# 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 propose annotated photos as a means for controlling a pervasive environment. We present 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 computer-enabled resources and networked appliances that hit the market nowadays have a simple form factor with a minimal 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 remotely. Such computer interfaces, driven by pervasive services and interweaven into the fabric of everyday life, give rise to a pervasive computing environment [8]. To interact with the environment, one must first become aware of its resources and their functionalities.

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 [1, 6]. This user interface is 'meta' because it acts as an interface for accessing other, application-specific user interfaces from a menu. However, the more resources that are embedded in a pervasive environment, the more difficult it becomes for end-users to locate them in a menu and to differentiate between similar resources based on a description. To make resource resolution from personal devices more natural, we propose photo-based user interfaces that can be created by end-users themselves. We believe interactive photos are a useful instrument to make immobile and invisible resources in an environment, such as lights and media services (e.g. represented by speakers on a photo), easily accessible in the digital world.

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

In this paper we present a software framework that enables users to picture an environment and use photos to interact with their surroundings. Our contributions our twofold. First we introduce Photobook, 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 [2]. Instead, we annotate the things that appear on a photo individually and use ontologies to connect things 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. Photobook goes one step further and delivers semantically enriched photos that can 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 [3]. 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, Photobook assigns a software-independent meaning to the things that appear on a photo, using the WordNet lexicon[4]. A photo then becomes interactive at runtime, by reasoning over an ontology that connects the things tagged on a photo with pervasive services. Likewise, [5] uses ontological and linguisitic knowledge for reference resolution in a virtual environment. A speech grammar, annotated with synonyms, is automatically generated during the conceptual modelling phase of a virtual environment.

## 3 A Semantic Photobook

The Photobook toolkit supports the creation of a semantic photobook: a set of digital images - from panorama images to fabricated pictures such as a floorplan - annotated and connected to each other via a interchangeable Photobook XML document. In Photobook, annotating a photo involves two steps. In a first step, different things of interest that appear on one or more photos are identified, labeled and assigned a number of keywords. For example, the piano in figure 1 is tagged as a piano in the sense of 'a keyboard instrument'. Photobook 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 thing. The second step consists of marking the things identified in the first step on a photo. This is achieved by drawing a rectangular area on a photo and linking this area with a thing, similar to tagging people in Facebook.

[2] http://www.facebook.com/
[3] http://www.flickr.com/
[4] http://wordnet.princeton.edu/

Hence different photos can link to the same things while they differ in e.g. viewing angle, distance to the subject, level of detail, etc.

Bringing together different photos and things in a digital photobook demands for efficient searching startegies. 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 things 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 things.

## 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.

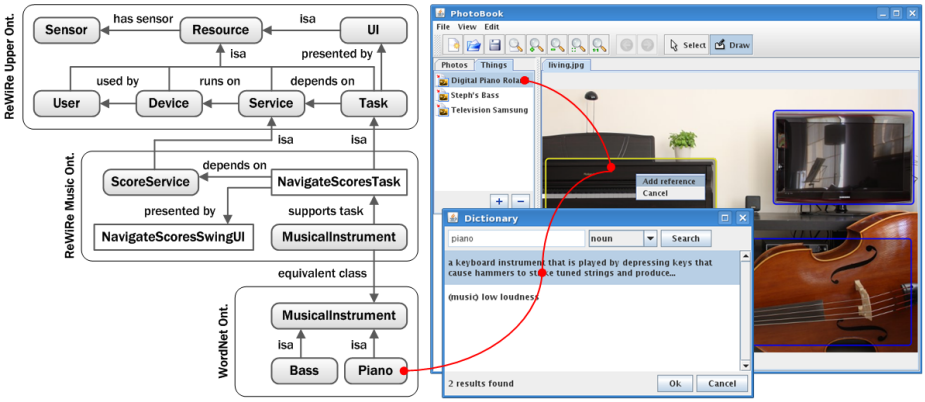



**Fig. 1.** A piano is marked and annotated on a photo. According to WordNet, a piano in the sense of a keyboard instrument is a musical instrument, which is linked with a domain ontology.

### 4.1 Pervasive services as back-end

As back-end, we use the ReWiRe framework to create and deploy dynamic pervasive services. A new service such as the score service is deployed in ReWiRe as an OSGi software component, called a 'bundle'. 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. In ReWiRe, domain ontologies are an application's backbone as they describe the resources relevant for the application, the relations that can exist between resources, events that can occur in a resource, etc. 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 - OWL things imported in ReWiRe via domain ontologies - describe a goal the end-user can achieve in the environment while services are functional components a task relies on. Tasks can thus be considered as an abstraction for concrete user interfaces: a ReWiRe client will select an appropriate presentation for a task based on the available modalities of the device it runs on (e.g. a graphical or a speech-based user interface) [7].

