

GeoMedia – A Framework for Producing Interactive, Geo-Referenced Video Contents

Andrei Papliatseyeu¹, Givoanni Cortese², and Oscar Mayora-Ibarra¹

¹ CREATE-NET, Via alla Cascata 56C, Trento Italy

² Interplay Software, Trento, Italy

{andrei.papliatseyeu,g.cortese}@ipsoft.it,

oscar.mayora@create-net.org

Abstract. This work-in-progress paper presents an approach for interactive geo-localized multimedia services. The GeoMedia platform will provide richer user experience within the context of city-level cultural or sport events. This will be achieved by means of media aggregation using geo-references and multi-view of pictures and videos aligned with the geospatial coordinate system. The composed imagery/video will be made available for the event participants via their mobile devices, such as smartphones. This will enable the users to reconstruct recent scenes and situations at certain locations of the medium-scale event, thus augmenting the participation experience.

Keywords: Geo-referenced content, location based services, context awareness.

1 Introduction

The advances of scientific knowledge in the fields of multimedia information management and augmented reality, and the increasing spread of enabling technologies (such as broadband connectivity and mobile handsets with advanced multimedia features) creates the adequate conditions for the development of new services for the web based on interactive video contents. In particular through images geo-registration techniques, it becomes possible to create interactive video geo-located, where the user can explore the space / scene depicted in the video, or change the observation point, or obtain information about a monument in the video by clicking on it [1]. This enables users to live a stronger experience of presence in the scene and the realism of it.

Mobile phones and smartphones are obviously ideal devices for capturing and utilizing multimedia content strongly related to the physical world [2, 3]. The application of interactive geo-located multimedia services has a large target market, and gathers different market sectors and applications (such as promoting tourism, cultural heritage, e-commerce, e-learning, asset management and public works) [4]. The growing competition in the tourism industry and the need for promotion and enhancement of local resources creates a continuous demand of innovation in the modalities for communication and marketing on the web. In the scope of activities

relevant to the promotion of tourism, an interesting scenario is that of sports and cultural events. Organizers of such events often require advanced tools for making information services or providing advanced collaboration tools benefiting the organization of participants and the spectators. To date, the state-of-the-art knowledge and current tools available for implementing interactive geo-localized multimedia services are very much at the research level, or still too expensive for widespread use. This paper presents an approach to make more immediate use of current knowledge and techniques for geo-localized applications in targeted contexts (such as massive sports or cultural events).

2 GeoMedia Scenario

The context which is taken as reference for the GeoMedia scenario is that of a massive event in a medium-size city (e.g. sport, cultural, touristic event). In this scenario, the organizers of the event intend to create services focused on provisioning multimedia content for the different participants and for spectators and tourists. An example of this could be the next Nordic Sky World Championship to be hosted in Trentino province of Italy. For this event it is planned that the main competitions will be on an extended area, partly or wholly equipped with network infrastructure such as Wi-Fi, or 3G. The organizers of the event will have an infrastructure capable of providing access to the web to visitors and other general users and an intranet with restricted access for specialized users (journalists, premium users). This reserved access network will provide special information and other added value services to participants of the event. One of such geo-localized services is the access to geo-referenced media. In this way, the special users will be able to reconstruct the recent events and situations that happen at certain locations of the sky track (e.g. pictures/videos of participants in that point) at different moments.

3 GeoMedia Framework

GeoMedia framework consists of a set of methods, algorithms and architectural approaches allowing organizers of city-level events to benefit the participants with the possibility of creating interactive video content using a new type of media aggregation including geo-references and multi-view. In this way, GeoMedia aims to communicate to the viewer a sense of involvement and engagement with the system and generally a richer sensory experience. The main feature of the contents produced with GeoMedia will be the possibility of being created from “imagery” being captured from photo / video cameras and sensors for positioning and orientation and recorded and aligned with a geospatial coordinate system. Thanks to this feature the presentation of content in GeoMedia will enable the multiple view of pictures considering the different geometry of the physical space represented in the video. GeoMedia enables the composition and production of imagery and interactive video, which are then made available to participants of the event via their smartphones. Even

during the delivery / use of content, the system's ability to acquire and interpret the position in space of the user, and therefore optimize the presentation of content (Context-awareness) will be one of the main features of GeoMedia framework.

