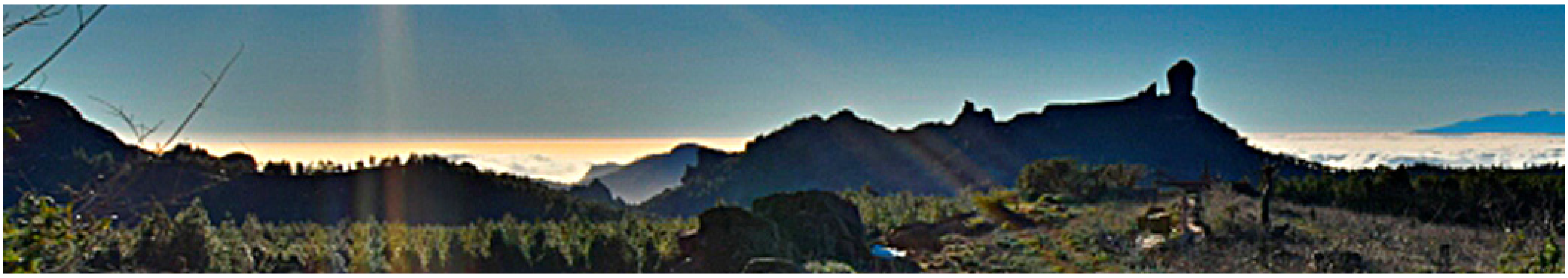
iberSPEECH2014

November 19-21 2014, Las Palmas de Gran Canaria



CONFERENCE PROCEEDINGS

iberSPEECH 2014

VIII Jornadas en Tecnologías del Habla

and IV Iberian SLTech Workshop

Escuela de Ingeniería en Telecomunicación y Electrónica
Universidad de Las Palmas de Gran Canaria
SPAIN

Organized by:

IDeTIC

División de Procesado Digital de Señal

RTTH

Spanish Thematic Network on Speech Technology (RTTH)

Supported by:

Universidad de Las Palmas de Gran Canaria

ULPGC 25 Aniversario 1989-2014

Selected papers published by:

Springer
Lecture Notes in Artificial Intelligence
Lecture Notes in Computer Science

LNAI Vol. 8854

PROCEEDINGS

iberSPEECH 2014

VII Jornadas en Tecnología del Habla
and
III Iberian SLTech Workshop

November 19-21, 2014
Escuela de Ingeniería enTelecomunicación y Electrónica
Universidad de Las Palmas de Gran Canaria
Las Palmas de Gran Canaria, Spain

Editors: Juan Luis Navarro Mesa
Pedro Quintana Morales
Alfonso Ortega
António Teixeira
Eduardo Hernández Pérez
Antonio Ravelo García
Iván Guerra Moreno
Manuel Medina Molina
Sofía Martín González

Organized by:

IDeTIC

División de Procesado Digital de Señal

RTTH

Spanish Thematic Network on Speech Technology (RTTH)

Supported by:

UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA

ULPGC 25 ANIVERSARIO 1989-2014

ISBN: 978-84-617-2862-6

http://iberspeech2014.ulpgc.es

The *Percepción Smart Campus* system

Carlos-D. Martínez-Hinarejos[1], Emilio Granell[1], David Rambla[2], Andrea Calia[2], Alejandro Luján[2], Guillermo Amat[3], Ángel Ramos[4], José-Miguel Benedí[1], and Alberto Sanchis[5]

[1] Pattern Recognition and Human Language Technology Research Center - Universitat Politècnica de València - Camino de Vera, s/n, 46022, Valencia, Spain
[2] Institute of New Imaging Technologies, Universitat Jaume I, Avda. Vicente Sos Baynat S/N, Castellón, Spain
[3] FHC25, Roger de Lauria 19 5-B, Valencia, Spain
[4] FHC25, Calle Rozabella, 4, Las Rozas, Madrid, Spain
[5] MLLP group, DSIC, Universitat Politècnica de València, Camino Vera s/n, 46022, Valencia, Spain

Abstract. This paper presents the capabilities of the *Smart Campus* system developed during the *Percepción* project. The *Smart Campus* system is able to locate the user of the application in a limited environment, including indoor location. The system is able to show routes and data (using virtual reality) on the different elements of the environment. Speech queries could be used to locate places and get routes and information on that places.

1 Overview of the system

The *Smart Campus* system, developed during the *Percepción* project, was defined to take advantage of the capabilities of mobile devices. Due to the nearly-universal of mobile devices such as smartphones, tablets, an so on, they can be used as a tool for getting information in a defined environment and locate interesting points. Moreover, the use of mobile devices could improve the productivity of the society by giving the users information that allow them save time and money, e.g., the fastest route to arrive to a given place, or the availability of a given resource (a book in the library, a free parking space, etc.).

