Demo Abstract:

PhotoNet+: Outlier-resilient Coverage Maximization in Visual Sensing Applications

Md Yusuf S Uddin
University of Illinois
Urbana, IL 61801, USA
mduddin2@illinois.edu

Md Tanvir Al Amin
University of Illinois
Urbana, IL 61801, USA
maamin2@illinois.edu

Tarek Abdelzaher
University of Illinois
Urbana, IL 61801, USA
zaher@illinois.edu

Arun Iyengar
IBM T.J. Watson Research
Center, NY 10598, USA
arun@us.ibm.com

Ramesh Govindan
University of Southern
California, CA 90089, USA
ramesh@usc.edu

ABSTRACT

This demonstration illustrates a service for collection and delivery of images, in participatory camera networks, to maximize coverage while removing outliers (i.e., irrelevant images). Images, such as those taken by smart-phone users, represent an important and growing modality in social sensing applications. They can be used, for instance, to document occurrences of interest in participatory sensing campaigns, such as instances of graffiti on campus or invasive species in a park. In applications with a significant number of participants, the number of images collected may be very large. A key problem becomes one of data triage to reduce the number of images delivered to a manageable count, without missing important ones. In prior work, the authors presented a service, called PhotoNet [2], that reduces redundancy among delivered images by maximizing diversity. The current work significantly extends our previous effort by recognizing that diversity maximization often leads to selection of outliers; images that are visually different but not necessarily relevant, which in fact reduces the quality of the delivered image pool. We demonstrate a new prioritization technique that maximizes diversity among delivered pictures, while also reducing outliers.

Categories and Subject Descriptors

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

Keywords

Redundancy reduction, outlier detection, visual sensing

1. INTRODUCTION

We define a participatory camera (sensor) network as one where participants contribute pictorial data, either on their own initiative or through participation in a corresponding data collection campaign. For example, in the aftermath of a natural disaster, relief workers and other first responders might survey an area in search of damage that is then pictorially documented and reported. Alternatively, residents of a neighborhood might pictorially document issues that require attention (e.g., graffiti on walls, trash piles, or hazardous potholes). Yet a third application might be to compile a list of most visited tourist landmarks from pictures contributed by local tourists. Participatory camera sensing applications are made popular by the vast proliferation of cameras and camera phones in the possession of the average individual, not to mention the richness of information contained in pictures compared to other sensing modalities [1].

Our camera sensing service runs on participants' phones (the clients) and on a destination server (the collection point). When pictures are taken using our application, they are locally stored on the phone. When two participant phones meet, they may gossip by exchanging a portion of their pictures. Similarly, when a phone connects to the destination server it uploads a portion of its pictures. The contribution of the service lies in prioritizing transmission of pictures both when two phones meet or when a phone meets the server, such that the most *representative subset* is sent (instead of sending all), in order to conserve resources. Resources may need to be conserved for many reasons. For example, participants, who upload pictures from their mobile phones, may have to pay for their data plans. Network resource constraints may also require data triage to fit the available capacity. In DTN-style communication and military scenarios, groups of soldiers in the field may have only a low or intermittent bandwidth channel to a remote base.

We do not make inherent assumptions regarding the type of network in which our service operates. For example, it could be a star network, where all phones have a direct way of connecting to the server. Alternatively, it could be a DTN, where the primary data propagation occurs via phone-to-phone gossiping. Either way, the decision we are concerned with is which pictures to send in what order when two nodes meet (either two clients, or a client and the server).

In this demonstration, we show that algorithms that maximize diversity to improve coverage, such as those proposed in previous literature [2], favor outliers as opposed to more representative content. The main contribution of our new prioritization scheme lies in combining coverage maximization with outlier elimination, to handle picture sets of poor quality in participatory camera sensing networks.

It is worth noting, at this point, that outlier elimination is not always a goal in a participatory camera network. In some applications, such as anomaly detection, outliers are

IPSN'12, April 16–20, 2012, Beijing, China.
ACM 978-1-4503-1227-1/12/04.

in fact what carries the relevant information. For example, an in-store security camera might report the same view all night, except when an intruder breaks in. A frame with the intruder in view might be the outlier, but it is also the frame that contains the most interesting information. Our service considers a different type of applications, where a community of users document relatively static conditions in the environment, such as damage or points of interest. In such cases, one is not looking for anomalies in reporting, but rather for *representative* depiction.

