Demo: Efficient Human Attention Detection in Museums based on Semantics and Complex Event Processing

Yongchun Xu[1)], Nenad Stojanovic[1)], Ljiljana Stojanovic[1)], Tobias Schuchert[2)]

[1)] FZI Research Center for Information Technology
Haid-und-Neu-Str. 10-14, 76131 Karlsruhe, Germany
{name.family name}@fzi.de
[2)] Fraunhofer Institute of Optronics, System Technologies and Image Exploitation (IOSB)
FraunhoferStr. 1, 76131 Karlsruhe, Germany
{name.family name}@iosb.fraunhofer.de

Abstract. In this paper we present a demo for efficient detecting of visitor's attention in museum environment based on the application of intelligent complex event processing and semantic technologies. The detection takes advantage of semantics: (i) in design time for the correlation of sensors' data via modeling of the interesting situations and annotation of artworks and their parts and (ii) in real-time for the more accurate and precise detection of the interesting situation. The results of the proposed approach have been applied in the EU project ARtSENSE.

Keywords: Sensor, Human attention, Complex Event Processing, Ontologies

1 Introduction

In this paper we describe a demo, which shows a semantic-based system providing personalized and adaptive Augmented Reality (A^2R) for the visitor, in which the digital contents react depending on the observed artwork and the user's engagement/attention state. In the demo we use semantic technologies for the correlation of sensors' data via modeling the so called interesting situation and use complex event processing to recognize the attention patterns in the event stream.

2 Overview

In order to enable an adaptive experience for the visitor to a museum, the demo is constructed around a four-phase OODA (Observe, Orient, Decide, Act) as shown on **Fig. 1**. In the Observe phase, our approach is concerned with the measurement of covert cues that may indicate the level of interest of the user. In order to consider how a user perceives an artwork, different sensors have been considered: The monitoring of visual behavior will allow the system to identify the focus of attention. The acoustic module should provide important information about environmental influences on patterns of visual attention or psychophysiology. Finally, a video-based hand gesture recognition provides an additional input modality for explicit interaction with the system (e.g., for selecting certain visual items, navigating through menus).

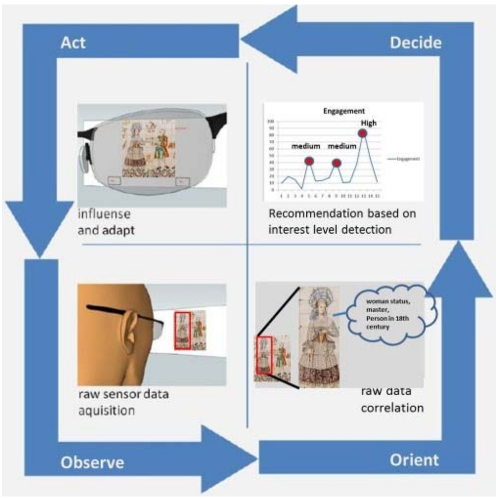

Fig. 1. OODA cycle

All data streams are collected and analyzed in real-time in order to yield a dynamic representation of the user attention state (phase Orient). In the Decide phase, covert physiological cues are used to measure the level of interest or engagement with artwork or with augmented content presented via the AR device. Based on the interpretation of this complex state, the provision of augmented content from a repertoire of available content is made. The presentation of selected content via the AR device (e.g. visual, audio) is subsequently executed during the final Act stage.

3 Semantic based attention

The challenge of this demo is to detect the attention of the visitor in the museum in real time according to the complex metadata. In most situations the attention of the visitors can be determined according to the gaze behavior of the visitors. In some cases the observed object is the attention object of the visitor, while in other cases the visitors pay attention to the information behind the observed objects. Thus, we distinguish between visual attention and content-related attention. **Fig. 2** summaries categories of attentions that are relevant in the museum context, including visual attention, content-based attention and audio disturbance.

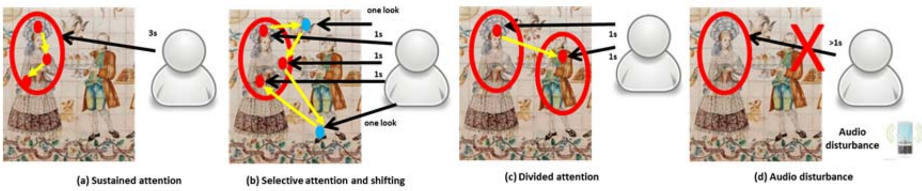

Fig. 2. . Different attention situations relevant for museums: (a) Sustained attention; (b) Selective attention and shifting; (c) Divided attention; and (d) Audio disturbance Implementation

Fig.2.(a) shows the sustained attention: the attention is focused over extended periods of times. Fig.2. (b) is then selective attention and shifting: The visitor changes his/her focus for a very short period and then focuses back on the previously selected artwork. Divided attention (see Fig.2.(c)) means sharing of attention by focusing on more than one relevant object at one time. When a visitor shifts his/her focus between two objects for certain time-period, we can say he/she is interested in the similarity of the objects. In the case that the audio disturbance happens during the visit and the

visitor reacts to this disturbance (see Fig.2.(d)), although the fixation of the visitor is detected, he/she pays no attention to fixated object but to the disturbance. Hence this situation should not be recognized as an attention.