These capabilities are extensively used in the concept of *Smart City*, i.e., a populated area that provides its inhabitants and managers with the infrastructure that allows them to obtain the maximum throughput of the available resources. In the *Percepción* project, the application is reduced to the environment of a university campus. Two systems were implemented, one for a medium-size university - *Universitat Jaume I* (UJI) of *Castelló* - and other for a small-size university - *Centro Universitario de Tecnología y Artes Digitales* (U-TAD). The UJI system comprises a total of 1997 relevant locations that can be required by the user. U-TAD system has a total of 234 locations.

The system would allow the user to employ its mobile device for the following activities:

– Obtain the current location, including indoor locations

– Show, by using Augmented Reality (AR), a route to arrive to a location.
– Show, by using AR, information on people (via *QR* codes), buildings, or other entities.
– Ask for a location by using speech.

In this system description, we will focus on the *SmartUJI Project*. This project consists of a smart-maps technology which integrates all the University data into a spatial information system. The main project's objective is to improve the monitoring and management of the university resources, and provide access to all of the university's data in a unified and homogeneous way as established in the *INSPIRE Directive of the European Parliament and of the Council* (2007/2/EC) [14, 1]. The platform is accessed through a web based application[6] and a set of integrated Android applications.

The following sections describe the main features of the different modules of the system: geographical location (Section 2), augmented reality based on geolocation and markers (Sections 3 and 4), and speech interface (Section 5). Section 6 describes possible features to be included in the systems in the future.

2 Positioning system

The *SmartUJI* Android applications use the positioning ability of mobile devices to provide the user location. Although the inclusion of *A-GPS* in Smartphones has solved the problem of outdoor positioning, performing indoor location is still an interesting research topic since *GPS* has no coverage inside buildings. Several technologies are being used to deal with the indoor location problem: Radio-Frequency Identification (RFID) [12], Bluetooth [7], Wireless Local Area Network [9, 10], ZigBee [11], Ultrasonic [4], Variations on magnetic field [3], Dead reckoning [5], or LED lights [6], among others. Nowadays, WiFi and magnetic field are two of the most promising technologies for indoor location due to the proliferation of *Smartphones* and the embedded WiFi connectivity and magnetic sensory system. In particular, a WiFi-based fingerprint *Indoor Positioning System* (*IPS*) based on *k-Nearest Neighbor* [2] has been integrated into the *SmartUJI platform* to support the *SmartUJI* applications.

The IPS was deployed as a new service in the *SmartUJI platform*. The mobile device has only to send the current fingerprint (a list of perceived WLAN intensities) and the service returns the current indoor location. Localisation errors are raised when the location cannot be estimated. An example of full location is shown in Figure 1, where the black arrow icon shows an estimation of the user's position (longitude, altitude and floor) inside an academic laboratory.

3 SmartUJI AR

SmartUJI AR is an augmented reality application that shows geographically located content using the *MetaioSDK*[7]. It allows the user to view the gathered data in an immersive and intuitive environment. As shown in Figure 2, it uses the camera to show which facilities or points of interest (*POI*) are near the user.

[6] `http://smart.uji.es`
[7] `http://www.metaio.com/sdk/`

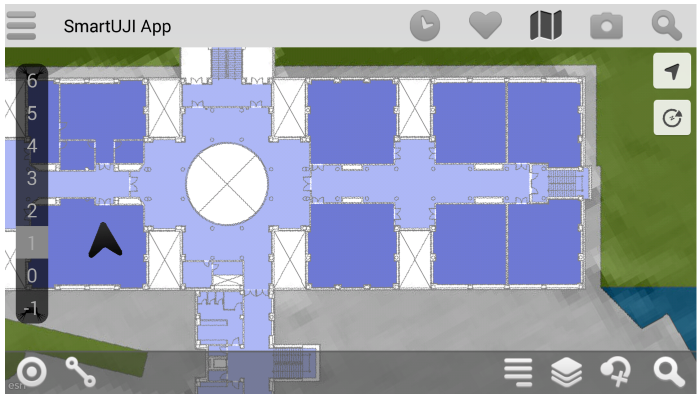


Fig. 1. Screenshot of the full positioning in a mobile device. The black arrow stands for the position where the user is located, this estimation is provided by the indoor positioning system.

