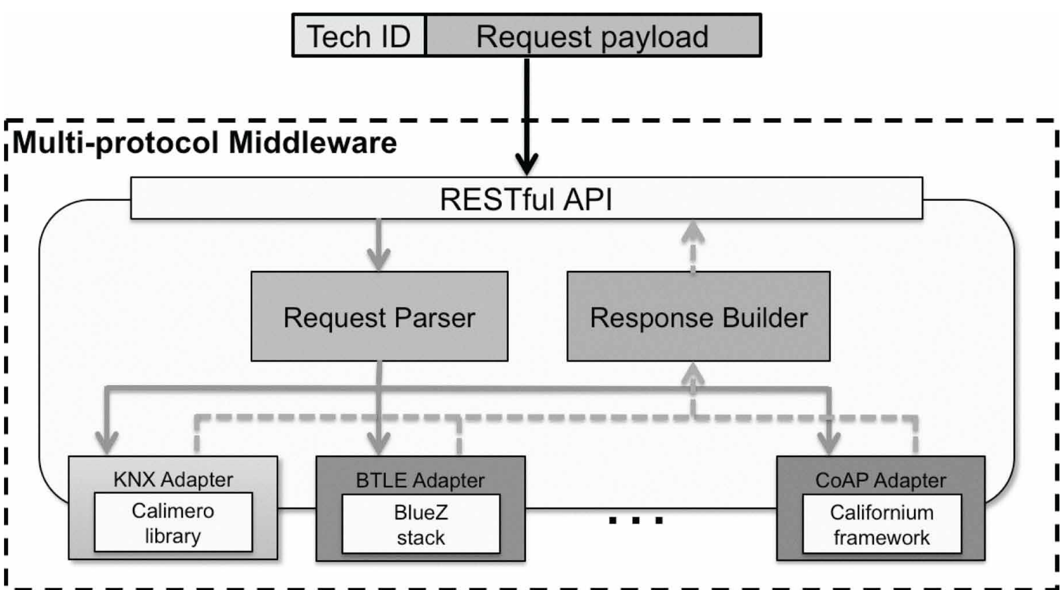
One of the main tasks of the services running on the processing center is to adapt the status of the environment according to the user profile and to the information coming from localization service. Exploiting IoT-aware technologies, the environment could be modified in real-time in order to provide the user with a real interactive experience. As an example, imagine that the museum has a special room where an historical war is represented by a mechanical animation managed by several IoT actuators. To maximize the impact of this animation, the system could decide to activate it only when the number of visitors in the room is higher than a predefined threshold. Furthermore, the system could reproduce different animations according to the user profile of the visitors present in the room. In the same way, lighting, temperature and other physical characteristics of a room could be controlled to perform special 4D effects. Obviously, the IoT technologies able to provide these features could be extremely heterogeneous since they are often compliant to different standards and protocols. In order to efficiently interact with such a kind of technologies, the services of the processing center exploit a multi-protocol middleware that allows a transparent access to the underlying heterogeneous devices, hiding the low-level communication details. In particular, on the one hand, it provides the services with high-level RESTful (Richardson, & Ruby, 2007) APIs to communicate with the physical network, whereas, on the other hand, it is equipped with specific software modules, called adapters, which communicate with the IoT devices in accordance to specific standards and protocols. (Figure 7). The modular structure that characterizes the middleware allows to easily extend it to new technologies, so guaranteeing flexibility and scalability. The key element that allows multi-protocol middleware to hide the underlying complexity is a "technology ID" embedded in the REST requests. Exploiting this ID, the request parser is able to identify the proper adapter to call. Actually, the middleware is equipped with three main adapters, which allow interacting with Costrained Application Protocol (CoAP) (Shelby, Hartke, & Bormann, 2014), KNX (KNX, 2014) and BLE compliant devices. The initial choice of KNX, CoAP, and BLE is due to their diffusion in both commercial and academic solutions available in the literature. KNX is the worldwide standard for home and building control, CoAP is one of the most used application protocol in the IoT, which provides a lightweight access to physical resources, and BLE is more and more the leading technology in commercial smart devices. However, any other IoT technology can be integrated into the system.

5. EXPERIMENTAL RESULTS

The evaluation phase was divided into two steps. The first one was dedicated to analyze the performance of the image recognition mechanism, independently of the rest of the architecture. On the other hand, the second step aimed at functionally evaluating the whole system in a real scenario.