The following figure schematically represents the GeoMedia system (Figure 1). The main components of the system highlight the different functionalities enabled by the system such as the presentation of videos in 3D environment, geo-referenced registration of video/images, scene analysis, etc) and architectural approaches (e.g. localization methods, techniques for determination of the camera position, encoding of geospatial information into the video stream, etc.).

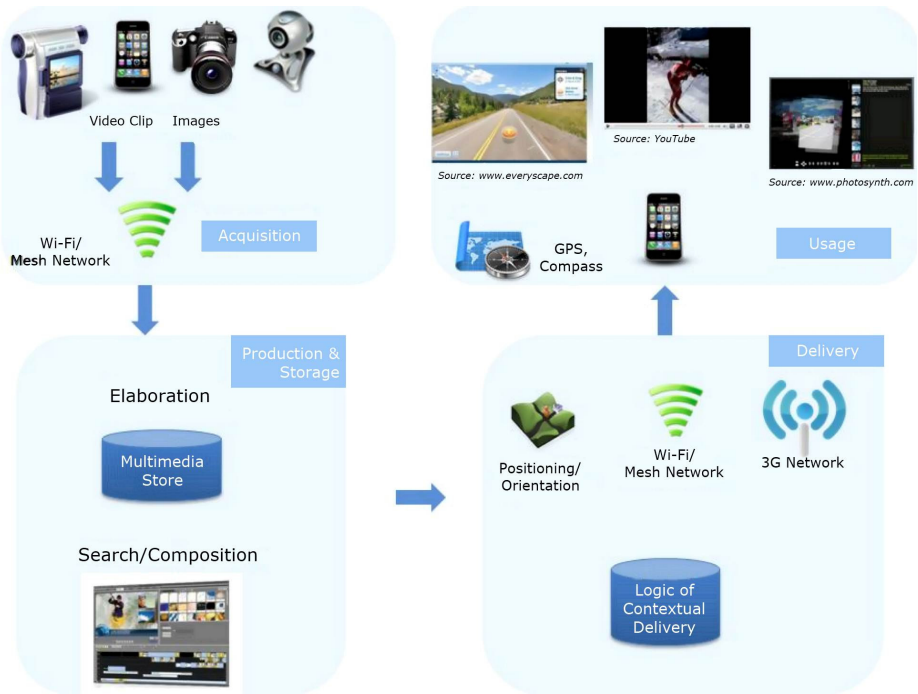


Fig. 1. GeoMedia Framework

The main components of GeoMedia framework include:

1. Acquisition Component
2. Pre-processing
3. Delivery
4. Presentation

Acquisition Component: In the acquisition phase, images and videos are acquired by consumer devices such as mobile phones, webcams, photo/video cameras, etc. These devices will be enriched when needed with other sensors for the acquisition of position and orientation information (GPS, compass, tilt sensors). The acquisition of

audiovisual content can occur either in the stage of organization of the event (for example, the creation of an interactive video that documents and allows to visualize the race route) and during the event itself (e.g. photo/video of the races). The multimedia information and location is transmitted to the content management system, using a network (e.g. mesh or Wi-Fi), or via removable storage devices.

Pre-processing: The pre-processing and analysis software processes the images and position data, extracting useful information for further processing. The geolocalized images are stored in a scalable repository equipped with advanced search functionality to find pictures on the basis of geospatial criteria. A set of software tools (Authoring framework) thus allows authorized users to access and edit the material of the CMS, and create presentations that will be published and made available to the end users. Some examples of authoring capabilities that will allow manipulation of multimedia and creation of interactive videos include: the creation of clips, composition, textual or graphical annotations (e.g. graffiti), creation of hyperlinks to other content media.

Delivery: The front-end towards the end user is represented by a delivery service that will be optimized for the transmission of interactive video to mobile terminals and the delivery of location-based information.

Presentation: The presentation at the end user side will be accessible through an intranet portal available from dedicated fixed terminals, interactive kiosks or mobile terminals. The information will be presented using video players enabled with graphics and control features for advanced interaction.

Finally, an appropriate network architecture consistent with the envisioned application scenarios will ultimately provide the required infrastructure for the proper functioning of services.

