Photo-Based User Interfaces: Picture It, Tag It, Use It

Geert Vanderhulst, Kris Luyten, and Karin Coninx

Hasselt University – transnationale Universiteit Limburg – IBBT
Expertise Centre for Digital Media
Wetenschapspark 2, 3590 Diepenbeek, Belgium
{geert.vanderhulst,kris.luyten,karin.coninx}@uhasselt.be

Abstract. Pervasive environments can be hard to configure and interact with using handheld computing devices, due to the mismatch between physical and digital worlds. Usually, smart resources in the user's vicinity are discovered and presented in a menu on the user's device from where they can be accessed. However, in environments with many embedded resources it becomes hard to identify resources by means of a textual description and to get aware of the tasks they support. As an alternative to menu-driven interfaces, we demonstrate annotated photos as a means for controlling a pervasive environment. We present as part of our approach a tool that enables people to picture their own environment and use photos as building blocks to create an interactive digital view on their surroundings. To demonstrate and evaluate our approach, we engineered a pervasive prototype application that is operated through a photo-based user interface and assembled using ontologies.

1 Introduction

Many new embedded computing resources and networked appliances that hit the market try to offer simple access to their functionality by minimizing the set of physical controls. As a consequence, these devices rely on digital user interfaces that migrate to personal devices from where they can be operated and configured by a remote user interface. Such computer interfaces, driven by pervasive services and weaving into the fabric of everyday life, give rise to a pervasive computing environment [8]. Moreover, in the vision of the Internet of Things [1] almost any object becomes part of a computer network which demands for proper ways to control all these objects. Awareness of the environment and its resources is key for efficient interaction.

Service discovery frameworks such as UPnP[1] and directory services provide methods to discover and use pervasive services, but only return limited information about these services. A more advanced approach integrates computer-enabled resources along with the tasks they support in a meta-user interface [2,6]. This user interface is 'meta' because it acts as an interface for accessing other, application-specific user interfaces from a menu. However, when the number of resources that are embedded in a pervasive environment increases, it becomes more difficult for end-users to locate them in a menu and to differentiate between similar resources based on a description. To overcome the complexity of pervasive environments we propose photo-based user interfaces that can

[1] http://www.upnp.org/

R. Meersman, P. Herrero, and T. Dillon (Eds.): OTM 2009 Workshops, LNCS 5872, pp. 610–615, 2009.

be created by end-users themselves. We believe interactive photos are a useful instrument to make immobile and invisible resources such as lights and media services (e.g. represented by speakers on a photo), easily accessible in the digital world.

In this paper we present a software framework that enables users to picture an environment and use photos to interact with their surroundings. Our contribution is the conjunction of tool support for photo-based user interfaces and a semantic binding between interface and back-end. First we introduce Photoporama, a tool for annotating photos by tagging the things of interest that appear on a photo. Second, we illustrate, using a prototype application, how ontologies help to glue interactive photos and pervasive services together.

2 Related Work

Strategies for annotating photos with semantic information have been widely studied. Various approaches use context-descriptive keywords or visual similarity for annotating photos (semi-)automatically. Guus et al. propose a subject matter ontology to describe the vocabulary and background knowledge of a photo's subject domain [3]. Instead, we annotate the objects that appear on a photo individually and use ontologies to connect objects with application logic. Facebook[2] applies a similar approach to mark people on photos and Flickr[3] supports notes that can be attached to regions drawn on a photo. By adopting familiar tagging methods, Photoporama enables non-expert users to create semantically enriched photos that serve as a user interface to steer pervasive applications.

The use of photos to discover and interact with pervasive services has been suggested before in [4]. In this work, a 'u-Photo' is tagged with eyemarks: physical entities that appear on a photo and that represent a pervasive service. Whereas u-Photo uses visual markers as tags for eyemarkers, Photoporama links objects that appear on a photo with richer semantics using the WordNet lexicon. This allows us to connect objects on a photo with services in the environment: the user can interact with the photo to access, configure and operate related services.

