

WiiNote: Multimodal Application Facilitating Multi-User Photo Annotation Activity

Elena Mugellini
University of Applied
Sciences of Western
Switzerland, Fribourg
Bd de Pérolles 80,
1705, Fribourg, Switzerland
+41 26 429 6870

elena.mugellini@hefr.ch

Maria Sokhn
University of Applied
Sciences of Western
Switzerland, Fribourg
Bd de Pérolles 80,
1705, Fribourg, Switzerland
+41 26 429 6978

maria.sokhn@hefr.ch

Stefano Carrino
University of Applied
Sciences of Western
Switzerland, Fribourg
Bd de Pérolles 80,
1705, Fribourg, Switzerland
+41 26 429 6748

stefano.carrino@hefr.ch

Omar Abou Khaled
University of Applied
Sciences of Western
Switzerland, Fribourg
Bd de Pérolles 80,
1705, Fribourg, Switzerland
+41 26 429 6589

omar.aboukhaled@hefr.ch

ABSTRACT

In this paper, we describe a multimodal application, called WiiNote, facilitating multi-user photo annotation activity. The application allows up to 4 users to simultaneously annotating their pictures adding either textual or vocal comments. Users use the Wii Remote device to select the whole picture or a specific region of it to be annotated. Annotations can be either free or structured, i.e. based on a domain specific data model expressed using MPEG7 standard or RDF language for ontology.

Categories and Subject Descriptors

H.1.2 [Information Systems]: User/Machine Systems – *Human factors, Human information processings.*

H.4. [Information Systems]: Information Systems Applications.

General Terms

Design, Experimentation, Human Factors.

Keywords

Multimodal System, Wii Remote, Multimedia Annotation, Semantic.

1. INTRODUCTION

With the widespread diffusion of digital cameras and their integration within mobile phones and PDA (Personal Digital Assistant) devices, the number of photos taken by people during their daily life activities never stops increasing. Photos are used to recall souvenirs as well as to share experiences with other people. They provide a specific view of “something” (a place, an object, etc.) at a particular moment on time. However finding the right picture at the right time becomes nowadays harder than ever because of such growing quantity of information. The availability of metadata annotation over multimedia content such as photos is known to enhance information retrieval and organization, particularly for large data set. Current digital photos contain some metadata information such as the time and date when the picture was taken, the size of the picture, etc. However these additional metadata provide mainly technical information, which is neither

sufficient nor adequate to facilitate user retrieval of such information. The greatest challenge for obtaining relevant metadata remains getting user to perform the large amount of tedious manual work that is required. So far several attempts have been done to develop user-friendly GUIs to ease annotation tasks ([1],[2],[3]). However these applications still provide traditional single-user interfaces which are very useful for expert users in professional environments but which are still too tedious to be used by end-users for no-professional activities (such as annotating the pictures of some vacation). Based on this observation, the aim of our work is to provide a more engaging photo annotation application by exploiting multimodality along with the concept of collaborative annotation (related to experience sharing, storytelling and memory recollection) in order to make such activity more appealing.

2. WiiNote PROTOTYPE

WiiNote is a multi-user application allowing simultaneously annotations of digital pictures or videos (which need to be previously segmented into a sequence of still images). Users use Wii Remote devices (sometimes nicknamed “Wiimote”) [4] to select the picture or the region within a picture to be annotated.

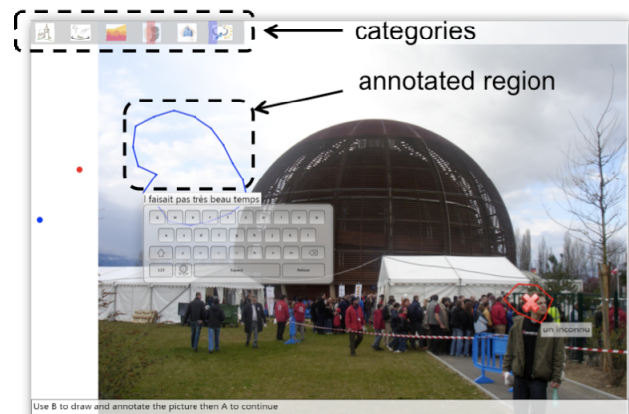


Figure 1 Wii Remote based textual annotation of picture

The Wii Remote device is first used to point toward and select the picture. Then, by pressing on the B button of the device, users can also select a specific region of the picture by free-hand drawing in order to provide a more granular annotation. Once the B button is released, a virtual keyboard and a text field popup (see Figure 1). The user can either edit the text using the keyboard or activating the

handwriting recognition panel (which slides out of the right side of the keyboard).

