Proposal for the BSc. AI 2013, UvA

O.R. Jozefzoon

Transformed

Overlay in AR

Visual marker

Conceptualized

Textual data

We present you with the Museum App[[1]](#footnote-1), an enhanced museum guidance system that uses trigger points to enrich the user with content during a tour around the city of Amsterdam. After experiencing the tour first hand we have decide to return to the museum to inform them about out thoughts and findings. Here a conversation with a member of the museum experience developing staff led to the conclusion that an added feature was needed to improve the user experience of the tour. This is where the idea emerged of transforming the existing data from the tour in a, from that point to be determined, certain manner to add an augmented reality experience to the existing tour(s).

**Research question.**

**The transformation of contextualized textual architectural concepts into visual markers to be presented as overlays in AR environments.**

Many applicants of location-aware applications propose the use of trigger zones to activate content and the use of servers to perform object recognition. Even though, location-awareness by the use of trigger zones is a highly efficient manner to display or use content, this dependency does not allow much room flexibility of the application. For example, an application does only work when it is near or on a trigger zone and does not work when a user is not within the range of the zone and also does it not work when a trigger is not provided. Furthermore, the use of a server creates a dependency from an external source, but with the improvement of mobile devices and object recognition algorithms this will not be necessary. Our aim is to reduce the dependency of the triggers by implementing an application that provides a user with content based on an observation of an object in the real world and on abstraction from a textual source. The textual data is abstracted from a database and translated in to useful description, based on architectural description from AAT[[2]](#footnote-2) of object/buildings, for our application. We propose a new approach in which our content is activated by the input from the data (text/image) and afterwards verified using geo-location and compass data from the mobile device.

**Literature review**.

Lameira et al. (2011) propose an application for real time object recognition without the use of a server. Even though, we are not completely ignoring the value of a server, we also aim to limit the use of the server to a maximum and depend for the biggest part of our computations on the power of the mobile device. The reason for this decision is based on the fact that mobile device, nowadays, are improved, have increased computational power and are, thus, able to perform complex computations within a reasonable amount of time. This decision allows us to limit the use of a server to perform only the act of a storage facility for our required data and make our application as independent and as fast as possible. Independent in the sense that the application is not dependent of an internet-connection to connect with the server, once the data is stored, and as fast as possible in the sense that the computation does not depend on an external source, but is almost completely self-sufficient.

Amlacher et al. (2008) demonstrate in their paper “geo-indexed object recognition from experimental tracks and image captures in an urban scenario, extracting object hypotheses in the local context from both (i) mobile image based appearance and (ii) GPS based positioning”. For our research we propose to use geo-indexed verification to check whether or not the assumptions, based on extracted object hypotheses conceptualized from text, made by our application are accurate. The result from the demonstration by Amlacher et al. (2008) provides us with a basis assumption to use geo-indexed data to improve the accuracy and coverage of our application.

According to a research conducted by Föchler et al. (2005) has shown that the use of a single-layer perceptron neuronal network is sufficient to achieve an object recognition accuracy of over 90%. This accuracy rate is achieved with the use of a mobile phone. Following the findings of Föchler et al. we aim to reach the same probability on a mobile device. Even though, we are not yet certain if the use of a single-layer perceptron neural network is the most efficient manner to achieve the desired result in combination with our application such a high recognition rate is certainly desirable for the use of our application. We propose the use of an object recognition algorithm, but without the use of an existing database of images. The application is supposed to recognize the object in real-time within a reasonable amount of time based on a conceptualized description of our textual data.

The client-side tracker as proposed by Gammeter et al. (2010) in ‘Server-side object recognition and client-side object tracking for mobile augmented reality’ is a valuable device, which we will use to memorize the position of an object even when out of screen, using visual and sensor based cues. This client-side tracker is necessary because the user will be guided towards an object to be recognized. Gammeter et al. does not use GPS information for object recognition and tracking. We, on the other hand, will use GPS information for the verification of the results presented by the application and, combined with the compass on the mobile device, as a device to guide the user in to the desired direction. Our use of the GPS and Compass information will provide the application with a more robust base to guide the user and to co-operate with the object recognition algorithm.

**Method and approach**.

In order to achieve our goal, of creating the correct conceptualization of our textual data for the use of object recognition, we will use a method of implementing and testing. In the first stage we will decide for a manner to conceptualize the textual data to be useful for our application. The conceptualization will be determined with the use of descriptions from the AAT and manually selected locations. The locations will serve as a reference pinpoint on which we will base our conceptualization. Once the correct representation of our data is determined we will start with the implementation of our application. The application will use the conceptualized data to present the user with the correct or clear description of an object in the real world. When it seems as if the correct description is not provided, we will explore the added feature of using the GPS and/or compass data for an extra verification. This verification might be necessary when the user, for example, does not find the object described or the application does not find a matching description for the object presented by the user.

**Evaluation**.

In this section we will describe how will the results of the research be evaluated? In a way this can be regarded as part of the method/approach, but it is important and therefore requires independent attention. Ask the question of how proof can be given that the results obtained are viable.

The main evaluation will occur after finishing the implementation of the application. However, we will incorporate unit test in our code to make sure that we are not providing our application with incorrect data our malicious chunks of codes. The unit tests are necessary for our application, because malicious codes may corrupt the accuracy and might negatively affect the coverage.

For our main evaluation real world testing will be necessary. This means that we will test if the application works as desired on the selected locations. If we are not provided with the desired results then we will have to decide whether the source of the error is our conceptualization of the text or whether the error could be found at the side of the object recognition algorithm(s) used. In case of the first we have to reconsider our conceptualization, in case of the latter an obvious reconsidering of algorithm will be necessary. We will probably mainly revise the algorithms, because the conceptualization will be given a fair amount of (re-) consideration and once the decision is made for an acceptable representation it is our task to implement a suiting algorithm.

**Plan**.

In this section (see figure 1) we will explicate the order in which the activities following from the above will be carried out in time. As well as how much time (effort in terms of hours or days) will be allocated to each task. A gantt chart is a very typical way to articulate plan. Further issues include the following. Are there multiple ways in which the activities can be organised? What are the landmarks, what decision options do these have, and on what criteria will that be evaluated? Is there a critical path? Can it become a problem? How can it be circumvented (is there a contingency plan)?

****

Figuur 1. Planning BSc scriptie 2013 - Jozefzoon 1

1. Created by The Amsterdam Museum [↑](#footnote-ref-1)
2. The descriptions are based on the Art & Architecture Thesaurus® Online vocabulary of the Getty Research

   Institute.

   [Source: http://www.getty.edu/research/tools/vocabularies/aat/] [↑](#footnote-ref-2)