Regarding the first step, in order to effectively test our painting recognition algorithm on real data, we employ the Maramotti Collection dataset. It is a completely unconstrained real-world

Figure 7. Multi-protocol middleware for interacting with heterogeneous IoT devices



museum dataset recorded at the Maramotti Museum of contemporary art of Reggio Emilia, Italy. This dataset features heterogeneous artworks such as paintings, sculptures and more, and different lighting conditions due to different museum rooms having different illumination requirements. The dataset contains more than 13000 frames at 960x540 resolution, acquired with a wearable device and annotated with the current visible painting. This scenario challenges our method in several ways: being a real museum, lighting conditions can vary greatly from one room to another. Furthermore, being acquired during a normal visit, artworks can be partially occluded or suffer from motion blur (see Figure 8).

In order to validate our choice of the SIFT descriptor, in Table 1, we evaluate several state of the art keypoint detectors and descriptors. We evaluate their results in terms of detection precision and recall and classification accuracy: this allows to loosely decouple the detection performance from the classification phase, effectively reflecting the fact that our method is based on two different steps. Furthermore, the best values of the employed thresholds are reported, along with the average per frame execution time of the algorithm. That is, each row of the table reports, for a given descriptor, the performance of the artwork detection (precision and recall), the accuracy that the method achieves in classifying it, i.e. the ratio of correctly recognized artworks. Furthermore, the subsequent columns respectively measure the average time required for the computation of the descriptor on a frame, and the adopted values of the two thresholds used in the method. This evaluation is indeed standard procedure in computer vision when analyzing the performance of an adopted local descriptor (Mikolajczyk & Schmid, 2005).

As Table 2 shows, the impact of different descriptors on the algorithm can be divided into performance and execution time. In fact, the best scoring techniques, such as SIFT, SURF and Maximally Stable Extremal Region (MSER)+SIFT, are among the slowest. This is due to these algorithms exploiting complex and robust keypoint detection techniques, along with high dimensional descriptors. Such robust methods result in perfect recall, very high precision and excellent recognition rates, but suffer from higher matching costs and higher execution times. The SURF average execution time results in being worse that SIFT due to the much higher number of detected keypoints.

Figure 8. Examples from the Maramotti collection dataset showing challenging situations such as motion blur, occlusions and non-painting artworks

Table 2. Comparative evaluation of several state of the art keypoint detectors and descriptors

Descriptor	Precision	Recall	Accuracy	Exec. Time	Theta_S	Theta_D
SIFT (Lowe, 1999)	0.871	1	0.964	0.689	230	0.4
SURF (Herbert, 2006)	0.946	1	0.907	0.756	0.3	0.35
ORB (Rublee, 2011)	0.867	0.979	0.060	0.502	300	0.05
MSER (Matas, 2004) +SIFT	0.975	1	0.928	0.704	230	0.04
SIFT+FREAK (Alahi. 2009)	0.789	0.464	0.088	0.626	300	0.03
BRISK (Leutenegger, 2011)	0.870	1	0.423	0.571	0.1	620

Faster techniques do exist, but come at the cost of performance: the ORB descriptors is indeed more than 30% faster than SURF and can achieve almost comparable detection performances. While this descriptor can effectively discriminate between artworks and the rest of the museum, it cannot correctly distinguish a painting from another, as testified by the very low accuracy.

Hybrid techniques are also possible, such as mixing the detection strengths of the MSER region detector, and the discriminative capabilities of the SIFT descriptor. This method performs similar to the standard SIFT approach, having a slightly higher precision due to the lower amount of keypoints, which reduces the number of false positives.

