Talking Digital Cities: Connecting Heterogeneous Digital Cities Via the Universal Mobile Interface

Tomoko Koda[1], Satoshi Nakazawa[2], Toru Ishida[3]

[1]JST CREST Digital City Project/Kyoto University, Kyoto Japan
koda@digitalcity.jst.go.jp
[2]Fuji Research Institute Corporation, Tokyo Japan
satoshi_nakazawa@fuji-ric.co.jp
[3] Department of Social Informatics, Kyoto University, Kyoto Japan
ishida@i.kyoto-u.ac.jp

Abstract. Digital Cities have been developed all over the world to provide regional city information. This paper proposes a new approach to integrate existing digital city systems via a universal mobile interface. The universal mobile interface integrates various media representations in connected digital city systems representing the same real city with audio semantic annotations. It is also applicable for a mobile use without developing a specific mobile interface. The demo system that connects three heterogeneous digital cities, namely, a 3D modeled virtual space, image-based virtual space, and map-based information space, successfully showed the connectivity of the universal interface. The current system provides the following functions: 1) a single and simple interface to navigate users into multiple digital cities, 2) a simultaneous walk-though into the connected digital cities, 3) audio descriptions of scenes using semantic annotation of objects in the connected digital cities, 4) prioritizing semantic annotation according to users' needs.

1 Introduction

Digital cities have been developed all over the world. Among them are efforts to make a virtual city where users can walkthrough, such as a 3D model of a real city, 2D map of a city, digital photo albums of a city, and so on [3]. As computers are becoming more powerful and cheaper, more digital representations of real cities will be developed. Moreover, the recent rapid growth of the digital camera market would increase the motivation of creating new digital cities.

However, digital cities that aim to provide a 'walkthrough' are developed in different ways, as no common design approach exists. At the same time, there is no common interface for users to walkthrough and manipulate those digital cities. An interface that would integrate heterogeneous digital cities representing the same city would be very helpful for users.

Digital cities are developed to provide regional information [3]. However, users also need information when not sitting in front of their computer. However, a precise 3D model of a real city cannot be displayed on a small PDA or a mobile phone. In

order to make full use of the existing digital cities, there is a need for a portable interface to information provided by digital cities.

The aim of creating the universal mobile interface is to connect various heterogeneous digital cities representing the same area. The universal mobile interface should enable a user to navigate through multiple digital cities by means of a single interface. The universal mobile interface has the following goals:

1. Single and simple interface: The universal mobile interface provides a single interface to multiple digital city systems so that a user can navigate oneself without learning specific commands to operate each system. The manipulation or movements within the city spaces should be intuitive and simple. Thus the universal mobile interface supports three input devices, namely, arrow keys, a joystick, and voice commands to move around the virtual spaces connected to the universal mobile interface.

2. Shared location: The universal mobile interface provides a simultaneous walkthrough experience of multiple digital cities. The universal mobile interface has its own coordinate system. The user's current location (coordinates) is converted to other coordinates that correspond to other virtual spaces. A mobile user with GPS can move around virtual spaces via the universal mobile interface. Detail description of the system is made in the later chapter.

3. Text based information: The universal mobile interface provides text-based information based on the user's current location by using text annotations attached to objects in the virtual spaces. The texts are converted to audio descriptions using text-to-speech. Users with vision problems or mobile users who do not have a display can in this way receive textual information from multiple virtual spaces.

4. Prioritizing location based text information: The universal mobile interface works as a filter to provide most useful text information for the user. Each virtual space has an initial priority value for giving text information to the user. The user can change each priority value according to his preferences.

2   Existing Digital City Systems

This section describes examples of existing digital cities that the universal mobile interface can integrate. Applicable digital cities should be organized in a spatial-geographical way, such as 3D model-based and 2D map-based cities.

2.1   3D Geometrical Models of a Real City

3D representations of a real city provide the most realistic model. One example is the Helsinki Arena 2000 developed by Helsinki Telephone Corporation. Helsinki Arena 2000 provides users an advanced web platform and services through a real 3D model of the city of Helsinki via broadband networks [9, 10]. Another example of a 3D model is FreeWalk, which was developed within the Universal Design of Digital City project [4]. Users enter into the FreeWalk virtual space via computer networks as avatars to navigate through the space and communicate with other avatars. Figure 1 shows a 3D VRML model of downtown Kyoto using FreeWalk.