SmartUJI AR collects all the data from the services, databases and geoprocesses provided by the *SmartUJI platform*. Then, these data are processed and properly showed to the user. The generated information is superimposed on the rear camera image in real-time. An appropriate icon is used to represent a certain facility or building in augmented reality. E.g., the *Library* (*Biblioteca*) is represented with an orange-white book icon in Figure 2. Then, the visible POI's are displayed using the user's orientation and current position via GPS or IPS (see Section 2). With this hybrid positioning system, GPS and IPS, it is possible to obtain the user's location anywhere inside the campus.

SmartUJI AR has a layer selector to choose what kind of POIs are going to be displayed. The main categories are: *Buildings*, *Restoration*, *Transports*, *Containers*, *Labo-*

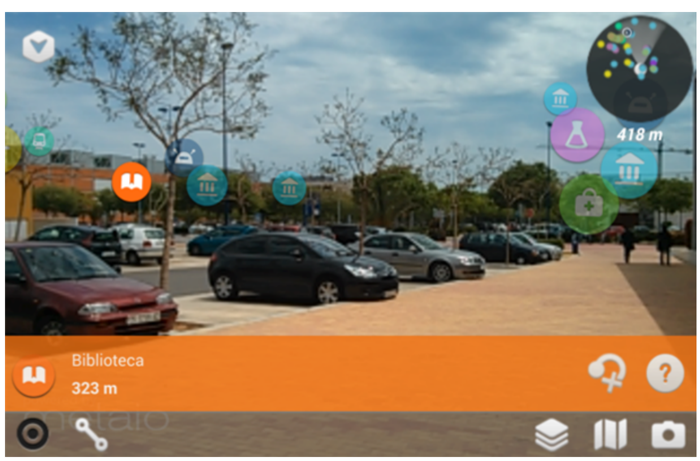


Fig. 2. Screenshot of the augmented reality application. It shows some facilities and the Library (*Biblioteca*) location is highlighted.

Fig. 3. Screenshot of the augmented reality application. It shows the path that the user should follow to reach its destination with white arrows.

ratories, and *Information Points*. Depending on the chosen categories, a different range of POI images will be displayed on the screen. Another functionality is the 2D radar in where all the POIs and the current vision field are represented. The user's vision fields moves at the same time as the user changes its orientation and position, so all the POIs that are inside the vision field are the same ones displayed on the screen.

The navigation functionality uses the *SmartUJI platform* to calculates a route between the user's current location and a POI. The obtained route is placed over the camera image as an 3D-arrows route. The user just has to follow the arrows to reach its destination as shown in Figure 3.

4 Marker based AR

This application is the piece of *SmartUJI* that focuses on providing useful information about faculties to the user. This information is provided through a navigable augmented reality interface. To support the AR behaviour, *Quick Response* (*QR*) codes have been used to work as a beacon. Furthermore, our approach relies on *QR* codes as AR markers as they can be also used to codify and store some information. In addition, the *QR* codes may contribute to geographical location whenever that information is read.

When the application starts, the device camera is opened and a simple interface is shown to the user. As can be seen in Figure 4(a), a target point in green indicates the user where to aim the camera to help the application find the marker and read it. Once the *QR* code is located and the information that was stored in it is decoded, a request is sent to the server in order to get the GUI description along with the matching faculty data. Meanwhile, a loading icon is displayed to show that work is in progress (Figure 4(b)).

The response from the server describes the GUI that the application has to parse and render. A main interface with navigation options and representing information about the faculty is what the server sends in first place, to see the rendered scene refer to Figure 5. Through this interface, the user is able to navigate to other levels of detail.

(a) (b)

Fig. 4. Marker based Augmented Reality examples.(a) Target screen, showed on start-up, (b) Loading screen, waiting for data.