4 GeoMedia Positioning and Geo-Tagging System

The GeoMedia system distinguishes two types of users: stationary and mobile ones. Stationary users are provided with a web interface, which displays a world map and a number of marked locations on it. In GeoMedia, a location is defined by a pair of coordinates and a name (e.g., “Shopping mall”). Each location has a number of virtual stickers associated with it. A sticker contains a multimedia file, its description and priority level. Stationary users can modify and add new locations, as well as upload new stickers.

Mobile users are equipped with location-aware devices that have Internet connectivity. The device periodically updates the system with the user’s current position, and receives a list of nearby locations and stickers attached to them. When a new high-priority sticker is detected nearby, it is automatically downloaded and presented to the user; low-priority stickers can be received manually by user’s request. In order to avoid abuse, the user can

configure criteria of automatic downloading. The activity inference and prediction module is responsible for building user movements profile and using it to predict future context of the user.

The architecture of the location tagging system is presented on Figure 2. Both stationary and mobile clients communicate with the system by HTTP via web front-end. The Map Builder module takes map images from an online service, e.g. Google Maps, and overlays it with location markers. When the user selects a location, a list of associated stickers is displayed. The user can then upload a new sticker, and it will be attached to the selected location. For the mobile user, the process is more straightforward: given the coordinates, the system generates an XML file with a list of locations and associated stickers in some vicinity of the user; basing on this list the mobile client downloads and displays stickers. However, the mobile part of the system presents a number of challenges, such as localization accuracy, limited resources, slow connections, privacy concerns, etc.

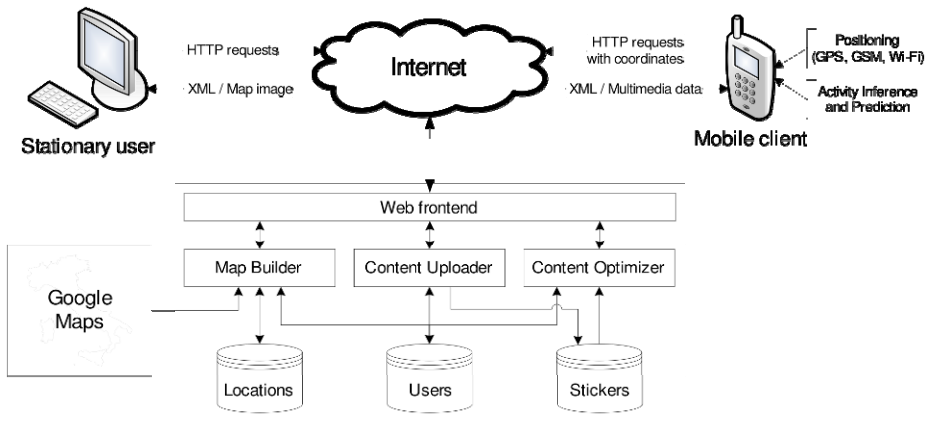


Fig. 2. Localization – Tagging System

To test the idea, we have implemented a single-user proto-type of the proposed localization-tagging system. Server part is running Apache Tomcat and Struts as web front-end; MySQL handles data management and storage. Stationary user’s interface (see Figure 3) is based upon Google Maps API, using JavaScript for GUI management and AJAX for asynchronous communication with the server. Side panel displays the list of stickers associated with selected locations, and provides an interface for sticker upload. Only images are displayed directly in the browser, other types of content can be down- loaded and viewed by third-party applications.

The mobile client is implemented in Java 2 Micro Edition (J2ME) and runs on any MIDP2-enabled phone (Nokia E61i in our case). The current implementation uses Java Location API and GPS to obtain user position; later on we plan to adapt PlaceLab framework. User position is sent to the server via Wi-Fi, 3G or GPRS connection every 5 sec., which provides a reasonable trade-off between network load and system responsiveness. Privacy of the positioning data is secured by JSR-179 (Java Location API) regulations [5].