Fig. 1. A 3D VRML model of Kyoto city using FreeWalk

Fig. 2. Image based virtual city of Kyoto using TownDigitizing

2.2   Image-Based Virtual Cities

Developing a 3D model of a real city requires time, money and expertise. Researchers have been investigating easier ways to build a digital city. One of those efforts is the use of digital photographs or digital photo realistic images for building a pseudo 3D digital city. One example of the above technologies is "3-D Photo Collage System" developed by the University of Tokyo [11]. The Photo Collage System provides non-experts a way to create a 3D space by associating 2D digital photos. Another example is "TownDigitizing" that uses omni-directional images instead of normal 2D images [7, 8]. Images are captured by an omni-directional camera that takes 360-degree images of the real city. The images are automatically fused based on the visitor's virtual location to reproduce what the user would see in a corresponding physical city. Figure 2 shows a TownDigitized version of downtown Kyoto. A next version of Town-Digitizing is currently under development in the Universal Design of the Digital City project. This version will capture life images with omni-directional

cameras installed inside a city, in order to create automatically an image based virtual space.

2.3   Map-Based Geographical Information Systems

Whereas 3D models and image-based digital cities try to represent a real city, map-based geographical information systems on the web put together various types of local information and organize it by spatial or geographic coordinates [13]. Numerous local GIS's have already been developed [12]. Among them are the Geographic Information System of Turin municipality [1], and KyotoSEARCH [6]. KyotoSEARCH extracts city related events and elements by analyzing web data. It calculates relations between these elements and events and describes them. Figure 3 shows a map of downtown Kyoto using KyotoSEARCH.
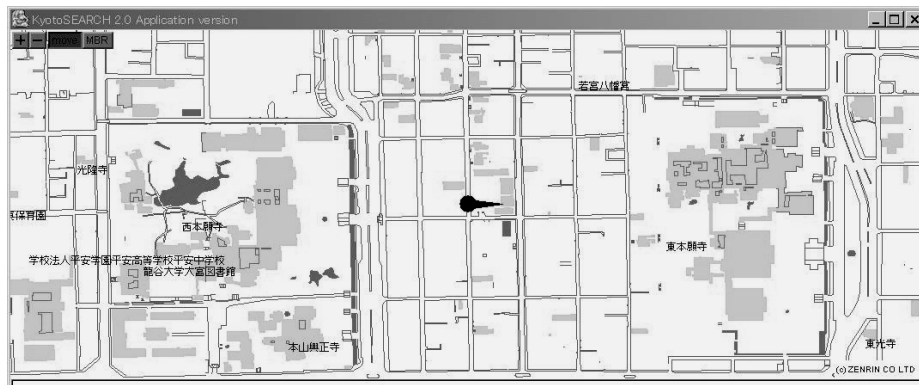

Fig. 3. Map of Kyoto city by KyotoSEARCH

2.4   Characteristics of Information from Digital Cities

The three types of virtual cities mentioned above provide specific information according to the nature of the technology used. The 3D model with FreeWalk is suitable for providing static and detailed information about buildings and shops, such as their size and color. However, developing a detailed 3D model is costly and time-consuming, and requires technical knowledge. Avoiding obstacles while walking through a 3D city with a joystick or cursor keys may be difficult for a novice user.

The image-based virtual city created with TownDigitizing can provide lively scenes of the real city at relatively low cost and in a short development period, and without high technical skills. Live cameras are useful for providing information on mobile entities of the real city, such as pedestrians and traffic on the road. However, walkable routes are limited to the places where photos of the real city are available. Users of an image-based virtual city cannot walk freely into the space.

The 3D model of a real city and the image-based virtual city are useful to provide detailed micro view of a user's location, while the map-based KyotoSEARCH tool can provide geographical and directional information through a macro view of the

user's location. However, a map-based system lacks visual images of the real city. Table 1 compares the characteristics of the three digital cities, and in section 3, we will discuss this in more detail. The universal mobile interface aims to provide combined city information by strengthening and complementing the characteristics of those virtual spaces each other.

