International Conference on Cloud Computing and Big Data 2015

# Using Global Camera Networks to Create Multimedia Content

Everett Berry and Yung-Hsiang Lu
Electrical and Computer Engineering
Purdue University, USA
{epberry, yunglu}@purdue.edu

Wei-Tsung Su
Computer Science and Information Engineering
Aletheia University, Taiwan
au4451@au.edu.tw

***Abstract*—Multimedia content created by users (in contrast to content created by professional studios) has been growing rapidly in recent years. Most of the content uses materials from the creators' personal collections, generated by the users, or downloaded from the Internet. It is difficult to create content that requires materials from far-away locations, long-term observations, or many sources. This paper presents a system using globally deployed network cameras to retrieve images or videos for creating personalized multimedia contents. The system currently has over 60,000 cameras across the world. Through the system, users can select some of the cameras and record data at desired time, frame rate, and duration. A demonstration creates a personalized news report for the Macy's Thanksgiving Parade in New York without actually visiting New York.**

***Keywords—camera networks; heterogeneous cameras; multimedia retrieval; cloud computing***

## I. INTRODUCTION

Since the Internet became widely accessible to the general public in 1990s, user-created materials (in contrast to the materials created by professionals) have become a major source of content. Every day, millions of messages are posted in blogs, millions of images are shared through social networks, and thousands of video clips are uploaded to sharing sites. Some sites allow profit sharing and encourage users to upload materials that can attract more viewers. This has created a new "ecosystem" in multimedia content creation and retrieval that follows these steps:

1. Users upload content that can attract many viewers.
2. When viewers search some topics, the sites retrieve and show relevant and popular content.
3. These viewers see advertisements at the sites.
4. The sites earn money by selling products to the viewers.
5. The sites share profits with the content creators.
6. The profits encourage more people to upload high-quality content that may attract even more viewers.

Many factors affect the popularities of videos [1-2]. A study [3] finds that most videos on YouTube are generated by users but the most popular content is produced by professionals. One explanation is that most user-created content is usually restricted to the materials available to the users. Users generally have the following ways to obtain materials for their content:

1. Personal collections of images or videos. The collections may be obtained from personal trips or from friends.
2. Content generated specifically for the videos to be uploaded.
3. Materials downloaded from the Internet. This, however, may create problems because the materials may be protected by copyrights.

In many cases, the creators are limited by the available content. Consider an example where a person living in Europe wants to create a news cast of the New-Year celebration in Hong Kong. This person has several options, for example (1) Fly to Hong Kong and shoot the images and videos and (2) Search the images and videos uploaded to Internet by others. Neither solution is ideal. The first is too costly and the person can capture the images or videos from only one specific location. The second is limited to what is uploaded by others; moreover, it is likely to have long delay after January 1st.

This paper presents a new approach to create multimedia content by retrieving publicly available images and videos from global camera networks. Many organizations deploy cameras and connect them to the Internet for various purposes, for example

- To inform the general public. Departments of transportation in many countries deploy traffic cameras for travelers to see the current traffic conditions.
- To attract potential customers. Some vacation resorts and national parks deploy cameras so that people may see beaches, mountains, lakes, or other scenery before making reservations.
- To provide real-time views of remote locations, for example, the laboratories in the Antarctica [4].

By using global camera networks, we have built a system that can retrieve images and videos for creating real-time multimedia content. This system currently has more than 60,000 cameras worldwide. Using our system, a content creator can select a set of cameras and record the data at specified time, for a given duration, and at the preferred frame rate. Moreover, the multimedia retrieval is programmable. We built a demonstration using the traffic cameras in New York City for the Macy's Thanksgiving Parade on November 27, 2014. We

are able to show many different views of the parade simultaneously, without actually going to New York. We envision that images and videos retrieved from global camera networks will become a major source for creating multimedia contents.

## II. Related Work

With the rapid decline of costs, digital cameras have been become widely available. A 20MP digital camera costs only $70 and a high-definition (HD) network camera costs less than $100. It is common for a single person to have multiple cameras in the person's mobile phone, tablet, laptop, and desktop. Some mobile phones and tablets even have cameras on both sides. Meanwhile, many video editing tools are available, allowing users to create high-quality multimedia content [5]. Content creation is not limited to traditional devices. Laibowitz et al. [6] use wearable devices to create multimedia content. Cameras have been used by many researchers for studying various topics related to the environment. The Phenocams in the University of New Hampshire [7] use cameras to monitor vegetation. The AMOS project [8] has retrieved millions of images from thousands of cameras.