### 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 things marked on a photo become interactive parts which allow to navigate through photos or manipulate the state of things in the environment, displayed on a photo. To simplify the integration of a photo-based user interface within an application, Photobook treats photos as user interface widgets with their own interaction events. As a proof of concept, we embedded Photobook as a user interface toolkit in ReWiRe. Hence annotated photos (Photobook XML) can be imported in ReWiRe and become a means to interact with the pervasive environment. When an annotated thing 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 thing in question.

### 4.3 Ontologies as glue

In order to link things 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 Photobook tags and ReWiRe domain ontologies. This is achieved by mapping a word sense on its corresponding OWL individual as discussed in [4] 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 a thing'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.

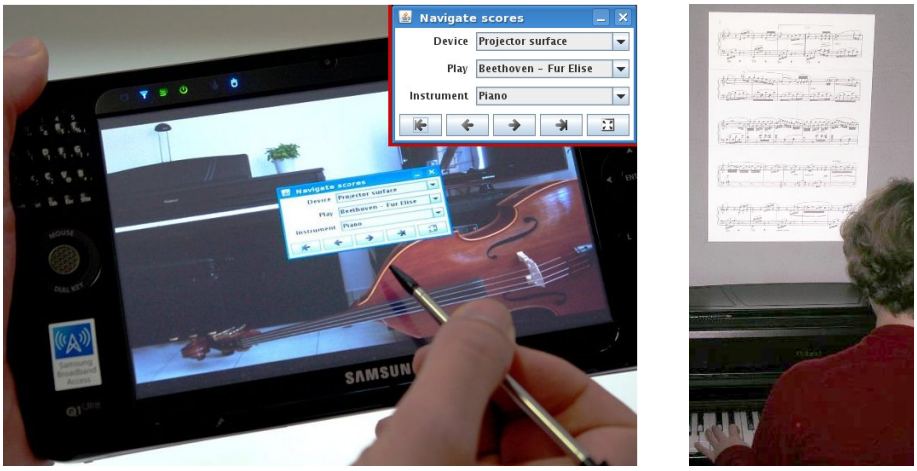
In our prototype application tasks are derived based on classes of resources; *any* musical instrument can be played using scores. However, if multiple instances of a resource exist (i.e. a piano in the living room and a piano in the hall), additional information is required to differentiate between 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 and conclusions

In our test environment, we deployed two ReWiRe score services, one running on a computer attached to a projector and another one on a notebook connected to a television set. We used a UMPC with ReWiRe and the Photobook toolkit installed to steer these services using a photo-based user interface as shown in figure 2. Two musical instruments, namely 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.

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 Photobook tool; assembly happens at runtime. Although our first experiments with Photobook were very promising, we have come up with 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 Photobook 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 EFRO (European Fund for Regional Development) and the Flemish Government. Funding for this research was also provided by the Research Foundation – Flanders (F.W.O. Vlaanderen, project number G.0461.05).




(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).

## References


1. Joëlle Coutaz. Meta-User Interfaces for Ambient Spaces. In *Proceedings of the 5th International Conference on Task Models and Diagrams for Users Interface Design (TAMODIA'07)*, pages 1–15, 2007.
2. A. Th. (Guus) Schreiber, Barbara Dubbeldam, Jan Wielemaker, and Bob Wielinga. Ontology-Based Photo Annotation. *IEEE Intelligent Systems*, 16(3):66–74, 2001.
3. Genta Suzuki, Shun Aoki, Takeshi Iwamoto, Daisuke Maruyama, Takuya Koda, Naohiko Kohtake, Kazunori Takashio, and Hideyuki Tokuda. u-Photo: Interacting with Pervasive Services Using Digital Still Images. In *Proceedings of the 3rd International Conference on Pervasive Computing (PERVASIVE'05)*, pages 190–207, 2005.
4. M. van Assem, A. Gangemi, and G. Schreiber. Conversion of WordNet to a Standard RDF/OWL Representation. In *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC'06)*. Genova, Italy, 2006.
5. Lode Vanacken, Chris Raymaekers, and Karin Coninx. Automatic Speech Grammar Generation During Conceptual Modelling of Virtual Environments. *Vis. Comp.*, 24(11):955–961, 2008.
6. Geert Vanderhulst, Kris Luyten, and Karin Coninx. 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'08)*, pages 51–56. IEEE Computer Society, 2008.
7. Geert Vanderhulst, Kris Luyten, and Karin Coninx. ReWiRe: Creating Interactive Pervasive Systems that cope with Changing Environments by Rewiring. In *Proceedings of the 4th IET International Conference on Intelligent Environments (IE'08)*, pages 1–8, 2008.
8. Mark Weiser. The Computer for the 21st Century. *Scientific American*, pages 94–104, 1991.