Table 1. Comparison of characteristics of digital cities that use different spaces.

	3D virtual spaces	Image-based city spaces	Map-based city information spaces
Feature	Detailed 3D model	Photo-realistic images	Scalable 2D map
World Dimension	3D	2-2.5D	2D
User space dimension	2D	2D	2D
Movable space	Anywhere except obstacles (walls etc.)	Along pre-defined routes	Anywhere
Interactivity	Very high	High	Low
Macro view	High	Not provided	Very high
Micro view	Very high	High	N/A
Real-time information	Possible but expensive	Possible with live camera, relatively easy	Possible (GIS)
Development Cost	Very high	Relatively low	Low

2.5 Scenario of Using the Universal Mobile Interface

Typical scenarios of using the universal mobile interface are as follows:

Scenario for an immobile user: A senior person or computer novice is accessing a virtual city of Kyoto via the universal mobile interface. He does not know how to use a mouse or keyboard. He selects the location where he wants to start exploring the city by telling the system "Shijo-kawaramachi." He uses voice commands to move forward, turn right, stop and so on to navigate himself into the space. The system presents the 3D model, images and map of his current location, while explaining shop names, their specialties, etc. The 3D virtual space tells the user: "The brown-colored shop on your left sells tea utensils." The image-based system shows a live image of the shop. The map system adds: "The shop is located at 50m south from the Shijo-kawaramachi intersection." The universal mobile interface provides three different views and audio descriptions of the same place.

Scenario for a mobile user: A first-time visitor of downtown Kyoto is walking around in the city with a mobile phone. The mobile phone is equipped with GPS and can run Java applications on it. The mobile phone is connected to multiple digital cities of Kyoto via the universal mobile interface. As the user moves forward along a street, the universal mobile interface gives audio information on landmarks ahead, shops to be visited, which way the street is leading, etc. The 3D virtual space gives the user detailed information on each landmark, such as "The brown building on your right is a famous bakery." The map system gives directional information such as "If you keep going along this street to the west for 1.6 kilometers, there is another bakery on your right." The image-based virtual space gives lively information to the user, such as "There is a long line in front of the bakery ahead of you." The user can

change priority values that are set to each virtual space by giving feedback to each virtual space while it is giving audio description to the user, such as "That sounds interesting." (while the 3D virtual space is giving information on the bakery). Then the priority value of the 3D virtual space increases so that it can give information more often than other virtual spaces. The mechanism of achieving the above scenario is described in the following section.

3   System Description

This section gives detailed system description of each component that realizes the goals of the universal mobile interface described in the section 1.
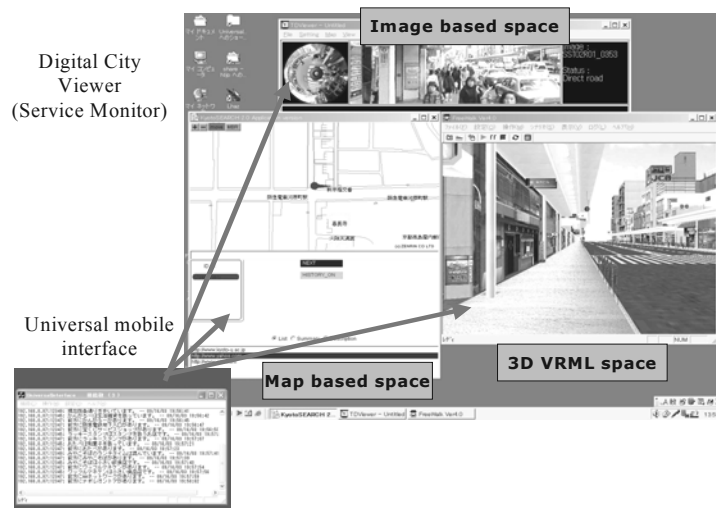

Fig. 4. The mobile interface window and service monitor, displaying 3 connected digital cities.