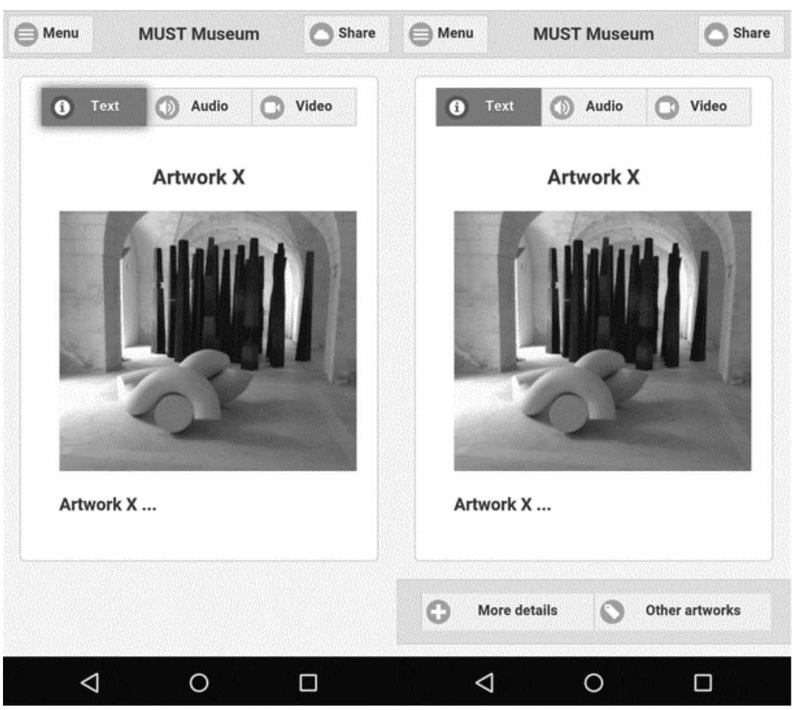
For what concerns the two thresholds, the best values reported in Table 2 are obtained via cross-validation. The different values needed by the various techniques are due to the different scales that the distance in the feature space can assume, e.g. the SURF distance is on a much smaller scale than SIFT.

Since the recognition process must be performed on a wearable device in real-time, frame rate subsampling is needed to reduce the computational effort the algorithm must take. Without loss in performance, the proposed technique analyzes one frame out of 25 (1 frame per second) resulting in the ability to recognize artworks at a grain of 1 second. Since the aim of the system is to provide an enhanced user experience, a 1s response time is more than enough to match the visitor movements inside the museum and respond to changes in the painting he is looking at.

For the second step of the validation phase, the painting recognition algorithm was integrated with the remaining components of the system in order to test the whole architecture in a real scenario. The initial system setup was carried out by deploying one beacon in each room in order to allow the wearable device to identify the exact room in which the user is located.

To functionally test the system, we chose two users with different cultural profile, namely “generic user” and “art-expert user”, and provided them with both the wearable device and a smartphone. As expected, when both users were in front of the same artwork, they received different cultural contents. As an example, Figure 9 shows two screenshots related to this circumstance. The first one refers to

Figure 9. Simple example about personalized contents on user’s smartphone



the "generic user profile", which guarantees the basic information concerning the observed artwork. The second screenshot refers to the "art-expert user", so the visitor can access to more information through further links, i.e. "more details" and "other artworks" of the same author.

Finally, the interaction with the environment was tested by defining two location-aware services that control lighting and music diffusion, respectively. In more detail, the first one was in charge to turn on the lights when at least one user was in the room, whereas the second one was able to start the music diffusion if at least two users were in the room. The lighting system was controlled by several CoAP-based embedded boards, whereas the music diffusion was managed by KNX-compliant devices. Of course, these services are only two simple example able to prove the effectiveness of the proposed architecture.

6. CONCLUSION

In this paper, we have presented an an IoT-based system for smart tourism services in a smart city and more specifically for smart museum personalized navigation. The proposed system presents a solution that allows to improve a visitor's cultural experience and answers the research question introduced in Section 1.

In fact, utilizing a BLE infrastructure and computer vision techniques, the tourist can obtain personalized information related to the artworks he is looking at without the need of explicit requests. The flexibility of the solution and the connection to the processing center allows the user to retrieve content from the Cloud and share his feelings with his friends through social networks. Moreover, the localization information can also be exploited by other services to adapt the environment to the users' movements and to notify, on the smartphone, the availability of further services, such as NFC micro-payments in specific museum areas.

To effectively answer the research question, we evaluate the proposed method in the real settings of the Maramotti Museum of contemporary art of Reggio Emilia, Italy and the MUST museum in Lecce, Italy. However, while being flexible and suited to answer to the needs of modern visitors, our system suffers from one main limitation. In particular, it is based on a specific hardware setting and therefore it involves additional hardware that must be provided by the museum itself.

ACKNOWLEDGMENT

This work is an extension of a work presented at AW4City 2015 workshop and is partially supported by the project "PON04a2_D - DICET LivingLab Di Cultura e Tecnologia – INMOTO – OR.C.HE.S.T.R.A.", a national project funded by the Italian Ministry of Education, University and Research (MIUR).

