

# COMPASS2008: The Smart Dining Service

Ilhan Aslan<sup>1</sup>, Feiyu Xu<sup>1</sup>, Jörg Steffen<sup>1</sup>, Hans Uszkoreit<sup>1</sup>, and Antonio Krüger<sup>1,2</sup>

<sup>1</sup> DFKI GmbH, Germany

<sup>2</sup> University of Münster, Germany

**Abstract.** The COMPASS2008<sup>3</sup> project is a sino-german cooperation, aiming at integrating advanced information technologies to create a high-tech information system that helps visitors to access location-sensitive information services during the 2008 Olympic Games in their preferred language, offering a variety of service-adaptive modalities available on the mobile devices. In this paper, we demonstrate one of the COMPASS2008 services, the *Smart Dining Service*, to showcase the new interaction concepts between multimodality, multilingual and location-sensitive information search.

## 1 Overview

We demonstrate in this paper that a systematic and conceptually thought-through combination of multimodal input and output techniques, multilingual and crosslingual communication and location-sensitive functionalities can greatly enhance tourist assistance systems. Such a combination can yield much more than an aggregation of functionalities. *Multimodal output* can help the user to understand information presented in one of the supported languages even if this language is not the user's mother tongue. *Multimodal input combined with crosslinguality* allows the user to interact with information not presented in his mother-tongue.

We have developed a useful service, called *Smart Dining*, which showcases this new technology combination.

## 2 Smart dining

### 2.1 Demo Scenario

The Smart Dining service is designed to help foreign tourists to orient themselves when it comes to finding their preferred dishes or a restaurant in an unfamiliar environment. Smart Dining can be utilized by users to submit dishes, drink or restaurant queries in their language and to obtain the appropriate suggestions

---

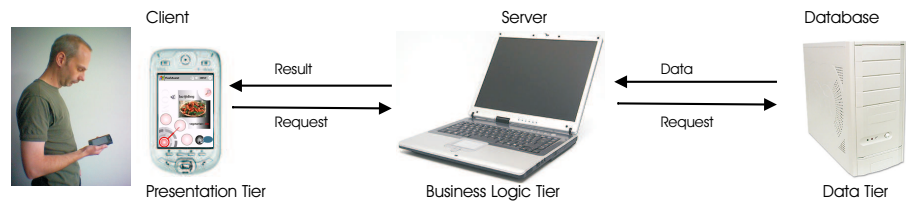
<sup>3</sup> COMprehensive Public informAtion Service project for the Olympic Games 2008 in Beijing, funded by German Federal Ministry of Education and Research, with grant no. 01IMD02A and selected for cofunding by the Chinese Ministry for Science and Technology.

both in their preferred language and Chinese. The Chinese information presented on a mobile device helps foreigners to communicate with the native speakers, such as restaurant personnel.

The Smart Dining service is available on mobile devices that support in addition to search and translation functionalities also multimodal interaction. Deficits of display size are balanced by the choice and combination of interaction modalities.

## 2.2 Architecture

The Smart Dining demo employs a three-tier architecture (see Fig 1).



**Fig. 1.** Overview of three-tier demo architecture

The **data tier** consists of a database that provides detailed multilingual and multimedia information about dishes, drinks and restaurants. It contains three sets of tables, one for each search category, namely, food, drink and restaurant. Within each set, there is one extra table for holding the language independent information, e.g., the phone number of a restaurant or the file location of a dish image. The language independent information is linked with the language dependent information in three other tables, one for each of the supported languages. Some attributes of a category are restricted to a fixed set of values, e.g. the taste of a dish or the cuisine of a restaurant. An additional table holds a multilingual dictionary where these values are listed in all three languages. This supports precise crosslingual queries. For crosslingual queries about attributes that have an unrestricted set of values, a less precise automatic machine translation is used. In the table set, an extra field is reserved for storing audio files which are the voice outputs of names in Chinese such as dish names. This audio information is extra developed to help users to pronounce names in Chinese.

The **business logic tier** is implemented in Java and runs as a server on a desktop PC. It manages user sessions and processes queries send by the client via HTTP using multilingual and crosslingual information access technologies ([2] and [4]). We allow not only form-based queries which are standard for database queries, but also free text search by applying free text indexing to all textual data in the databases. Users can fill or select their preferences of a dish, a drink or a restaurant in the query forms. In case of free text retrieval, users can submit a free text query such as "sweet and sour Shanghai dish" straightforwardly

without stepping through the form fields. The free text search can be entered by either handwriting or speech. The advantages of this approach is that we leave it open to users how precise they want to formulate their queries and support speech as an additional modality. In order to search at the same time the related information in Chinese, the language dependent part of queries are translated into Chinese when needed, and then the whole queries are mapped to a sequence of SQL statements that yield the requested information. A special process is responsible for combining the free text retrieval and database retrieval results. We use XML as our data interchange format between components.