2. THE OPERATING PRINCIPLE

The contribution of our new service lies in maximizing diversity while removing outliers in the delivered subset of collected images. An explicit goal is to estimate relevance of a picture to the mission *without having to understand the semantics* of what is in a picture, since this would be very complex and application-specific. The scheme has two components; one for maximizing coverage and one for outlier elimination, as discussed below.

Coverage maximization: Our scheme separates out the application-specific notion of "similarity" between images into the definition of a *distance metric*, $d(x, y)$, defined on any pair of images x and y to denote the degree of similarity in their visual content and metadata (such as location). The distance metric allows images to be represented as points in a logical multidimensional feature space, where the proximity of points designates information overlap between the corresponding objects. If two points lie very close to each other, they are partially redundant. We further assume that there exists a certain distance threshold beyond which there is no information overlap. Let this constant be τ. Hence, it is useful to imagine that each object logically covers a hyper-sphere with radius $\frac{\tau}{2}$ so that the spheres of two objects overlap when their distance is smaller than τ. The volume of a sphere is called the *coverage* of the object.

Note that, due to overlap, the total coverage of a set of objects is generally less than the sum of the coverages of the individual objects. The total coverage of all objects in a set can thus be treated as a quantitative estimation of the diversity of the set. The diversity maximization problem is then to chose a subset of objects whose total coverage is maximum, subject to an aggregate resource constraint (e.g., storage capacity) that limits the number of objects chosen.

In practice, pictures taken by participants would typically fall into groups (each group representing pictures of the same scene at the same place), such that logical distances between pictures within the same group (or cluster) are much smaller than those among different groups. This naturally leads to partitioning objects into a set of *clusters*. Our service implements a coverage-maximizing algorithm for picture selection that leverages clustering to reduce problem complexity. The details of the algorithm are beyond the scope of this abstract.

Outlier elimination: It turns out that clustering offers an elegant way of separating the concern of outlier detection from the concern of diversity maximization. Intuitively, by assigning appropriate *relevance weights* to clusters, we can first get rid of low-ranked clusters (the outliers) to address relevance, then collect objects from the remaining clusters, thereby maximizing diversity for non-outlier clusters.

Short of "understanding" each picture, we can only approximately *estimate* relevance, which we do from the behavior of data collection agents themselves. Presumably, they are motivated to collect relevant information. Hence, if more sources report an observation, it is more likely that the observation is relevant. Note, however, that the converse is not true. Sometimes items may be isolated not because they are irrelevant and do not generate interest, but rather because they are in the vicinity of only very few observers. If there were more people in their vicinity, more pictures may have been taken of them. Hence, some consideration to the level of isolation of the *location* of pictures needs to be made in outlier determination. Intuitively, a scene should be considered an outlier not only because it is different but because others who are present at the scene are not taking pictures of it. Correspondingly in our context, a picture is treated as an outlier, if it is geographically collocated with a popular picture set, but is *visually* significantly different from the group.

3. THE DEMONSTRATION SCRIPT

In this demonstration, the audience will be presented with a table-top landscape comprised of different "damage scenes" in an imaginary city that has just witnessed a natural disaster, as well as some scenery that is less relevant to damage documentation and response efforts. The audience will be presented with smart phones and asked to imagine that they are volunteer first-responders in that virtual city, instructed to document and report urgent concerns of relevance to the rescue mission by taking pictures of them and sending those to the local rescue center. They will be allowed to take as many or as few pictures as they like. For example, they can choose to photograph one or many of the table-top scenes. They can take multiple pictures of the same scene, or not. They can also choose to "confuse" the system by advertantly taking pictures of no relevance to the mission. A base-station (at the demo location) will receive all pictures, run our algorithm and choose a small subset to represent the "most urgent" rescue needs. These will be displayed and compared to the subset of pictures chosen by other baselines, such as random selection and FIFO, to show that our service provides better coverage without outliers.

Acknowledgment

Research reported in this paper was sponsored by ONR grant N00014-10-1-0172, DTRA grant HDTRA-1-10-1-0120, NSF grants CNS 06-26825 and 1040380, and the Army Research Laboratory under Cooperative Agreement Number W911NF-09-2-0053. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation here on.

4. REFERENCES

[1] Mani Srivastava, Tarek Abdelzaher, and Boleslaw Szymanski. Human-centric sensing. *Philosophical Transactions of the Royal Society, A*, 370(1958):176–197, 2012.

[2] Md Yusuf Uddin, Hongyan Wang, Fatemeh Saremi, Guo-Jun Qi, Tarek Abdelzaher, and Thomas Huang. Photonet: a similarity-aware picture delivery service for situation awareness. In *Proc. of RTSS*, 2011.