3 A Semantic Photobook

We define a photobook as a set of digital images, including camera captured photos as well as fabricated images such as a floorplan. Each photo in a photobook is enriched with information about the objects that appear in the photo and references to other photos. For example, a door on one photo could refer to a photo showing the room behind that door. Hence one can browse through a photobook in a similar way as navigating a web page. We provide the Photoporama toolkit[4] to create photobooks and ease the integration of a photo-based user interface in a pervasive application.

Annotating a photo involves two steps. First, different objects that are related to the usage of pervasive services are identified in the photobook and assigned a number

[2] http://www.facebook.com/
[3] http://www.flickr.com/
[4] http://research.edm.uhasselt.be/photoporama/

of keywords. For example, the piano in figure 1 is tagged as a piano in the sense of 'a keyboard instrument'. Photoporama uses the WordNet dictionary to disambiguate between the different senses a word might have. The user selects the word sense she has in mind from a list and attaches it as a tag to the object. The second step consists of marking the objects identified in the first step on a photo. This is achieved by drawing a rectangular area on a photo and linking this area with an object, similar to tagging people in Facebook. Hence different photos can link to the same objects while they differ in e.g. viewing angle, distance to the subject, level of detail, etc.

Bringing together different photos and objects in a digital photobook demands for efficient searching strategies. A tag cloud, composed of the various annotations on a photo and weighted by their frequency in the photobook, allows users to quickly navigate to objects and browse through photos. Moreover, the linguistic relations that apply between words such as synonyms and hypernyms – 'is a' relationships, e.g. 'musical instrument' is a 'hypernym' for 'piano' – are exploited to search for available objects.

4 Engineering Photo-Based Applications

In this section we discuss the process of designing a photo-based user interface for a pervasive application. We have built a prototype application that displays scores for a piece of music on a screen or wall surface. A 'score service' takes as input a piece of music and a musical instrument which identifies the type of scores – piano scores differ from e.g. guitar scores. When a score service receives input, it searches for corresponding scores in a shared database and renders them on the local device. The user can then remotely navigate through these scores via a distributed user interface. We designed a photo-based user interface to operate the application and linked it with the pervasive score services by means of an ontology as depicted in figure 1.

In our test environment, we deployed two score services, one running on a computer attached to a projector and another one on a notebook connected to a television set. We

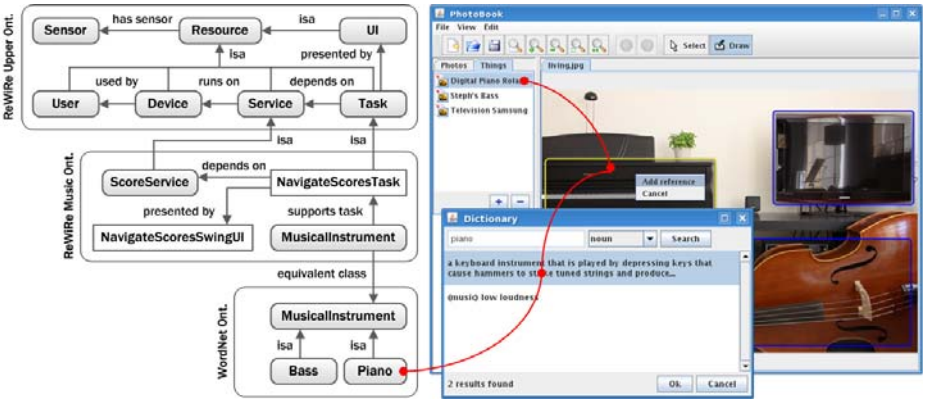
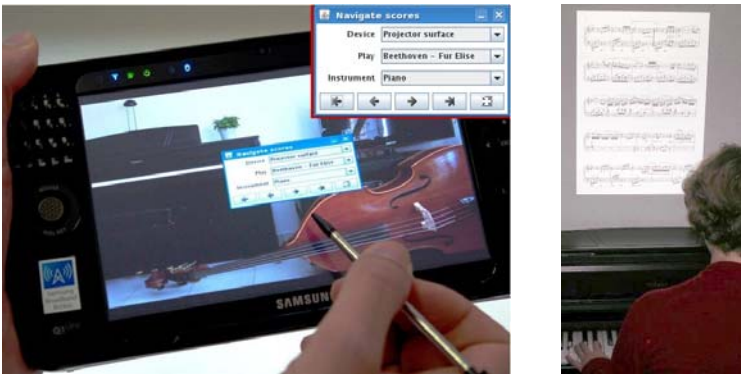