3.1   Single and Simple Interface

The universal mobile interface is a windows-based application that communicates with various digital city systems via TCP/IP. A user can connect to digital cities via the universal mobile interface by selecting a specific geographical point (i.e., an intersection, a department store, etc.), or by using a longitude and latitude of a geographical point. A user can move around the connected digital cities with arrow keys, joystick, or voice commands. Figure 4 shows the universal mobile interface window and service monitor that displays three connected digital cities, namely, an image-based virtual space, 3D VRML model, and 2D map of downtown Kyoto. All three digital cities display the same location.

3.2   Location and Movement Control – Shared Location

To achieve synchronous location and movement control among various digital cities, the universal mobile interface has its own coordinate system. When a user moves around the connected digital cities, the user actually moves within the coordinate system of the universal mobile interface. Figure 5 shows the diagram of location and movement control provided by the universal mobile interface.
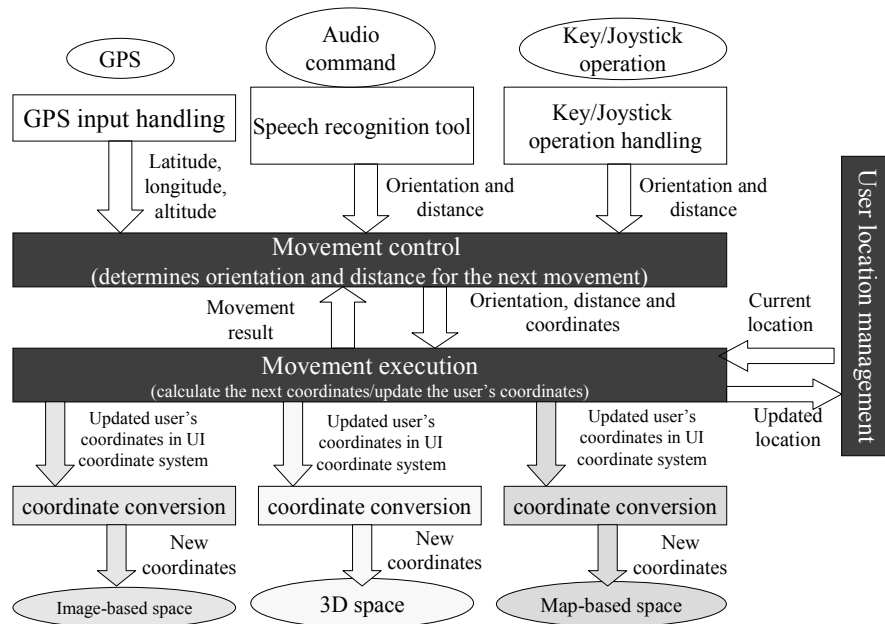

Fig. 5. Location and movement control between the interface and connected digital cities.

The three connected systems have their own coordinate systems. For example, the image-based system has a two dimensional coordinate system that uses relative position of each photographic image, while the map-based system has a two dimensional one (Table 1). Thus, a user's position is represented differently in each system. The universal mobile interface solves this technical problem in the following two ways.

Synchronizing Location and Orientation
The universal mobile interface has its own 3 dimensional coordinate system based on longitude, latitude, and altitude to achieve synchronous location and orientation control. The user's current location is represented by the universal mobile interface's coordinates and maintained by the universal mobile interface system. To each of the connected system we added a function that converts coordinates between these systems and the universal mobile interface. This function enables synchronization of the user's location and orientation by just sending the location and orientation data from the universal mobile interface to each connected system (Figure 5). Advantages

of having a standard and independent coordinate system in the universal mobile interface are:

- To be able to move into any location when a user is outside the area of either of the connected systems. It is rare that the connected systems have exactly the same area as a movable space.
- To be able to maintain the user's current location when one of the connected systems is not operating.

Synchronizing Movements

It is important for synchronizing movements to adjust a timing to send the user's location from the universal mobile interface to the connected systems. In the concrete, each connected system receives the user's location data, calculates the user's location based on its own coordinate system, changes the coordinates of user's location to the new ones, then sends a notification of a completion of the movement operation to the universal mobile interface. The universal mobile interface waits until it receives the notifications from all the connected systems.

