Challenges in Mobile Multimodal Application Architecture

Toshihiko Yamakami
CTO Office, ACCESS
2-8-16 Sarugaku-cho, Chiyoda-ku, Tokyo, Japan
Toshihiko.Yamakami@access-company.com

Abstract

The multimodal application had been considered for promising mobile applications in early 2000. An industrial forum terminated its 7-year effort to make standardization of mobile multi-modal applications. Many researchers and industrial players still think mobile multimodal applications as rich opportunities of new services. The author attempts to analyze the gap between these two different aspects of the current mobile multimodal application, and presents some implications.

1 Introduction

The mobile Internet quickly penetrates every-day life. The Internet has been rapidly mobilized by increasing penetration to wireless users. It will come as no surprise that the mobile Internet occupies an important role in the Internet in the coming decades. The mobile Internet users outnumbered the PC Internet ones in Japan in 2005 and this dominance will be expected to persist in the next decade.

2 Landscape of Mobile multimodal Services

2.1 Underlying Technologies

There is little doubt that the underlying capabilities available in mobile handsets are improving in a never-ending manner. Many high-end commercial devices are equipped multiple CPUs with multiple hundred MHz horsepower. The penetration of 3G(IMT-2000) handsets is increasing. For example, 78 %of wireless subscribers used 3G in Japan at end of September 2007. It means more bandwidth is readily available for mobile handsets. The 3G phones in Japan are equipped 2 GByte Micro-SD memory, 3 Mega pixel camera, and 100 MB memory. All these capabilities are improving day-by-day in network, CPU and memory.

2.2 Standardization in Industrial Fora

The mobile phones are equipped with voice capabilities from Day One. They were considered as an excellent device for multimodal applications. Also, it was recognized that the lack of standardized framework will lead to a disaster as early as 2000. WAP Forum started a work item effort called multimodal Multi-device Services (MMMD service) in 2000 is to enable the user to interact with services through multiple modalities or devices. There was no objection that multimodal is expected to be a key element of the mobile user interface. When WAP Forum was reengineered to OMA (Open Mobile Alliance), the work item was successfully transferred to OMA. And it was an approved work item in 2002. After a 5-year struggle in OMA (following 2 years in WAP Forum), OMA decided to terminate the work item without producing any technical specification except an Architecture document [1] in 2007. The Architecture document describes how different modalities or devices can be synchronized so that interaction in one or multiple modality or devices is appropriately reflected in the other registered modalities and devices as well as in the application logic or data model. Many people have no doubts that the future mobile devices will be empowered by multimodal applications. They can still see much enthusiasm in this field. However, it is unlikely that they will renew their efforts in a short-range time frame. It is important to take an in-depth view that the hidden challenges we have in the mobile multimodal service standardization.

2.3 Underlying Operating Systems

The third-party software can be portable to mobile handsets, especially Linux is now a promising platform for a mobile handset. They can see multiple Mobile Linux fora, for example, LiPS and LiMo, for standardizing mobile Linux platform. This also increases easiness of mobile multimodal application deployment.

2.4 Software on a Mobile Handset

Many mobile handsets are equipped with a micro-browser. Many micro-browsers embedded in mobile handsets accommodate a million line of codes, just demonstrating how much software capabilities are available in recent handsets.

3 Views on Three Challenges

Considering the gap between the expectation and adoption in industrial level, the author considers the following three challenges in the mobile multimodal applications:

- Challenges in Encapsulation
- Challenges in multimodal Contexts
- Challenges in Content Authoring

3.1 Dependencies in Underlying Platforms

One observation in the challenges in mobile multimodal services is that the multimodal applications significantly depend on the specific capabilities of a device. The tight-bound-ness to the underlying capabilities lead to either a) generalization obscures the real issues, b) design without specific knowledge of the underlying environment does not solve problems in the real world, or c) rapid progress of the underlying environment constantly threats the higher-layer standardization. The author observes that the rigorous efforts to unbind this relation to the lower-layer implementation details were not so successful compared to the other parts of the Internet. Considering the progress of the underlying environment, the rapid progress in the underlying environment itself may be an obstacle to the mobile multimodal services than a leverage to it.

3.2 General vs. Specific multimodal Services

The multimodal application architecture is very general in W3C or in OMA. This covers the wide variety of multimodal applications. This generality is good for design and early stages, however, they will notice the fundamental departures of different multimodal contexts. We can identify the three different contexts for multimodal services:

- Single application with multimodal interactions with a user
- Single application with different modal interactions to different users
- Single application with single modality with a user at a time (different contexts provide different modal interactions with the user)

The general model can solve all these three cases. However, it should be noticed that the three application domains have different requirements for interaction design and configuration. Authoring and configuration are critical in network applications and multimodal one is no exception. The author considers that the further investigations on classifications of multimodal contexts can ease deployment of multimodal applications or not. When the unexpected challenge refrain us from building useful frameworks for multimodal applications, we may need to split the general model. However, it will give a risk that the classifications can be easily obsolete in the future (especially in the mobile application domain). It should be noted that the PC Internet does not leverage the audio-related multimodal applications except plain video. The public place has natural constraint on audio applications, which may lead to significant restrictions to demonstrate multimodal services in public.

3.3 Authoring

It is still a challenging issue how we can construct commercially feasible multimodal applications. Writing codes in distributed environment may be compared writing codes in an assembly language for a machine. They may provide high-level functional languages to facilitate authoring, however, it will unavoidably expose the underlying complexity in a debugging phase. These two-level challenges for authoring commercially feasible multimodal applications are unbeaten challenges for widely adoption of mobile multimodal applications. This is a two-fold challenge with a challenge for general multimodal applications and a challenge to deploy complicated software in mobile handset environments.

4 Conclusions

The mobile multimodal services have been a promising domain since early 2000s. There are two aspects that we can observe, one is the increasing capability in mobile handsets, and the other is stagnancy in mobile industrial standardization. The author presents three views on these conflicting situations to identify the challenges in mobile multimodal service design and standardization.

References

[1] Open Mobile Alliance™.
OMA Multimodal and Multi-device Enabler Architecture V1.0.
http://www.openmobilealliance.org/, October 2006.