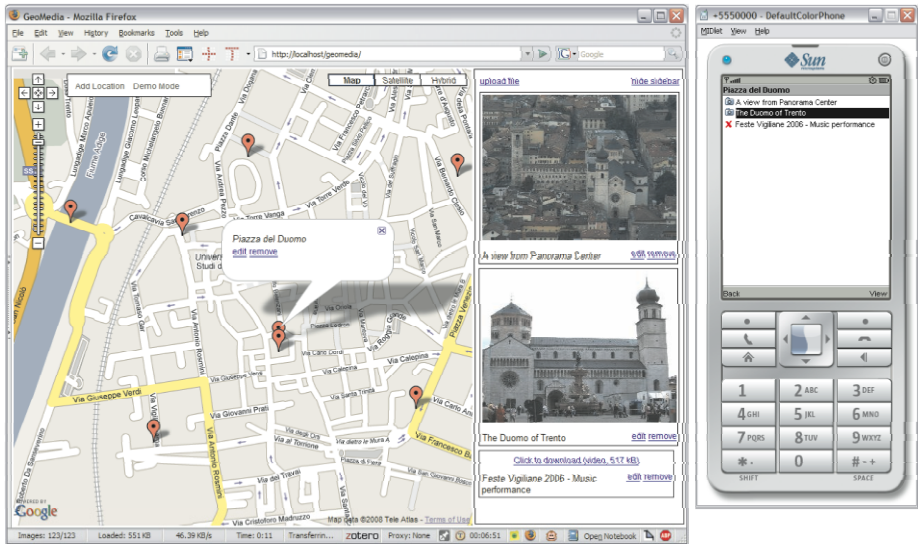


Fig. 3. Preliminary Web and Mobile Client Interfaces for Location-Tagging

5 GeoMedia Envisioned Services

The envisioned services of GeoMedia platform will be illustrated through some examples of possible applications inspired in the sports scenario mentioned above. The main services are divided according to the scenarios defined for both, the end users and the content creators.

Services for Content Consumers:

"Navigator" – Joe is a journalist who follows the event. Using a service offered by the Organizing Committee Joe can see the road races, represented in the form of an interactive video with navigation controls. The videos are annotations and "buttons" that allow him to view useful information, documents, pictures, videos, records of the competitors going through that location, etc. A variation of this service can consent to access information about nearby attractions, shops, restaurants, services and events in the area.

"What am I seeing" – Mary is watching the competitions close to the racetrack. She points an area in the racetrack with her phone video camera and the display gets a series of augmented descriptive information of the passage. On the image that is seeing are also offered a list of images / video clips relating to that passage (at precedent days, etc.).

"Media Alert" – Jan is an athlete in training. When approaching a specific area he is notified on his smartphone about the availability of video information that his trainer had previously registered on the site. By orienting the smartphone in different

directions in the same site a series of images are proposed to him highlighting the most relevant viewpoints.

Services for Content Producers:

"Annotation and Composition" – Joe is a journalist. Connecting to the CMS he can get some of the video clips about the races. Using the authoring tools of the system he can create a new interactive video composed of several connected clips and can write textual or graphical information associated with specific segments of the videos, create links associated with some items shown in the videos which refer to other media in the CMS or in external web pages. Once he is done with his video he can publish it to the Event webpage.

"Location Tagging" – The event organization wants to communicate useful information for athletes when they enter a certain area. In order to do this, they use the navigation tool for positioning in the respective area and tag it with the content. Otherwise, some athlete or trainer points and takes a photo/video on a certain location and post it to the system for delivering to a specific colleague or group of colleagues when reaching the tagged area.

6 Concluding Remarks

In this paper we presented a work in progress dedicated to building a framework for interactive geo-referenced multimedia provisioning. The main feature of the GeoMedia platform will be the possibility to create the content from "imagery" captured from photo/video cameras and align it with a geospatial coordinate system using positioning and orientation sensors. Thanks to this feature, the GeoMedia will be able to present the content from multiple views considering different geometry of the physical space represented in the video. The paper also described a number of location-based services which GeoMedia framework will enable.

References

1. Google Maps Street View. Online, <http://books.google.com/help/maps/streetview>
2. Ludford, P., Frankowski, D., Reily, K., Wilms, K., Terveen, L.: Because I Carry My Cell Phone Anyway: Functional Location-Based Reminder Applications. In: Proc. of CHI 2006, pp. 889–898. ACM Press (2006)
3. Marmasse, N., Schmandt, C.: Location-Aware Information Delivery with ComMotion. In: Thomas, P., Gellersen, H.-W. (eds.) HUC 2000. LNCS, vol. 1927, pp. 157–171. Springer, Heidelberg (2000)
4. Schiller, J.H., Voisard, A.: Location-based services (2004) ISBN: 978-1-55860-929-7
5. JSR 179: Location API for J2ME. Online, <http://jcp.org/jsr/detail/179.jsp>