REFERENCES

Alam, K. M., Saini, M., & Saddik, A. E. (2015). Workload Model Based Dynamic Adaptation of Social Internet of Vehicles. *Sensors (Basel, Switzerland)*, *15*(9), 23262–23285. doi:10.3390/s150923262 PMID:26389905

Alexandre, A., Ortiz, R., & Vandergheynst, P. (2012). Freak: Fast retina keypoint. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).* IEEE.

Amazon S3. (2015). Retrieved from http://aws.amazon.com/it/s3/

Anthopoulos, L. (2015, September). Defining Smart City Architecture for Sustainability. *Proceedings of the 14th IFIP Electronic Government (EGOV) and 7th Electronic Participation (ePart) Conference*, Thessaloniki, Greece.

Anthopoulos L. G., Gemma P., Fathy R. A., & Sang Z. (2015). Setting the framework for an ICT architecture of a smart sustainable city. *International Telecommunication Union, Focus Group On Smart Sustainable Cities: Technical Specifications.*

Atzori, L., Iera, A., & Morabito, G. (2010). The internet of things: A survey. *Computer Networks*, *54*(15), 2787–2805. doi:10.1016/j.comnet.2010.05.010

Blue, Z. (2014), Retrieved from http://www.bluez.org

Fischler, M. A., & Bolles, R. C. (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography.Communications of the ACM, 24(6), 381-395.

Hardt, D. (2012). The OAuth 2.0 Authorization Framework RFC 6749.

Herbert, B., Tuytelaars, T., & Van Gool, L. (2006) Surf: Speeded up robust features. *Proceedings of theEuropean Conference on Computer vision* (pp. 404-417). Springer Berlin Heidelberg.

Karaman, S., Bagdanov, A., Landucci, L., D'Amico, G., Ferracani, A., Pezzatini, D., & Del Bimbo, A. (In press. (2014). Personalized multimedia content delivery on an interactive table by passive observation of museum visitors. *Multimedia Tools and Applications*.

Kovalenko, O., Mrabet, Y., Schouten, K., & Sejdovic, S. (2015). Linked Data in Action: Personalized Museum Tours on Mobile Devices. *Proceedings of theDevelopers Workshop, ESWC2015*.

Kuflik, T., Stock, O., Zancanaro, M., Gorfinkel, A., Jbara, S., Kats, S., & Kashtan, N. et al. (2011). A visitor's guide in an active museum: Presentations, communications, and refection. *Journal on Computing and Cultural Heritage*, *3*(3), 11. doi:10.1145/1921614.1921618

Kuusik, A., Roche, S., & Weis, F. (2009). Smartmuseum: Cultural content recommendation system for mobile users. *Proceedings of theFourth International Conference on Computer Sciences and Convergence Information Technology*, Tallinn, Estonia (pp. 477–482). doi:10.1109/ICCIT.2009.257

Lau, E., & Chung, W. Y. (2007). *Enhanced RSSI-Based Real-Time User Location Tracking System for Indoor and Outdoor Environments* (pp. 1213–1218). Gyeongju: Enhanced RSSI-Based Real-Time User Location Tracking System for Indoor and Outdoor Environments Convergence Information Technology. doi:10.1109/ICCIT.2007.253

Leutenegger, S., Margarita, C., & Roland, Y. S. (2011). BRISK: Binary robust invariant scalable keypoints. *Proceedings of the IEEE International Conference on Computer Vision (ICCV).* IEEE.

Lowe, D. G. (1999) Object recognition from local scale-invariant features. *Proceedings of the Seventh IEEE international conference on computer vision* (Vol. 2). IEEE.

Mainetti, L., Mighali, V., & Patrono, L. (2015a). An IoT-based User-centric Ecosystem for Heterogeneous Smart Home Environments. *Proceedings of theIEEE International Conference on Communications IEEE ICC '15*, London, UK. doi:10.1109/ICC.2015.7248404

Mainetti, L., Mighali, V., & Patrono, L. (2015b). A Location-aware Architecture for Heterogeneous Building Automation Systems. *Proceedings of the14th IFIP/IEEE Symposium on Integrated Network and Service Management IM '15,* Ottawa, Canada (pp. 1065-1070). doi:10.1109/INM.2015.7140434

Mainetti, L., Mighali, V., Patrono, L., & Rametta, P. (2014). Discovery and Mash-up of Physical Resources through a Web of Things Architecture. *Journal of Communications Software and Systems,* 10(2), 124-134.

