

---

# ZoneTag: Rich, Community-supported Context-Aware Media Capture and Annotation

**Shane Ahern, Simon King, Mor Naaman, Rahul Nair, Jeannie Hui-I Yang**

Yahoo! Research Berkeley  
1950 University Ave, Suite 200  
Berkeley, CA 94704, USA  
<http://yahoresearchberkeley.com>  
{sahern simonk, mor, rnair, jeanniey}@yahoo-inc.com

## **Abstract**

ZoneTag is a rich mobile client that enables context-aware upload of photographs from cameraphones. In addition to automatically supplying location metadata for each photograph, ZoneTag supports media annotation via context-based tag suggestions. Sources for tag suggestions include past tags from the user, the user's social network, and the public, as well as names of real world entities such as restaurants, events and venues near the user's location. A seamless interface makes it easy to assign tags to a photo, forming the basis for a richer personal media organization system. We believe that lowering the barriers to tagging has great potential for effective retrieval.

## **Keywords**

Photo labeling, tagging, multimedia management, retrieval, tag suggestions, context-aware, location-aware

## **ACM Classification Keywords**

H.5.1 [Information Interfaces and Presentation (e.g., HCI)]: Multimedia Information systems

## Introduction

Consumer photography has made it exceedingly simple for people to capture images, which they do at an ever-growing rate. The growing rate of photo capture is driven by the proliferation of capture devices (such as digital cameras and cameraphones) as well as decreasing storage costs. At the same time, however, creation of semantic metadata about the photo content remains an elusive goal. This type of metadata is required in order to facilitate retrieval from large collections of photographs and media. Completely automated tools, though improving, have not proven sufficient in providing complete and accurate semantic metadata. Manual annotation interfaces, even on desktop computers, remain time-consuming, often presenting the user with an overwhelming number of photos to annotate. Semi-automatic approaches based on content analysis and often on contextual information have made some headway. Recent emphasis on contextual metadata for photo organization is indicated by a growing number of other projects (e.g. Meaning[1], Context Watcher[2], Lifeblog[3]) that automatically associate contextual metadata with media captured on cameraphones. ZoneTag is unique in its use of this automatic metadata to generate tag suggestions. We believe that a semi-automatic approach providing tag suggestions at the point of capture is a promising step towards lowering the barrier to participation in a media annotation system.

ZoneTag supports location-based sharing of metadata. Allowing metadata to be shared creates an opportunity to leverage community effort to benefit each of the individuals within the community. Additionally, sharing tags throughout a community presumably leads to some degree of tag convergence, making it easier for

individuals to retrieve not only their own media, but others' as well. In shared-metadata systems we can consider three levels of user participation: creators, synthesizers, and consumers. The challenges in designing a metadata-entry system are making it easy for creators to generate metadata, easy for synthesizers to use it, and beneficial for consumers who do not contribute at all. Technically this means easy input (especially challenging on a mobile device) for the creators, good algorithmic support for tag suggestions for the synthesizers, and techniques for leveraging this in a UI that does not require participation.

ZoneTag supports easy photo capture and upload to Flickr [4] in as few as two keystrokes from the mobile device. The immediate benefit for all consumers is that the photo is automatically annotated with location data for browsing by location even without any user effort. For synthesizers, the system offers the user the opportunity to add tags, selected from a location-sensitive searchable list. For creators, tags can be entered on the phone (assisted by T9 predictive text and auto-complete on all previously-selected tags), or entered on the web.

## System Description

The two primary components of ZoneTag are the client application, running on Nokia Series60 cameraphones, and the ZoneTag server, a PHP/MySQL application providing location translation and suggested tags to the client as well as processing uploaded images and metadata from the client and passing the images and tags to Flickr.

The ZoneTag client runs as a background process, monitoring the cell tower to which the phone is

currently connected (which provides a rough location) and possibly communicating via BlueTooth to a GPS device, providing much more exact location information (and also providing a mapping between cell tower and physical location.) When the phone switches cell towers (or every 10 minutes if stationary) the client contacts the server for a new set of suggested tags, appropriate to the new context. Suggested tags are pre-fetched because network latency is too great to provide acceptable user experience fetching tags at the moment a photo is captured. The decision to pre-fetch was based on previous work on the Mobile Media Metadata (MMM) system [8] which provided some degree of context sensitive tags but received several complaints about the latency in a browser-based tagging interface [9]. When a photo is captured, the ZoneTag application comes to the foreground of the phone's user interface and provides an integrated tagging interface that allows the user to quickly annotate and upload the photo. In addition to the user-selected tags for a photo, the ZoneTag client sends the server current location data (cell info and GPS if available.)

