Visualization Framework for Inter-Media Comparison using Image Flows

Masahiko Itoh* The University of Tokyo
Masashi Toyoda† The University of Tokyo
Cai-Zhi Zhu‡ National Institute of Informatics
Shin'ichi Satoh§ National Institute of Informatics
Masaru Kitsuregawa¶ National Institute of Informatics, The University of Tokyo

Abstract

To understand recent societal behavior, it is important to compare how multiple media react to real world events and how each medium reacts to other media. This paper proposes a framework for inter-media comparison through visualizing images extracted from different types of media. We extract blog image clusters from our six-year blog archive and search for similar TV shots in each cluster from a broadcast news video archive by using image similarities. We then visualize such flows of images on a timeline to explore visually changes in activities and interests of people and differences and/or similarities between media such as image clusters that become hot topics on only blogs or that become popular on blogs earlier than on TV.

1 Introduction

The first photo of the plane ditching during the "Miracle on the Hudson" on January 15, 2009 appeared and spread on Twitter and was then used in TV news. In the case of the "Chelyabinsk Meteor" on February 15, 2013, many people reported videos of the incident on YouTube, and then mass media reused them on TV programs. Our use of media has changed dynamically in the last decade, and this affects our societal behavior. Mass and social media affect each other.

Visualization of image flows in multiple media resources such as blogs and TV news helps us to understand the difference in exposure time of topics between media by checking frequencies of the appearance of topical images in each medium, the influences of media on each other by examining the difference in timing of bursting, or which medium first provided the information by tracking the origins of these images. It is therefore useful for marketing, politics, and sociology to extract and visualize such image flows from various kinds of media resources.

Eccles et al. [1] visualized transitions of time sequential events using a 3D space. There has been much research visualizing and analyzing temporal changes in trends on set of images [3, 2]. However, no study has visualized and explored transitions of trends of images related to multiple topics among different kinds of media resources in a 3D space.

This paper proposes a framework for inter-media analysis through image flows extracted from blogs and TV to understand societal behaviors. Both blogs and TV generate huge amount of flows of images every day. Comparing such enormous image flows from blogs and TV is one of our big challenges to capture activities over both media. Images extracted from blogs and TV are visualized in a 3D space. For each topic and medium, images are piled up like a time series histogram, and are arranged so that the user can easily compare differences in exposure and timing.

*e-mail: imash@tkl.iis.u-tokyo.ac.jp
†e-mail: toyoda@tkl.iis.u-tokyo.ac.jp
‡e-mail: cai-zhizhu@nii.ac.jp
§e-mail: satoh@nii.ac.jp
¶e-mail: kitsure@tkl.iis.u-tokyo.ac.jp

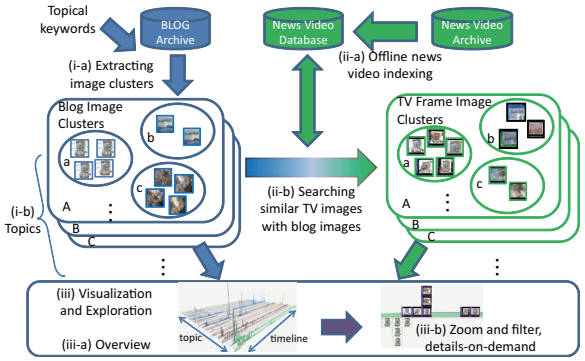

Figure 1: Overview of framework for visual inter-media comparison

2 Framework for Visual Media Comparison

Our framework aims for comparing appearances of topical images in social media and mass media and tracking the origins of these images. As the first step to this purpose, we utilize two media archives. One is a blog archive that includes two million blog feeds and one billion posts for seven years. The other is a broadcast news video archive that includes 12,498 news videos, lasting more than 6,000 hours, on six channels for 19 months.

We built a system that enables users to compare exposure of topical images in both media and to detect which medium preceded the other on the topic. Our system retrieves relevant blog articles about a given set of queries from the blog archive, from which images and surrounding text are extracted. The extracted blog images are first clustered into sets of near duplicate images on the basis of visual similarity (Fig. 1 (i-a)). We use Lowe's implementation of Scale Invariant Feature Transform (SIFT) features to calculate visual similarity. Each image cluster represents a fine-grained topic in the user's interest. To capture broader topics, these image clusters are clustered again on the basis of textual similarity calculated by the cosine of the term frequency vectors weighted by tf-idf (Fig. 1 (i-b)). We consider these sets of image clusters as topics.

For tracking the origins of blog images, we use the blog image clusters (sets of near duplicated images) as queries for retrieving corresponding shots in the broadcast news video archive. Unlike our previous work [4], this framework does not use textual information but image similarity to extract corresponding images. To retrieve shots from the large-scale news videos, we build a scalable shot retrieval index (Fig. 1 (ii-a)) that enables us to retrieve similar shots from a given image cluster (Fig. 1 (ii-b)). For each shot, multiple frames are sampled at a rate of 1 fps, and then for each frame, hessian affine Root SIFT features are extracted, and BoW quantization is applied to obtain a normalized BoW representation. We then build an inverted index for all normalized BoW vectors in the database. We therefore retrieve similar news video shots for each image cluster (e.g. a, b, and c in Fig. 1) for a given topic (a set of