The connected systems need to have a function that changes the user's location, orientation and movement in their own coordinate systems. There is a case that the connected system does not allow a user to move to any location. For example, the image-based system has fixed routes (Table 1) and the relative position of photo images are represented as points on a map. The universal mobile interface calculates the points on the map and selects the closest image from the user's location. In this way, the universal mobile interface provides a transparent interface in terms of movement and location control both for users and connected systems that have their own virtual spaces.
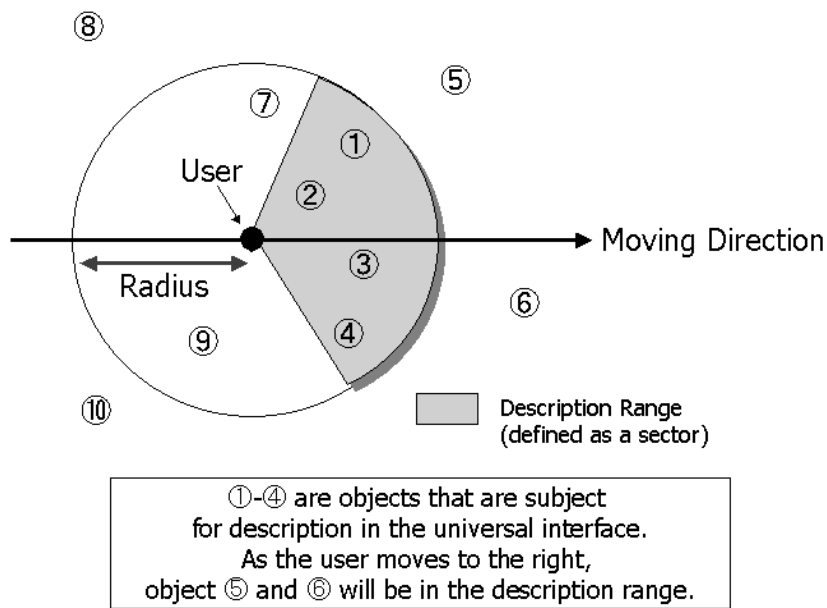

Fig. 6. The scope of objects described in the universal mobile interface

3.3 Semantic Annotation – Text-Based Information

The universal mobile interface provides text-based information based on the user's current location. This text-based description of the scenes uses semantic annotations attached to visual objects (i.e., streets, buildings, signs) in the connected virtual spaces.   The semantic annotations are converted to audio descriptions using text-to-speech. Users with vision problems or mobile users who do not have a display can in this way receive audio descriptions from multiple virtual spaces.

As the user moves around the connected digital cities, the universal mobile interface receives semantic annotations of each object seen in front of the user's point of view from the connected digital cities. As shown in figure 6, a description range is a sector where the user's location is in the center. A definition file defines the sector's radius and center angle. The closer objects are to the user, the higher priority they have to be described. Objects in the description range change as the user moves around the digital cities. The text description of the scene is provided as follows:

- Objects are classified with ontology-based definitions. Figure 7 shows the "buildings" class tree used in the universal mobile interface.
- Each class has its own attributes defined in an annotation definition file and inherits its upper classes' attributes. For example, a MacDonald's shop has attributes such as "color: yellow, category: hamburgers, number of seats: 65, seasonal menu: chicken, etc."
- The Universal mobile interface uses a template definition to generate textual descriptions of an object using its attributes. For example, the above MacDonald's can be described as "The Macdonald's shop has 65 seats. Its seasonal menu is chicken."

Using characteristics of connected virtual spaces: As described in 2.4, each connected virtual space has characteristic information. The universal mobile interface can generate various annotations by using a semantic annotation mechanism. For example, the 3D virtual space can have attributes of shop objects, such as the shape and color. The image-based virtual space can have attributes such as the popularity of the shop by using live images. And the map-based space can have attributes such as the exact location of the shop, its address and direction. In this way, the annotations described in the scenario in 2.5 can be achieved.