Fig. 1. A piano is marked and annotated on a photo. According to WordNet, a piano in the sense of a keyboard instrument is a musical instrument, which is linked with a domain ontology.



(a) (b)

Fig. 2. Interacting with a photo-based user interface running on a UMPC (a) pops up a migrated user interface to steer the projection of scores above a piano in the physical world (b)

used a UMPC with the ReWiRe [7] service platform pre-installed and the Photoporama toolkit to steer services using a photo-based user interface as shown in figure 2. Two musical instruments, a piano and a double bass, were marked on a photo and annotated using WordNet. When a user selects an instrument, the ontologies that describe the pervasive environment are queried and a navigate scores task is found and listed in a menu on the photo. When invoked, one is presented with a user interface that displays the selected instrument type (e.g. piano or bass) and asks for a preferred output device for displaying the scores for a piece of music. This input is then passed to the score service on the selected output device that renders the scores which she can navigate from her UMPC. To enhance navigating scores while playing an instrument, the navigation buttons are rendered ten times their original size so that they can easily be tapped.

4.1 Pervasive Services as Back-End

As back-end, we use the ReWiRe framework to create and deploy dynamic pervasive services. The ReWiRe platform runs on the computing devices that give rise to the pervasive environment and employs a centralized runtime environment model which is shared amongst devices. In this environment model, ontologies describe the environment's topology and instances of these ontologies represent the environment's current state. An upper ontology defines a number of generic concepts and is merged with domain-specific ontologies at runtime. Figure 1 shows a domain ontology for the score application which is imported into the environment model when a score service is deployed. It shows that a 'musical instrument' supports a 'navigate scores task' which is presented by a corresponding user interface. Note that tasks – defined as executable user processes using the Web Ontology Language (OWL)[5] – are essential to describe the goals the end-user can achieve in the environment while services are functional components a task relies on.

[5] `http://www.w3.org/2004/OWL/`

4.2 Photos as User Interface

When used as a user interface, a set of photos replaces the windows, dialogs and other widgets found in traditional form-based user interfaces. The objects marked on a photo become interactive parts which allow to navigate through photos or manipulate the state of objects in the environment, displayed on a photo. To simplify the integration of a photo-based user interface within an application, Photoporama treats photos as user interface widgets with their own interaction events. We used Photoporama to create a meta-user interface on top of the ReWiRe framework. Hence annotated photos become a means to interact with the pervasive environment. When an annotated object on a photo is selected, one is presented with supported tasks in a context-sensitive menu. Selecting a task then results in a user interface being displayed to interact with one or more services related to the selected object.

4.3 Ontologies as Glue

In order to link objects on a photo with resources in the pervasive software system, we use ontologies as a binding between the user interface and the application logic. Moreover, WordNet is used to mediate between Photoporama tags and ReWiRe domain ontologies. This is achieved by mapping a word sense on its corresponding OWL individual as discussed in [5] and mapping concepts in a domain ontology on WordNet individuals via OWL's built-in "equivalentClass" property. In practice, a domain ontology designer must link the OWL classes he creates with WordNet while an end-user simply has to tag her photos to realize these bindings. Hence, by observing an object's tags, its corresponding resource class(es) in the pervasive system can be semantically resolved along with a list of tasks that are supported by this type of resource. For example, the piano in figure 1 supports a 'navigate scores task' because it relates with the 'musical instrument' concept defined in the domain ontology. According to WordNet, 'musical instrument' is a hypernym of 'piano' and thus matches with the 'musical instrument' concept defined in the domain ontology which is denoted equivalent with a WordNet 'musical instrument' by the ontology designer.