These studies demonstrate the value of using cameras for conducting different types of environmental studies. We are not aware of any existing system that uses global camera networks to create personalized multimedia content. If our system is not unique, it is perhaps the largest.

## III. EventCam System

Our prototype system is named EventCam and it is used to record images and videos of events, programmable by users, occurring in different parts of the world. EventCam contains four major components: (1) camera database, (2) web interface, (3) content storage, and (4) content retrieval. The camera database contains information about more than 60,000 cameras from various sources, as well as the protocols needed to retrieve data from these heterogeneous cameras. The web interface allows users to select the cameras and to set the parameters for recording data from these cameras. The retrieved data may be downloaded and stored in the user's computer or in cloud instances. The details of these components are described as follows.

### *A. Camera Database*

The cameras in our system come from multiple sources, for example, departments of transportation, national parks, research laboratories related to environmental protection, and universities. These cameras are discovered by searching the Internet or signing data sharing agreements with camera owners. One challenge of handling the wide range of cameras is that different methods are needed to retrieve data from these cameras. Currently, about 45% cameras are in the USA and another 45% are in Western Europe. The remaining 10% are in Asia, Africa, and the Southern Hemisphere. We are working on finding more cameras across the world and adding them into our system. Fig. 1 shows the distribution of the cameras in the USA. The northeast area near New York has the highest density of cameras.

Creating this camera database requires significant effort. First, we have to discover cameras by searching the Internet. Based on a market report, millions of network cameras are sold each year [9]. Thus, we expect much more cameras can be added to our system. Second, we have to find methods to retrieve the data. Some cameras have IP addresses and are connected to the Internet directly. An HTTP GET request can retrieve the data from these cameras. However, different brands of cameras have different paths for the GET request so the database must include the brands of the cameras. Third, because other cameras are connected to web servers and we have no direct access to the cameras, we must retrieve data from the web servers. In some cases, the web pages are dynamically generated and the images' URLs include the timestamps when the images are acquired. Retrieving data from these web sites requires additional effort.

Another challenge is finding the cameras' locations. Some sources (such as departments of transportation) provide accurate locations in longitudes and latitudes. The other sources do not provide locations and we have to use several methods (such as IP address to location mapping) to locate the cameras. The mappings from IP addresses to geographical locations are not always accurate. Moreover, given the large number of cameras, it is inevitable that some cameras may be disconnected for various reasons. Our system takes a snapshot from each camera weekly to ensure that the camera is still accessible.

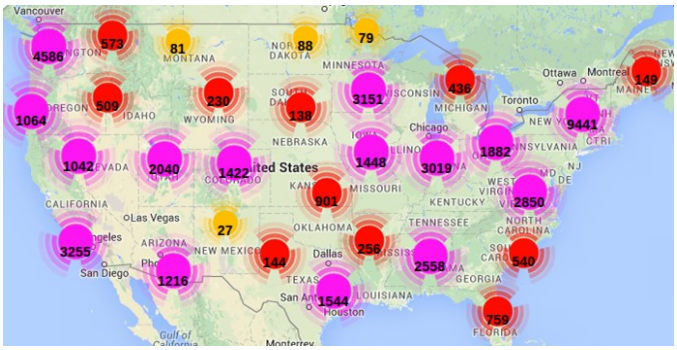


Fig. 1 The distribution of the cameras in USA.

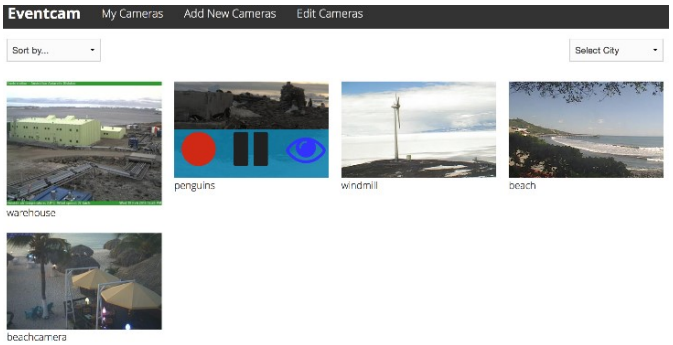


Fig. 2 The EventCam user interface. A user can record, pause, or view a live feed from the cameras.

### B. *Web User Interface*

The system is controlled by a web interface as shown in Fig. 2. A user has to register and login to the system. Then, the user can select cameras based on their geographical locations such as country, state (or province), and city. After selecting a location, the user may choose to see recent snapshots from all cameras in this area. The snapshots allow the user to see the content seen by each camera. The user may select cameras by clicking the snapshots. After selecting the cameras, the user needs to create a configuration which includes specific times to start recording, the duration to record, and the minimum desired frame rate.