3.4 Prioritizing Location-Based Text Information

The connected digital city systems have their own semantic annotations to describe the same object, i.e., an 3D model based digital city would describe the object as a brown square building, while a map based system would describe it with its actual address. Each connected digital city system has its own priority value. A priority value controls a priority of each system to "speak", (in other words, its audio description of an object is presented to a user) about an object within the pre-defined scope. A user can change each system's priority value by using an audio command "No interest." or "Tell me more." while an audio description is made. Figure 8 shows a system component that generates descriptions of objects and priority control. As shown in the mobile user's scenario in 2.5, this prioritizing mechanism is especially necessary in a mobile scene. A mobile user may not want to listen to architectural
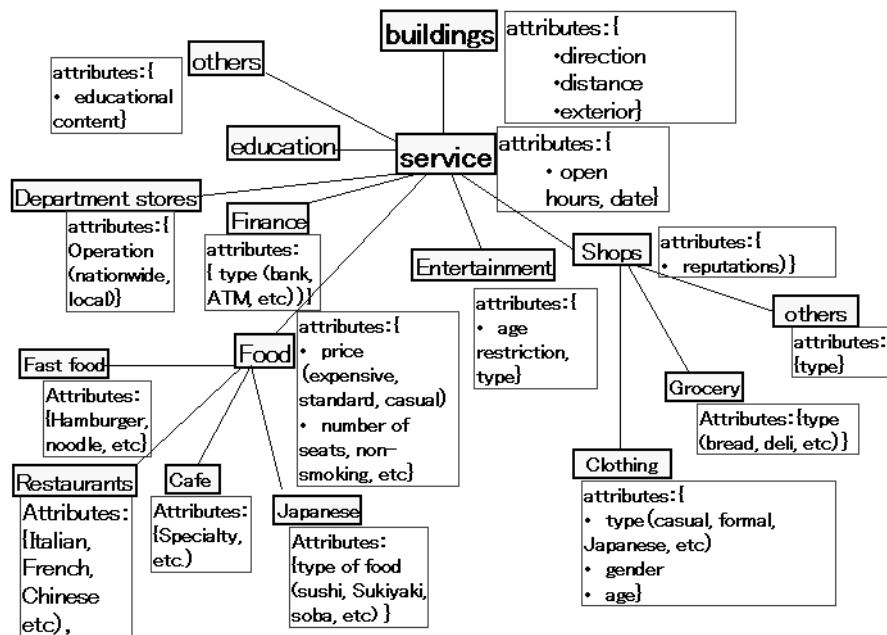
Fig. 7. Ontology of "Buildings" used in the Universal Mobile Interface

information of a department store when looking for other information. If her interest is on "finding other bakeries", then she may want to change the priority value set to the map-based system higher. If she wants to know how crowded another bakery is, she can set the priority value of the image-based system higher.

The universal mobile interface provides on-demand city information by allowing mobile users to select what they want to know most and skip unnecessary information.

4 Evaluations of the Current System and Future Development

4.1 Current System Implementation

The current universal mobile interface system connects three heterogeneous digital cities with a spatial/geographical organization, namely, a 3D modeled virtual space, image-based virtual space, and map-based information space successfully. Although some digital cities have a more topical architecture and lack a spatial/geographical organization, the current system does not include those cities as connectable cities.

The universal mobile interface is designed to use GPS in the future as it uses longitude/latitude for its coordinate system. We will be able to use the universal mobile interface on a mobile phone in a real city when high capacity mobile phones are available in the market.
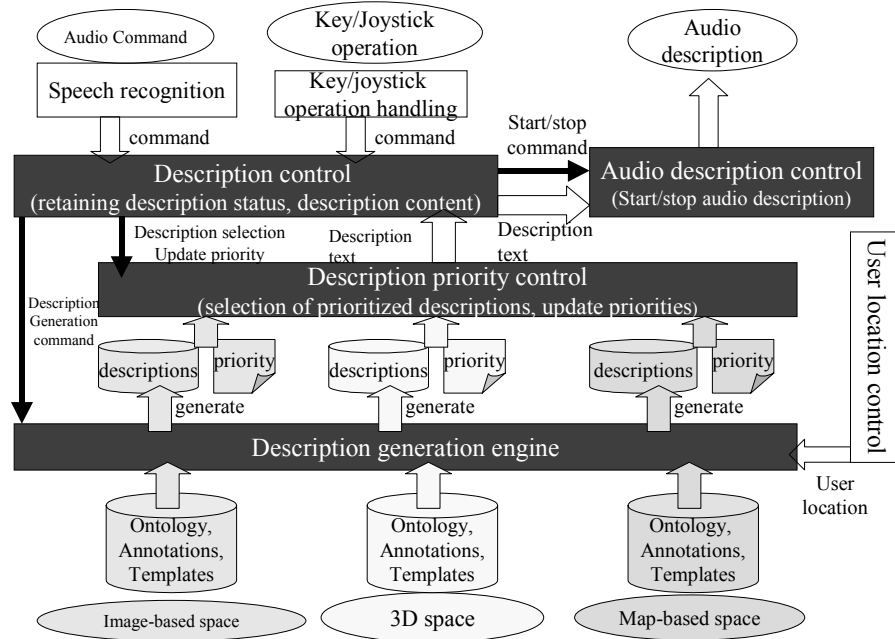
Fig. 8. System components that generate descriptions of objects and controls priority of each digital city to "speak."