Moreover users can tag pictures by adding some vocal comments. So far, due to some hardware constraints, only one user at a time can add vocal comments. In order to activate the voice recognition, the user has to rotate the Wii Remote in a vertical position. This movement turns the color of the annotation text field into blue (see Figure 2, on the left) and activates a microphone that starts recording the user vocal comment. As soon as the user starts to speak, the voice is recognized using the Windows Vista Speech Recognition tool [5] and translated into text which is displayed in the annotation text field. The vocal comment is also stored as a wave file that can be accessed and replayed later on. The “red cross” icon in the middle of the annotated region of the picture (see Figure 2, on the right) allows user to delete already added notes.

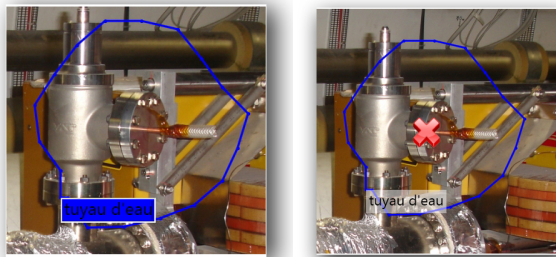


Figure 2 Voice-based annotation

As shown in Figure 3, each annotation is associated to a picture and contains three types of information (the coordinates of the annotated region, the user who made the annotation – each user has to identify him/herself at the start of the application - and the category the note belongs to in case of structured annotations). Multiple annotations can be associated to a picture.

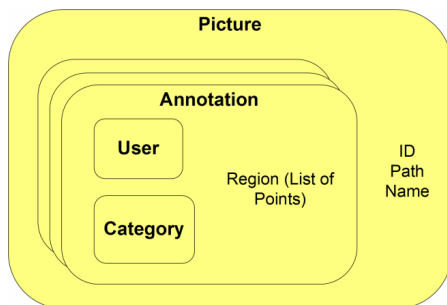


Figure 3 Picture annotation data structure

Annotations can be free or structured i.e. based on a domain specific data model such as an ontology. In the case of ontology based annotation, the main concepts of the ontology are displayed to the user on the interface as icons (see “categories” Figure 1, on the top). Annotations done by selecting one of these icons are directly linked to the related ontological concept and can be used later on to improve information retrieval by exploiting semantic reasoning. Different ontologies can be used according to the chosen domain of annotation (e.g. tourism, theatre, etc.). In our prototype we have focused on the tourism domain since this activity is strongly related

to the production of pictures as well as memories recollection and sharing. As shown in picture Figure 4, the application integrates different annotation formats (RDF, MPEG7, XML, etc.).

Once the annotation of a picture is completed user can continue the slideshow and annotate the other pictures. Annotations can be modified and enriched at any time later on either using the WiiNote application or other annotation tools which supports the aforementioned data format (such as CALISEMA tool [1]).

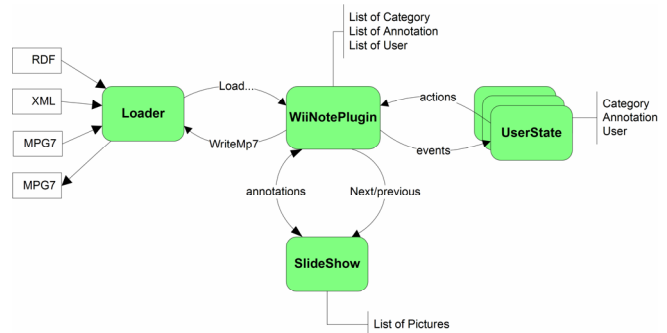


Figure 4 WiiNote application overview

3. CONCLUSIONS

This paper has presented the WiiNote prototype that allows annotating digital pictures in a multimodal fashion by combining both Wii Remote and voice based interaction. As next step we plan to integrate some gesture recognition functionalities in order to control slideshow, and to add multi-picture annotation functionality in order to annotate several pictures at the same time.

4. ACKNOWLEDGMENTS

This research work has been supported by RCSO-TIC within the framework of Memoria-Mea project. Our thanks to Julien Eberle for his valuable contribution.

5. REFERENCES

- [1] Sokhn, M., Mugellini, E., and Abou Khaled, O. 2009. Knowledge Management Framework for Conference Video-Recording Retrieval. In Proceedings of the 21st International Conference on Software Engineering and Knowledge Engineering (SEKE'09), Boston, USA, July 1-3, 2009.
- [2] PhotoStuff - An Image Annotation Tool for the Semantic Web, <http://www.mindswap.org/2003/PhotoStuff/>.
- [3] Petridis, K., Anastasopoulos, D., Saathoff, C., Kompatsiaris, Y., and Staab, S. 2006. MOntoMat-Annotizer: Image annotation, linking ontologies and multimedia low-level features. In proceedings of the 10th Intl. Conf. on Knowledge Based, Intelligent Information and Engineering Systems (KES 2006), Bournemouth, UK, October 9-11 2006.
- [4] Wii Remote device website, http://en.wikipedia.org/wiki/Wii_Remote.
- [5] Windows Vista Speech Recognition <http://www.microsoft.com/enable/products/windowsvista/speech.aspx>