When the server receives a new photo it translates the location information from the client into human-readable labels (i.e. city, state, country, ZIP code) which propagate to Flickr as tags as well as location data. If there is no known location mapping for a given cellID, users can 'teach' the server new locations via a web interface. These new locations are then propagated throughout the user's social network (as determined by Flickr family, friends, and contacts) with preference given to locations obtained from users with less social distance. More details about the cellID to location translation can be found in [5].

ZoneTag provides tag suggestions based on the photo location, the the tag history in this location from the user, their social network, and the public. Suggested tags are displayed on the handset ordered by the likelihood of being used in the current location context. This likelihood score is based on the number of times a tag has been used in the current location, weighted according to the social distance from the user to the tag creators (or synthesizers who have subsequently used the tag.) The location(s) in which tags have previously been used also factors into the weighting with tags used in nearby locations weighted less than tags used in the exact location.

Suggested tags are drawn not only from tags entered by other ZoneTag users, but also from external geo-referenced data sources, including Yahoo! Local, Upcoming.org, and geo-tagged Flickr photos (not necessarily created by ZoneTag users.) Note that while tags from these external sources often describe features of the location itself (restaurants, tourist attractions, etc.), tags from other ZoneTag users do not necessarily describe the location, but often are names or adjectives describing objects frequently photographed near the location. For example, a user may get tag suggestions such as 'whiteboard', 'meeting', or coworkers' names at work and 'home', 'dinner', or pets' names as suggestions at home.

In addition to providing suggested tags to the ZoneTag client on the handset, the system provides a web interface to tag suggestions, linked from the Flickr website. The ZoneTag server periodically crawls ZoneTag images on the Flickr website for tag changes made via Flickr's web interface, which can impact the suggested tags propagated back to the handset client.

This feature allows users to enter new tags on the web and simply select these tags later on the mobile interface, without the hassle of phone-based text entry.

### **User Experience**

Using the ZoneTag application, a user can upload a newly captured image from her cameraphone to Flickr in two quick clicks. After an image is captured, ZoneTag displays a "Post to Flickr?" dialog over the newly captured image. At this stage, the user can review the image and decide whether she wants to upload it. If so, more options are shown, but the user is given the option to upload the image immediately, keeping the settings and tags entered for the previous photo.) The option screen allows the users to control various features, and add or change the metadata tags associated with the image. The controls include the option to set the image privacy settings; the option to specify a title for the photo; and the option to enter the photo's location data if it is unknown. Most importantly, the user has an option to select or type in tags (textual labels) to appear on the Flickr photo page, using the tag suggestions as described above.

### **Deployment**

ZoneTag is currently deployed to 500 users (130 active) taking photos in over 50 countries. ZoneTag is available for download for Nokia Series 60 2<sup>nd</sup> and 3<sup>rd</sup> Edition phones, as well as select Motorola phones, at <http://zonetag.research.yahoo.com>.

ZoneTag powerful context and location-based tag suggestions and capture features allow the system to

serve as an innovative platform to investigate various aspects of media capture and annotation systems, including user privacy [6] and user motivations for tagging [7].

### **References and Citations**

- [1] <http://meaning.3xi.org/>
- [2] <http://portals.telin.nl/contextwatcher/>
- [3] <http://europe.nokia.com/nokia/0,,71739,00.html>
- [4] <http://www.flickr.com>
- [5] Shane Ahern, Marc Davis, Simon King, Mor Naaman, Rahul Nair. Reliable, User-Contributed GSM Cell-Tower Positioning Using Context-Aware Photos. In The Eighth International Conference on Ubiquitous Computing (UbiComp '06), September 2006, Orange Country, CA, USA.
- [6] Shane Ahern, Dean Eckles, Nathan Good, Simon King, Mor Naaman, Rahul Nair. Over-Exposed? Privacy Patterns and Considerations in Online and Mobile Photo Sharing. In Proceedings, CHI 2007, San Jose, CA, USA.
- [7] Morgan Ames and Mor Naaman. Why We Tag: Motivations for Annotation in Mobile and Online Media. In Proceedings, CHI 2007, San Jose, CA, USA.
- [8] Davis, M., King, S., Good, N., and Sarvas, R. From Context to Content: Leveraging Context to Infer Media Metadata. In Proc. MM 2004. ACM Press (2004), 188-195.
- [9] Sarvas, R., Herrarte, E., Wilhelm, A., and Davis, M. Metadata Creation System for Mobile Images. In Proceedings of the Second International Conference on Mobile Systems, Applications, and Services (MobiSys2004) in Boston, Massachusetts. ACM Press, 36-48, 2004.