4.2 Is the Information Up to Date?

The universal mobile interface is connected to three existing digital city systems. Each connected system was developed in a different time period. There are some shops that don't exist any more, and some that exist in one system don't exist in other systems. This inconsistency of digital cities is inevitable, since cities are alive and change constantly. In order to complement the lack of up-to-date city images, semantic annotation can add recent information on changed objects. For example, "This bank's name has been changed to … after a merger." can be generated by adding attributes to the ontology definition file.

4.3 Semantic Annotation Generation

The current implementation requires creation of annotations of each object. Such content creation is a labor-intensive and costly task.

One solution to the problem is to make use of existing geographical information systems. As in [3], GISs will be one of the key technologies for digital cities that provide regional information. Kyoto city already has a GIS of Kyoto called InfoMap [2] in addition to Kyoto SEARCH. InfoMap provides an extended image map

interface system that enables users to browse a lot of information using 2600 homepages located in the Kyoto metropolitan area. The universal mobile interface can be developed to use such GISs for creating object annotations.

Another solution is to develop an annotation-authoring tool. As the semantic web plays a major role in the current WWW, especially with respect to city information spaces and GISs, semantic annotations of objects in digital cities can be automatically gathered and converted to semantic annotations by a new authoring tool.

4.4   Autonomous Prioritizing

The current description priority setting is initiated by a user. In order to make this priority control mechanism to fully work in a mobile situation, the mechanism should include pre-set user preferences (i.e., user A wants to know directional descriptions; user B wants to get detailed information about restaurants). A general priority setting can be done by a user preference file, while adapting the user's situation can be initiated by the user.

5   Conclusion

The universal mobile interface integrates various existing media representations (digital city systems) representing the same real city. This approach is different from existing approaches that would develop a specific mobile interface for each digital city system, such as a scalable interface for a 3D city information system for mobile services [14]. The universal mobile interface has the following distinct characteristics.

- **Connect**: The universal mobile interface connects heterogeneous digital cities with a single interface.
  Three different virtual spaces are currently connected to the universal mobile interface, and if other virtual representations of downtown Kyoto would become available, they can be connected to the universal mobile interface too. In other words, the universal mobile interface provides a "universal" system interface to connect heterogeneous digital cities.
- **Carry**: The universal mobile interface makes digital cities portable with audio semantic annotation.
  The main goal of the universal mobile interface is to provide on-site, on-demand city information for a mobile user. The current efforts to develop digital cities are divided into two groups: one is to create realistic digital cities with 3D model or digital photo's, the other is to develop a mobile city information system specially dedicated to mobile users. The former systems aim to provide "virtual space walkthrough experience" to a user. We may be able to enjoy visually represented digital cities via a mobile phone or PDA in the near future. However, the universal mobile interface assumes different use. Mobile users cannot always look at a display when they are busy doing other things. The universal mobile interface provides audio information of a city like background music. Thus, the universal mobile interface is "portable" and can be used by city walkers independently of the

situation they are in. The universal mobile interface does not require specifically developed mobile systems.

- **Extract**: The universal mobile interface extracts the required information.
  The universal mobile interface extracts information required by the user, following a prioritizing mechanism. This function is especially important for mobile use, and it also solves the problem that audio annotation is slow, and still busy explaining a scene whereas the user already moves on to the next situation.