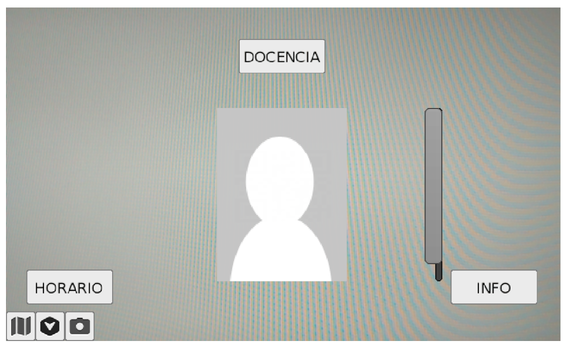
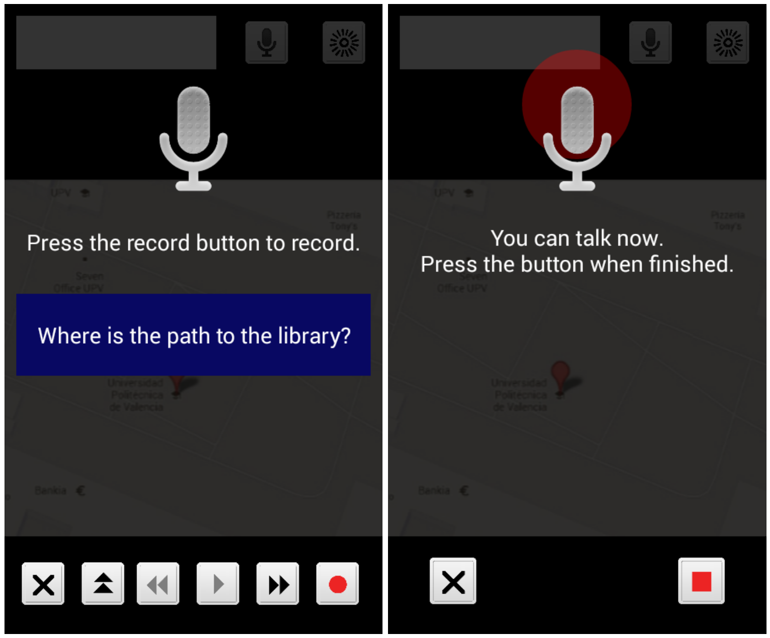


Fig. 5. Main screen, navigable menu, where users can get information about a university member: teaching hours and tutorials (HORARIO), courses (DOCENCIA), and more info (INFO).

5 Speech interface

The ASR system is used to locate points of interest in the university campus. The system is composed of a client application, a speech recogniser, a semantic analyser, a dialogue manager, an adaptation manager, a data manager and a database, with the following features:

- **Client application**: developed for Android 2.1 and higher versions; it is composed of two speech-related activities, one for speaker adaptation and the other for speech recognition.
- **Speech recogniser**: it is a version of the iATROS system [8] that supports TCP sockets audio input and speaker adaptation.
- **Semantic analyser**: FreeLing [13] was used.
- **Dialogue manager**: determines the current state of the dialogue; dialogue space is defined by a finite-state automaton; each state has associated a different language model for speech recognition.
- **Adaptation manager**: manages the speaker adaptation (generation of adaptation sentences and speaker adaptation based on the acquired audio signal).



(a) Speaker adaptation activity. (b) Speech acquisition activity.

Fig. 6. Speech interfaces.

- **Data manager**: connects clients with speech recogniser, provides system responses, and updates semantic models according to the user queries and the database.
- **Database**: the same than for the geographical part.

The system allows the speech recognition in three languages (Spanish, Valencian, and English). The language to be decoded depends on the client device configuration. Acoustic, language, lexical, and semantic models were generated for each language.

The speech interface is launched from the main client application by touching the proper icon. When launched, detects if the device/speaker is registered for the selected language. In case it is not registered, launches the speaker adaptation activity (Figure 6(a)). This activity shows the user the text to be uttered and allows to send the acquired audio signal when all the sentences were completed.

In any case, the speech activity is launched (Figure 6(b)). This activity starts with an initial dialogue state (i.e., no data was present and all possible queries are admissible). The activity follows a push-to-talk protocol that allows the user a complete control on the speech signal that is recorded. After finishing the recording, the audio signal is sent to the server and it is recognised and processed against the database. Three different situation may appear from the user perspective:

1. An only item was retrieved from the database: the corresponding identifier is returned to the client and the speech activity sends it to the geographical manager.

2. A small number (5 or less) of items was retrieved from the database: a list with the items is generated and sent to the client, that shows it and awaits for user selection; if the "None of the above" option is selected, it returns to speech acquisition.
3. A large number (more than 5) of items was retrieved from the database: returns to speech acquisition, showing an informative text (based on the new dialogue state) and the last recognition result.

In any case, internal dialogue state is updated according to the data the user gave and the database query results.

In case any error appears (server not available, sentence not recognised, etc.), the control is returned to the main client application.

6 Discussion and future directions

In this article we presented the complete *Percepción Smart Campus* system for the *SmartUJI* environment. The described system complements with the other Android applications of *SmartUJI*. For example, from anywhere in the application is possible to launch the other parts. Furthermore, this application can enhance the others parts providing a precise localisation of the user. This can improve the localisation algorithm because we know the position of the user based on the known position of the *QR* code the user is scanning. Also, the application integrates with the underlying operating system in various ways.