4 The role of Semantic Technologies

The demo is based on knowledge-rich, context-aware, real-time artwork interpretation aimed at providing visitors with a more engaging and more personalized experience. Indeed, we propose to combine annotation of artworks with the time-related aspects as key features to be taken into account when dealing with interpretation of artworks. Thus, the aspects of the museums modeled by ontologies are classified into:

- Static aspects which are related to the structuring of the domain of interest, i.e. describing organization of an artwork and assigning the metadata to it;
- Dynamic aspects which are related to how a visitor's interpretation the elements of the domain of interest (i.e. artworks) evolve over time.

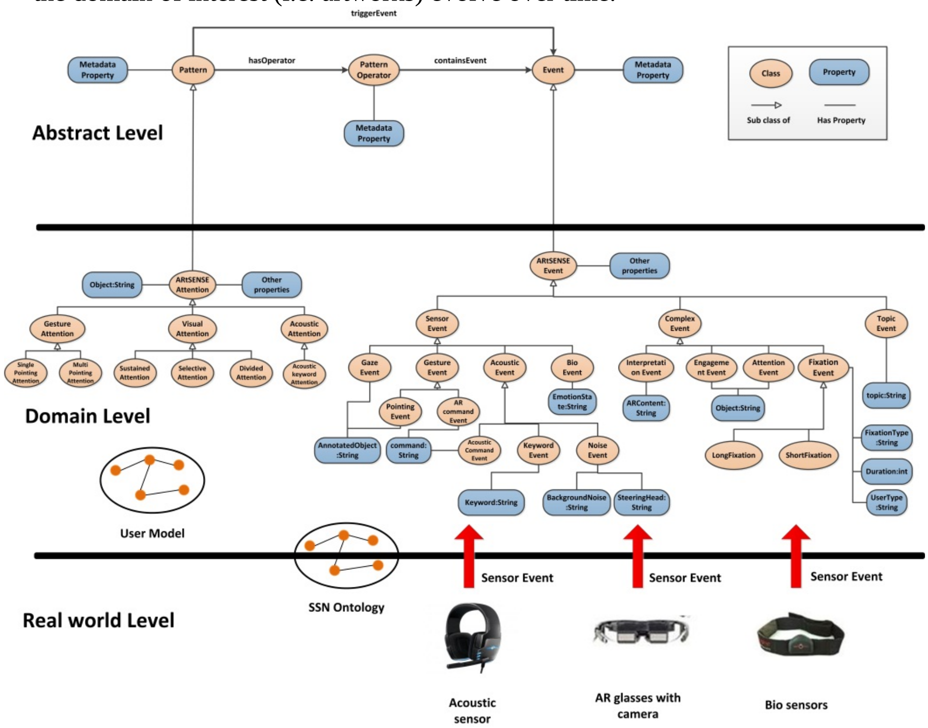

Fig. 3. Ontology for sensor data and patterns of human attention detection

Fig. 3 shows the ontology for sensor data and patterns of human attention detection that is used to describe what happened, when it happened, what the cause was and what the situation meant. We distinguish between three levels. Whereas the abstract level represents a meta-model for events and patterns, at the domain level[1] we defined the ARtSENSE pattern and event ontology, which describes the types of events and patterns relevant for the ARtSENSE context. For example, the most important situations of interest in ARtSENSE are attentions (see section 3), which are modeled as a pattern hierarchy. The real world level describes the sources of the events. In the case of the ARtSENSE these sources are sensors including acoustic sensors, bio sensors and camera sensors, which can be modeled by the existing sensor ontologies, such as SSN Ontology[2].

5 Demo setting

The demo will be performed using following hardware equipment:

- A Poster of Valencia Kitchen in MNAD (Museo Nacional de Artes Decorativas, Madrid Spain) as artwork
- Vuzix Star 1200 AR glasses with camera
- M-Audio Fast Track Pro audio card and BEYERDYNAMIC MCE 60.18 mic
- Zephyr HxM Bluetooth Heartrate sensor

6 Demo Implementation

Fig.4. shows the architecture of the ARtSENSE system. The system consists of sensors, sensor adapters, publisher/Subscriber, knowledgebase, metadata annotation tool, ETALIS and Interpretation. All components communicate through ActiveMQ by publishing and/or subscribing to events.

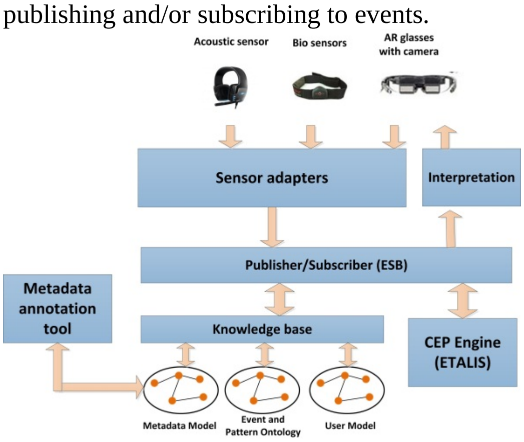

Fig. 4. System architecture

The following sensors are used: see-through glasses with integrated camera that can track the gaze of visitors and display the augmented reality (AR) content to visitors; acoustic sensor senses the acoustic information surrounding visitors such as environment noise or the content that visitors are listening to, and bio sensor observes the biological signals of visitors like heart rate.

A recent live demonstration of the system during a workshop in the Louvre museum can be found at http://www.youtube.com/watch?v=BnbGllVQMYQ.

[1] The user model is used to model parameters of visitors such as age, gender, etc.
[2] W3C Semantic Sensor Network Incubator Group, *Semantic Sensor Network Ontology*