In our prototype application tasks are derived based on classes of resources; *any* musical instrument can be played using scores. However, if multiple instances of a resource exist (i.e. a piano in the living room and a piano in the hall), additional information is required to differentiate between similar resources. This is particularly the case for stateful computer-augmented resources such as the light on the piano: a specific light serves as input for a service that steers the lights in an environment. In this situation, it is useful to tag the piano light with a reference to its representation in the pervasive software framework, e.g. through a URI that differs in namespace from WordNet tags.

5 Discussion

We believe a major advantage of our approach is the loose coupling between user interface and pervasive services: a software developer creates services and designs domain ontologies mapped on WordNet, an end-user pictures her own environment the way she likes it and annotates her photos using the Photoporama tool; assembly happens at

runtime. Awaiting a more formal evaluation, our first experiments with Photoporama already surfaced a number of enhancements. Currently, the state of the environment (e.g. is a light switched on or are scores projected?) can be observed in the digital world using application-specific user interfaces, but it would be nice if photos directly provide feedback about the real world. This can be achieved by introducing a 'state layer' in Photoporama with semi-transparent or animated images connected to resource properties in ReWiRe, such as a yellow glow around a light or musical notes to indicate that scores are being projected. Furthermore, photos can be a useful instrument to help users manually allocate resources for pervasive services by selecting them on a photo.

Acknowledgments. Part of the research at EDM is funded by ERDF (European Regional Development Fund), the Flemish Government and the Flemish Interdisciplinary Institute for BroadBand Technology (IBBT).

References

1. ITU Internet Reports 2005. The Internet of Things, `http://www.itu.int/osg/spu/publications/internetofthings/`
2. Coutaz, J.: Meta-User Interfaces for Ambient Spaces. In: Coninx, K., Luyten, K., Schneider, K.A. (eds.) TAMODIA 2006. LNCS, vol. 4385, pp. 1–15. Springer, Heidelberg (2007)
3. Schreiber (Guus), A.T., Dubbeldam, B., Wielemaker, J., Wielinga, B.: Ontology-Based Photo Annotation. IEEE Intelligent Systems 16(3), 66–74 (2001)
4. Suzuki, G., Aoki, S., Iwamoto, T., Maruyama, D., Koda, T., Kohtake, N., Takashio, K., Tokuda, H.: u-photo: Interacting with pervasive services using digital still images. In: Gellersen, H.-W., Want, R., Schmidt, A. (eds.) PERVASIVE 2005. LNCS, vol. 3468, pp. 190–207. Springer, Heidelberg (2005)
5. van Assem, M., Gangemi, A., Schreiber, G.: Conversion of WordNet to a Standard RDF/OWL Representation. In: Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC 2006), Genova, Italy (2006)
6. Vanderhulst, G., Luyten, K., Coninx, K.: Put the User in Control: Ontology-Driven Meta-level Interaction for Pervasive Environments. In: Proceedings of the 1st International Workshop on Ontologies in Interactive Systems (ONTORACT 2008), pp. 51–56. IEEE Computer Society, Los Alamitos (2008)
7. Vanderhulst, G., Luyten, K., Coninx, K.: ReWiRe: Creating Interactive Pervasive Systems that cope with Changing Environments by Rewiring. In: Proceedings of the 4th IET International Conference on Intelligent Environments (IE 2008), pp. 1–8 (2008)
8. Weiser, M.: The Computer for the 21st Century. Scientific American, 94–104 (1991)