### C. *Data Retrieval and Recording*

After setting these parameters, the user can proceed to record the data from the selected cameras. EventCam can store the data on the user's local disk or in the cloud. If only a few cameras are selected and the frame rate is low, storing the data at the user's computer is convenient. Using cloud instances can retrieve data from many cameras simultaneously at a high frame rate. If cloud instances are used, the data can be transferred to the user's computer after the scheduled recording has completed. EventCam automatically names each file by adding the timestamps when the data are acquired. The timestamps help the user manage the files.

It is a research problem of determining the number of cloud instances to meet the requirements of retrieving and storing the data from all the selected cameras. Many factors affect the data rates (Mbps). First, some cameras provide video streams (such as MP4) and some other cameras provide MJPEG (motion JPEG). MP4 has inter-frame compression and requires fewer bits per pixel. Both MP4 and MJPEG allow the cameras to send multiple frames for one retrieval command. The other cameras support only single-frame JPEG files; each retrieval command can obtain only one frame. The cameras' resolutions and frame rates play important roles in the data rates. Second, the locations of the cameras and the cloud instances can affect the network round trip time (RTT) and indirectly affect the frame rates. Our measurements using Amazon Web Services and Microsoft Azure reveal that the frame rates may drop noticeably when one 8-core cloud instance retrieves data from more than 100 cameras simultaneously. Efficiently allocating cloud resources is currently being investigated in our research group.

### D. *Programmable Content Retrieval*

EventCam allow users write programs to retrieve multimedia data in an event-driven basis. For example, users maybe only interest in the videos on the crowded streets. Then, users could write a program to detect the crowded streets using OpenCV-Python, one the famous image processing library. A number of cloud instances will be lunched to execute the programs for processing multimedia content from all the selected cameras simutanenously.

## IV. Demonstration: A News Report

To demonstrate EventCam, we created a news report for the Macy's Thanksgiving Parade in New York on November, 27, 2014. New York City is chosen for three reasons: First, the route of the parade is known in advance. Second, there is a high density of cameras at the intersections of the streets. We can select the cameras along the route. Fig. 3 shows the route of the parade and the locations of the traffic cameras. Third, the parade is a scheduled event and we can set the parameters for recording the data.

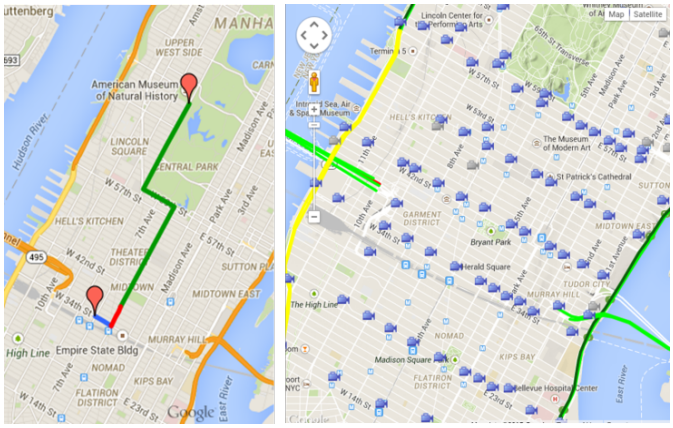


Fig. 3 (left) The route of the Macy's Parade. (right) The locations of the traffic cameras in New York City.

Before the parade, we used EventCam to browse all traffic cameras along the route and selected the cameras that could have clear views. Then, we set the recording parameters and the system recorded the images from these cameras during the parade. Using EventCam, we could record the parade from many different locations simultaneously. After recording the images, we created a video news report for the parade, similar to a typical TV news report. This report is uploaded to YouTube. A snapshot of the video is shown in Fig. 4. (For blind review, the URL of the video is not revealed here.) Using traffic cameras has one unique advantage: the cameras are usually mounted at traffic signal poles, much higher than the viewing crowd. Hence, the parade can be seen clearly without being blocked by anyone. We are able to create this news report by using EventCam even though no one in our group went to New York.