It is worth mentioning that all the applications, functionalities and modules have been designed to interact between them, to be fully interoperable within the *SmartUJI platform*, and to enhance the user's experience. Also, the applications are integrated with the underlying operating system in various ways. E.g., *SmartUJI applications* are able to call a university's member through the operating system's dialler application.

However, the system could benefit from other data sources that are not integrated at this moment in the whole application. For example, user identification can be used to know features on the user that could enhance the interaction in both location and speech recognition: if a student is inscribed in a given center or subject, it will be more usual for s/he that her/his location is on a given part of the campus, or that the uttered speech is on locations related to the user situation (e.g., asking for a specific classroom could be centered to the classrooms of the center where the student is inscribed). There are future plans to incorporate this data on the system and improve interaction, apart from using data pertaining to systems different to the *SmartUJI* system (e.g., data on social networks that allow to locate the user or knowing her/his preferences).

References

1. Benedito-Bordonau, M., Gargallo, D., Avariento, J., Sanchis, A., Gould, M., Huerta, J.: Uji smart campus: Un ejemplo de integración de recursos en la universitat jaume i de castelló. In: Actas de la IV Jornadas Ibéricas de Infraestructuras de Datos Espaciales (2013)
2. Cover, T., Hart, P.: Nearest neighbor pattern classification. IEEE Trans. Inf. Theor. 13(1), 21–27 (Sep 1967), http://dx.doi.org/10.1109/TIT.1967.1053964

3. Guo, Y., Chen, Y., Liu, J.: Indoor location estimation based on local magnetic field via hybrid learning. In: Extreme Learning Machines 2013: Algorithms and Applications, Adaptation, Learning, and Optimization, vol. 16, pp. 189–207. Springer International Publishing (2014)
4. Ijaz, F., Yang, H.K., Ahmad, A., Lee, C.: Indoor positioning: A review of indoor ultrasonic positioning systems. In: Advanced Communication Technology (ICACT), 2013 15th International Conference on. pp. 1146–1150 (Jan 2013)
5. Jin, Y., Toh, H.S., Soh, W.S., Wong, W.C.: A robust dead-reckoning pedestrian tracking system with low cost sensors. In: Pervasive Computing and Communications (PerCom), 2011 IEEE International Conference on. pp. 222–230. IEEE (2011)
6. Kuo, Y.S., Pannuto, P., Dutta, P.: Demo — luxapose: Indoor positioning with mobile phones and visible light. In: The 20th Annual International Conference on Mobile Computing and Networking. MobiCom '14 (September 2014)
7. Li, H.: Low-cost 3d bluetooth indoor positioning with least square. Wireless Personal Communications pp. 1–14 (2014), `http://dx.doi.org/10.1007/s11277-014-1820-1`
8. Luján-Mares, M., Tamarit, V., Alabau, V., Martínez-Hinarejos, C.D., i Gadea, M.P., Sanchis, A., Toselli, A.H.: iatros: A speech and handwritting recognition system. In: VJTH'2008. pp. 75–78 (2008)
9. Machaj, J., Brida, P., Pich, R.: Rank based fingerprinting algorithm for indoor positioning. In: Proceedings of the 2nd the International Conference on Indoor Positioning and Indoor Navigation (IPIN'2011). pp. 1–6 (2011)
10. Marques, N., Meneses, F., Moreira, A.: Combining similarity functions and majority rules for multi-building, multi-floor, wifi positioning. In: Proceedings of the 3th the International Conference on Indoor Positioning and Indoor Navigation (IPIN'2012) (2012)
11. Martí, J.V., Sales, J., Marín, R., Jiménez-Ruiz, E.: Localization of mobile sensors and actuators for intervention in low-visibility conditions: The zigbee fingerprinting approach. International Journal of Distributed Sensor Networks (2012)
12. Montaser, A., Moselhi, O.: RFID indoor location identification for construction projects. Automation in Construction 39(0), 167 – 179 (2014), `http://www.sciencedirect.com/science/article/pii/S092658051300109X`
13. Padró, L., Stanilovsky, E.: Freeling 3.0: Towards wider multilinguality. In: Chair), N.C.C., Choukri, K., Declerck, T., Doan, M.U., Maegaard, B., Mariani, J., Moreno, A., Odijk, J., Piperidis, S. (eds.) Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12). European Language Resources Association (ELRA), Istanbul, Turkey (may 2012)
14. Sanchis, A., Arnal, A., Molina, W., Sanchis, V., Daz, L., Huerta, J., Gould, M.: viscauji: campus inteligente como ide local. In: Actas de la III Jornadas Ibéricas de Infraestructuras de Datos Espaciales (2012)