Our current demo system connects three heterogeneous digital cities of Kyoto and successfully showed the connectivity of the universal interface. We believe our approach of integrating digital cities will be more useful when semantic web technologies are widely used for GISs.

Acknowledgements

References

1. Bolatto, G., Sozza, A., Gauna, I., and Rusconi, M.: The Geographic Information System (GIS) of Turin Municipality, In: Ishida, T., Isbister, K. (eds.): Digital Cities: Experiences, Technologies, and Further Perspectives. Lecture Notes in Computer Science 1765, Berlin Heidelberg New York, Springer-Verlag (2000) 97-109
2. Hiramatsu, K.: Log Analysis of Map-Based Web Page Search on Digital City Kyoto, In: Tanabe, M, van den Besselaar, P., Ishida, T. (eds.): Digital Cities II: Computational and Sociological Approaches. Lecture Notes in Computer Science 2362, Berlin Heidelberg New York. Springer-Verlag (2002) 233-245
3. Ishida, T.: Understanding Digital Cities. In: Ishida, T., Isbister, K. (eds.): Digital Cities: Technologies, Experiences, and Future Perspectives. Lecture Notes in Computer Science 1765, Berlin Heidelberg New York, Springer-Verlag (2000) 7-17
4. Ishida, T.: Digital City Kyoto: Social Information Infrastructure for Everyday Life, Communications of the ACM, Vol. 45, No. 7 (2002)
5. T Ishida: Activities and Technologies in Digital City Kyoto. In P. van den Besselaar, S. Koizumi (eds), Digital Cities 3. Lecture Notes in Computer Science, Vol. 3081. Springer-Verlag, Berlin Heidelberg New York (2005) pp. 162-183
6. Kambayashi, Y., Cheng, k., Lee, A.: Database Approach for Improving Web Efficiency and Enhancing Geographic Information Systems, 2001 IRC International Conference on Internet Information Retrieval. (2001) 159-176

7.   Koizumi, S, Dai, G., Ishiguro, H.: Town Digitizing for Building an Image-Based Cyber Space. In: Tanabe, M, van den Besselaar, P., Ishida, T. (eds.): Digital Cities II: Computational and Sociological Approaches. Lecture Notes in Computer Science 2362, Berlin Heidelberg New York, Springer-Verlag (2002) 357-370

8.   Koizumi, S. and Ishiguro, H.: Town Digitizing, Omnidirectional Image-based Virtual Space. In P. van den Besselaar, S. Koizumi (eds), Digital Cities 3. Lecture Notes in Computer Science, Vol. 3081. Springer-Verlag, Berlin Heidelberg New York (2005) pp. 242-253.

9.   Rinturi, R., Koivunen, M., and Sulkanen, J. Helsinki Arena 2000-Augmenting a real city to a virtual one. In: Ishida, T., Isbister, K. (eds.): Digital Cities: Experiences, Technologies, and Further Perspectives. Lecture Notes in Computer Science 1765, Berlin Heidelberg New York, Springer-Verlag (2000) 83-96

10. Rinturi R. and Simula, T.: Virtual Helsinki . In P. van den Besselaar, S. Koizumi (eds), Digital Cities 3. Lecture Notes in Computer Science, Vol. 3081. Springer-Verlag, Berlin Heidelberg New York (2005) pp. 110-137.

11. Tanaka, H., Arikawa, M., and Shibasaki R. A 3-D Photo Collage System for Spatial Navigations, In: Tanabe, M, van den Besselaar, P., Ishida, T. (eds.): Digital Cities II: Computational and Sociological Approaches. Lecture Notes in Computer Science 2362, Berlin Heidelberg New York, Springer-Verlag (2002) 305-316

12. The University of Edinburgh GIS WWW Resource List
     http://www.geo.ed.ac.uk/home/giswww.html

13. U.S Census Bureau The Geographic Information Systems FAQ!
     http://www.census.gov/geo/www/faq-index.html

14. Vainio, T., Kotala, O., Rakkolainen, I., and Kupila, H.: Towards Scalable User Interfaces in 3D City Information Systems. In: Paterno, F. (ed.) Mobile Human Computer Interaction. Lecture Notes in Computer Science 2411, Berlin Heidelberg New York, Springer-Verlag (2002) 354-358