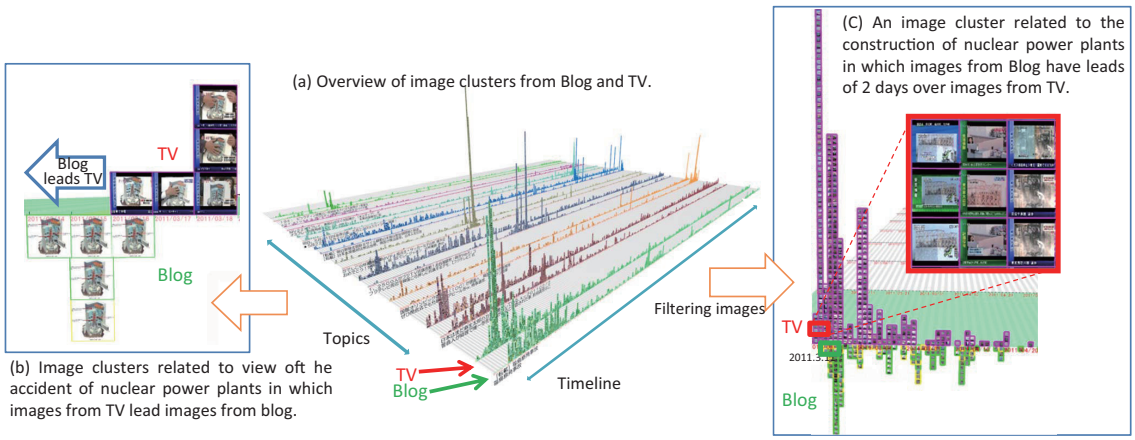

Figure 2: Visualization and exploration of image clusters from blogs and TV.

blog image clusters such as A in Fig. 1). Then retrieved shots are also gathered into a corresponding topic in news videos.

The corresponding topics in blogs and news videos are arranged in a 3D space (Fig. 1 (iii)) on the basis of their timestamps at which they are posted or broadcasted. Though our previous work [4] visualized image flows from TV to compare TV and blogs, it visualized images from only one kind of media resource at once and had no mechanisms for dynamically filtering images on the basis of temporal correlation to explore image flows interactively. Section 3 provides the detail of functions for visualization and exploration of image flows.

3 Visual Exploration Environment

Our system provides the following functions for requirements to analyze trends visually from flows of images on various topics.

(1) A function for visualizing time series of image flows is required. We adopt a histogram of images by stacking images on a timeline (Fig. 2). This design enables us to find out the beginning of the topic, bursting points, and a lifetime of the trends while confirming the contents of images.

(2) A function is necessary for comparing multiple image flows, which consists of multiple image clusters, on different topics. We hence arrange multiple histograms of images in a 3D space (Fig. 2 (a)). This design allows us to observe different situations between different topics, sequence of trends, and events at the same time in different topics.

(3) To compare similarities and differences between different media resources, we arrange image flows extracted from two different data resources: front and behind (Fig. 2 (a)) or top and bottom (Fig. 2 (b)(c)). This enables us to find interesting events that cannot be extracted using only one media resource, such as events that become popular on blogs earlier than on TV.

(4) To explore image clusters with interesting characteristics dynamically and easily, we provide a dynamic query function. It enables us to dynamically filter out unnecessary image clusters on the basis of attribute values of image clusters such as the number of images, similarity of image flows, and lead/lag of days among two image flows and cross-correlation of them (Fig. 2 (b)(c)). By using this function, we can find out which medium influences the other and what kind of events become popular in only blogs or TV or both blogs and TV.

(5) To confirm the contexts of the image flows, we provide functions for accessing original contents on both blogs and news video about selected images on the timeline.

4 Case Studies

Examples shown in Fig. 2 (b) and Fig. 2 (c) visualize images on a topic related to the accident at nuclear power plants in Fukushima after the Great East Japan Earthquake on 11 Mar. 2011.

Fig. 2 (b) visualizes an image cluster related to the construction of a nuclear power plant, in which images from blogs have leads of two days over images from TV. We find these images by using the dialog for dynamic query while setting appropriate values for cross-correlation and values for leads of days for blogs over TV. By accessing original contents, we can understand an explanation of the structure of and situation at the nuclear power plant in Fukushima by a researcher at MIT was spread through the Web, and then TV programs started to use the same image to explain the situation of the accident.

Fig. 2 (c) shows image clusters related to the outside views of the nuclear power plants. In this case, we select some representative images such as broken buildings of the power plants, the scene of an explosion at the power plants, and aerial photograph of the power plants and then filter out other image clusters. Fig. 2 (c) shows that most of these images first appeared on TV and then spread on the Web. By accessing original news video, we can see that TV uses the same photos (not videos) as blogs because information sources were very limited.

5 Conclusion

In this work, we have proposed an interactive visualization and exploration framework for inter-media comparison through the flows of images. By visualizing image flows in multiple media resources, we can recognize changes in trends of people's ideas, experiences, and interests through exploring the chronological flows of images. Flows of images enable us visually to grasp situations and timings of events in which people have participated.

One future work will be trying an opposite flow for searching images in which we generate query images from TV and then search for similar images from blogs. This framework can be applied to other media such as images on Twitter or videos on YouTube to compare influences between different kinds of media resources.

Acknowledgements

This work was supported by the Multimedia Web Analysis Framework towards Development of Social Analysis Software program of the Ministry of Education, Culture, Sports, Science and Technology, Japan.

References

[1] R. Eccles, T. Kapler, R. Harper, and W. Wright. Stories in GeoTime. In *Proc. VAST '07*, pages 19–26, 2007.

[2] G. Flake. Is Pivot a Turning Point for Web Exploration? In *http://www.ted.com/talks/gary_flake_is_pivot_a_turning_point_for_web_exploration.html (TED.com video - Feb/Mar 2010)*.

[3] A. Gomi and T. Itoh. A Personal Photograph Browser for Life Log Analysis based on Location, Time, and Person. In *Proc. of SAC '11*, pages 1245–1251, 2011.

[4] M. Itoh, M. Toyoda, and M. Kitsuregawa. Visualizing Time-Varying Topics via Images and Texts for Inter-Media Analysis. In *Proc. IV '13*.