The **presentation tier** runs on a Pocket PC 2003 based PDA. It is responsible for the user interaction. This includes multimodal query generation and adaptive information presentation. It combines a pedestrian navigation system, which is described in [1] and multimodal interaction concepts based on a shopping assistant described in [3]. The pedestrian navigation service is extended to highlight restaurants in the local environment as provided by the Smart Dining service. The multimodal concepts described in [3] are extended with multilinguality, e.g. speech and handwriting is allowed in multiple languages. In addition, the presentation tier provides a Flash based, vivid graphical user interface module.

### 2.3 Multimodal and Crosslingual Information Access

The user can query the database in many ways. The query grammar allows the specification of one or more dish attributes. These are either predefined values (e.g. meat, fish, vegetarian or hot, sweet-and-sour) or dynamic values (e.g. the name of the dish, its ingredients or preparation). To support crosslingual retrieval, query terms are translated into other languages so that the query also returns dishes in a language different from the query language.<sup>4</sup> Using the *unimodal* speech modality to formulate a query, a user could say:

- speech: *"Show me only sweet-and-sour dishes."*

The user can also utilise a location-based restriction to get dishes that are available at a restaurant near his current position, using the ability of the mobil device to transmit the user's location. The result of such queries is a list of dishes, that matches the criteria the user provides, in addition to any context data that is provided by the handheld device. At this point, the user has three options: He can bring up an overview map with restaurants, he can pick an entry and check the dish details or he can further refine his search. The overview map shows the user's current position and all restaurants that offer any dish contained in the result list. The user can then pick a restaurant for further details about the restaurant and the offered dishes. The user could do this with a simple *multimodal* query:

- speech: *"Show me all dishes of this restaurant."* + gesture: tap on a restaurant on a map

---

<sup>4</sup> A corresponding query is possible for drinks, using the appropriate attributes.

If he decides to eat at the chosen restaurant, he can activate the navigation service that guides him to the restaurant. If not all the information is available in the user's language, he can use a *crosslingual (and multimodal)* query for an indicative translation:

- speech: *"Translate this Chinese writing."* + gesture: tap Chinese writing on the display

The dish description includes several output modalities: a transcription of the dish's name in the user's language, its name in Chinese characters, a picture of the dish and a link to an audio file that plays the Chinese dish name.<sup>5</sup> The latter can be used to order the dish in a restaurant.

So far, we have described the Smart Dining service using dish search as entry point. Another entry point is required for a scenario where the user has not decided yet what he wants to eat, but is interested in finding a restaurant in the first place. The user guidance is similar to the one in the dish search. Restaurant details are also available in different modalities: a transcription of the name in the user's language, a picture or video presentation of the restaurant, its name in Chinese characters, etc.

### 3 Conclusion

The COMPASS smart dining service represents new opportunities to combine multimodal, multilingual and location-based services for useful information access. This combination provides users with a highly versatile range of options to express their different search requirements, which is also reflected in the presentation of the results. Furthermore, this service is general enough to adapt to other cities and countries.

### References

1. A. Krüger, A. Butz, C. Müller, C. Stahl, R. Wasinger, K.-E. Steinberg, and A. Dirschl. The connected user interface: realizing a personal situated navigation service. In *IUI '04: Proceedings of the 9th international conference on Intelligent user interface*, pages 161–168, New York, NY, USA, 2004. ACM Press.
2. H. Uszkoreit. Cross-lingual information retrieval: From naive concepts to realistic applications. In *Proc. of the 14th Twente Workshop on Language Technology*, 1998.
3. R. Wasinger, A. Krüger, and O. Jacobs. Integrating intra and extra gestures into a mobile and multimodal shopping assistant. In *Pervasive*, pages 297–314, 2005.
4. F. Xu. *Multilingual WWW — Modern Multilingual and Crosslingual Information Access Technologies*, chapter 9 in Knowledge-Based Information Retrieval and Filtering from the Web, pages 165–184. Kluwer Academic Publishers, 2003.

---

<sup>5</sup> Currently we are working on a text to speech solution for Chinese output generation.