Mainetti, L., Mighali, V., Patrono, L., Rametta, P., & Oliva, S. L. (2013). A novel architecture enabling the visual implementation of web of Things applications. *Proceedings of theInternational Conference on Software, Telecommunications and Computer NetworksSoftCOM '13* (pp. 1-7). doi:10.1109/SoftCOM.2013.6671847

Mainetti, L., Patrono, L., & Vilei, A. (2011). Evolution of wireless sensor networks towards the Internet of Things (Survey). *Proceedings of the 19th International Conference on Software, Telecommunications and Computer Networks, SoftCOM,* Split-Hvar-Dubrovnik, Croatia (pp. 16-21).

Matas, J., Chum, O., Urban, M., & Pajdla, T. (2004). Robust wide-baseline stereo from maximally stable extremal regions. *Image and Vision Computing*, *22*(10), 761–767. doi:10.1016/j.imavis.2004.02.006

Mighali, V., Del Fiore, G., Patrono, L., Mainetti, L., Alletto, S., Serra, G., & Cucchiara, R. (2015). Innovative IoT-aware Services for a Smart Museum. Web Applications for Smart Cities: bringing together Government, Businesses and Citizens, AW4City 2015 in conjunction with WWW'15 Conference (pp. 547-550). ACM 978-1-4503-3473-0/15/05. DOI: . Florence. Italy. doi:10.1145/2740908

Mikolajczyk, K., & Schmid, C. (2005). A performance evaluation of local descriptors. *IEEE Transactions on* Pattern Analysis and Machine Intelligence, *27*(10), 1615–1630.

Militano, L., Nitti, M., Atzori, L., & Iera, A. (2015). Using a Distributed Shapley-value Based Approach to Ensure Navigability in a Social Network of Smart Objects. *Proceedings of theIEEE International Conference on Communications IEEE ICC '15,* London, UK. doi:10.1109/ICC.2015.7248402

Moreno, M. V., Hernandez, J. L., & Skarmeta, A. F. (2014). A New Location-Aware Authorization Mechanism for Indoor Environments. *Proceedings of the28th International Conference on Advanced Information Networking and Applications,* Victoria (pp. 791 – 796). doi:10.1109/WAINA.2014.160

Naphade, M., Banavar, G., Harrison, C., Paraszczak, J., & Morris, R. (2011). Smarter cities and their innovation challenges. Computer, 44(6), 32–39.

Odroid-XU. (2015). Retrieved from http://www.hardkernel.com/main/products/prdt_info.php

Raspberry Pi. (2015). Retrieved from http://www.raspberrypi.org

Richardson, L., & Ruby, S. (2007). *RESTful web services*. O'Reilly Media.

Rublee, E., Rabaud, V., Konolige, K., & Bradski, G. (2011, November). ORB: an efficient alternative to SIFT or SURF. *Proceedings of theInternational conference on Computer Vision (ICCV)* (pp. 2564-2571). IEEE. doi:10.1109/ICCV.2011.6126544

Sartori, A., Yanulevskaya, V., Salah, A. A., Uijlings, J., Bruni, E., & Sebe, N. (2015). Affective Analysis of Professional and Amateur Abstract Paintings Using Statistical Analysis and Art Theory. ACM Transactions on Interactive Intelligent Systems, 5(2). doi:10.1145/2768209

Sen, S., Roy, S., & Sarkar, S. K. (2014). Proposal for Enhancing Museum Visiting Experience Implementing Active RFID Technology. *Proceedings of theInternational Conference on Advances in Computing & Communications*. doi:10.1109/ICACC.2014.76

Shelby, Z., Hartke, K., & Bormann, C. (2014). The Constrained Application Protocol (CoAP). RFC 7252.

Sparacino, F. (2002). *The museum wearable: real-time sensor-driven understanding of visitors' interests for personalized visually-augmented museum experiences* (pp. 17–20). Museums and the Web.

KNX Standard. (2014). Retrieved from http://www.knx.org/

Subbu, K. P., & Thomas, N. (2014). Poster abstract: A location aware personalized smart control system. *Proceedings of the13th International Symposium on Information Processing in Sensor Networks,* Berlin (pp. 309 – 310). doi:10.1109/IPSN.2014.6846777

Wang, J., Zixue, C., Jing, L., Yota, O., & Zhou, Y. (2012). A location-aware lifestyle improvement system to save energy in smart home. *Proceedings of the4th International Conference on Awareness Science and Technology,* Seoul, Korea (pp. 109 – 114). doi:10.1109/iCAwST.2012.6469598