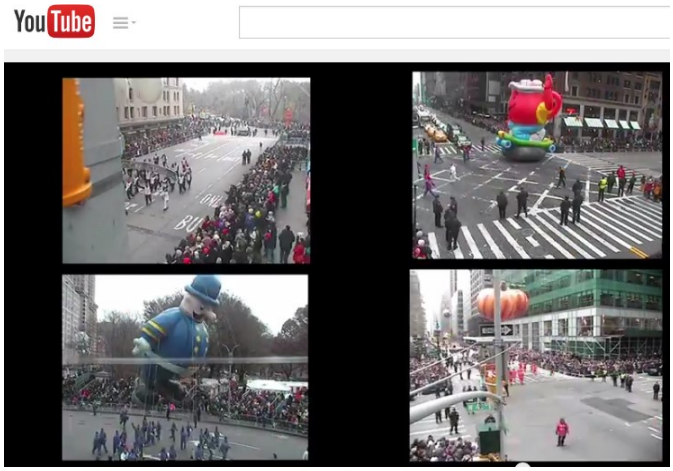


Fig. 4 Using EventCam, we can show four (or more) views simultaneously.

Despite the initial success of EventCam, the system requires many improvements. For future work, we will increase the coverage by adding more cameras into the system. The user interface will allow more options than using locations for finding cameras and recording the data. Also, linking EventCam with an Internet search engine will help users find scheduled events as well as recent news so that the users can more easily find cameras covering the areas they are interested in.

## V. Additional Use Cases of EventCam

Besides making personal videos, EventCam can be used for many other purposes. Here are some possible applications.

### A. Breaking News

EventCam can be used to report breaking news in various scenarios, for example,

- when the camera crew cannot arrive immediately
- when transportation reaching the scene is difficult
- when the scale of the event covers a very wide range and the camera crew cannot provide full coverage (such as flooding)
- most important, when it is too dangerous to send a camera crew. The Huffington Post reports (November 25, 2014) that a news anchor was almost hit by tear gas when reporting a protest. Wildfire, volcano eruptions, and military conflicts are other examples when the reporters' own lives can be in danger.

Journalists sometimes take great risks for obtaining live video footages. The Committee to Protect Journalists reports that 1,100 journalists have been killed since 1992. Adopting technologies like EventCam may reduce the danger of journalists. Moreover, if EventCam can provide live reports immediately after an event occurs and before any news crew arrives, emergency response officials may be able to obtain early assessments of the scene.

### B. Environmental Studies

Another application of EventCam is long-term environmental studies. Many cameras in our system watch natural scenery. If the data are periodically recorded for long durations, they can reveal insightful knowledge about our environment. EventCam provides a simple user interface for environmental researchers to record the data from many cameras.

### C. Transportation and Urban Planning

Many cameras in the system watch roads and urban infrastructures. This system may be used to understand how the infrastructures are used, for example, how many vehicles pass a road, when congestion starts, etc. Moreover, the navigation based on real-time videos from cameras selected in the route can be provided to users.

## VI. Conclusion

This paper presents a system called EventCam for recording data from many cameras around the world. The recorded videos and images can be used for creating timely and personalized multimedia contents. Moreover, users could write programs to analyze the data from many cameras simultaneously. Currently, we are working on (1) providing more public cameras and even allowing users attach their own cameras to the proposed system, (2) providing more options selecting cameras which let users easily select the cameras they need, and (3) investigating the efficient cloud resource allocation to meet the requirements of processing data from multiple sources.

## References


[1] Flavio Figueiredo et al., 2011. The tube over time: characterizing popularity growth of youtube videos. In Proceedings of the ACM international conference on Web search and data mining.

[2] Adiya Abisheva et al., 2014. Who watches (and shares) what on youtube? and when?: using twitter to understand youtube viewership. In Proceedings of the ACM international conference on Web search and data mining

[3] Gijs Kruitbosch and Frank Nack. 2008. Broadcast yourself on YouTube: really?. In Proceedings of the ACM international workshop on Human-centered computing.

[4] Department of the Environment, Australian Antarctic Division, http://www.antarctica.gov.au/webcams.

[5] Susmit Bagchi, 2008. A framework for designing personalized ubiquitous multimedia. In Proceedings of the international conference on Entertainment and media in the ubiquitous era.

[6] Mathew Laibowitz et al., 2010. Multimedia content creation using societal-scale ubiquitous camera networks and human-centric wearable sensing. In Proceedings of the international conference on Multimedia.

[7] University of New Hampshire Phenocam. http://phenocam.sr.unh.edu/webcam/.

[8] Nathan Jacobs et al., 2007. Consistent Temporal Variations in Many Outdoor Scenes. In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition.

[9] Marketandmarkets.com. Network camera and video analytics market. September 2012. Report Code: SE 1238.