

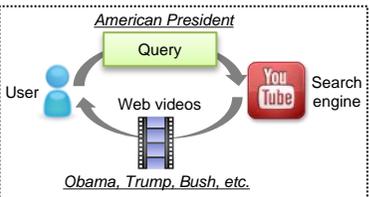
(UCD-P1.5) TRACKING HIERARCHICAL STRUCTURE OF WEB VIDEO GROUPS BASED ON SALIENT KEYWORD MATCHING INCLUDING SEMANTIC BROADNESS ESTIMATION

Ryosuke Harakawa, Takahiro Ogawa, and Miki Haseyama
Graduate School of Information Science and Technology, Hokkaido University

© 2016 Haseyama Laboratory

1. Introduction

Background of Web video retrieval



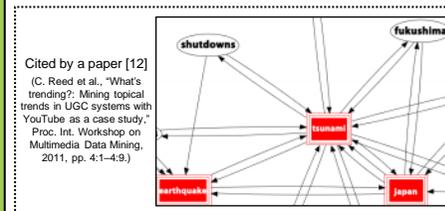
Time stamps including desired topics are different depending on each user.

Ex. Given a topic "American president", some users desire Web videos about "Donald Trump" and some desire "Barack Obama".

Problem Existing search engines cannot provide retrieval results considering trends of topics over time.

2. Related Work

Conventional studies to track Web video groups over time for Web video retrieval [8-12]



Cited by a paper [12] (C. Reed et al., "What's trending?: Mining topical trends in UGC systems with YouTube as a case study," Proc. Int. Workshop on Multimedia Data Mining, 2011, pp. 4-1-4-9.)

Web video sets with similar topics

Realize retrieval by providing Web video groups obtained via flat clustering considering time information

Merit

Can retrieve desired Web videos by grasping trends of topics over time

Problem

Cannot accurately track Web video groups since their semantic broadness is ignored

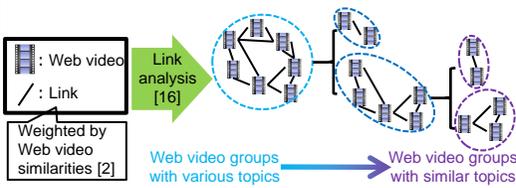
Ex. Cannot distinguish Web video groups with different semantic broadness (e.g. "iPhone" and "Electronic device")

3. Proposed Method

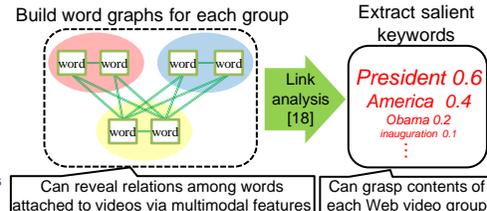
Novel method to track Web video groups considering semantic broadness for Web video retrieval

Phase I: Extraction of the hierarchical structure of Web video groups

(i) Extraction of the hierarchical structure

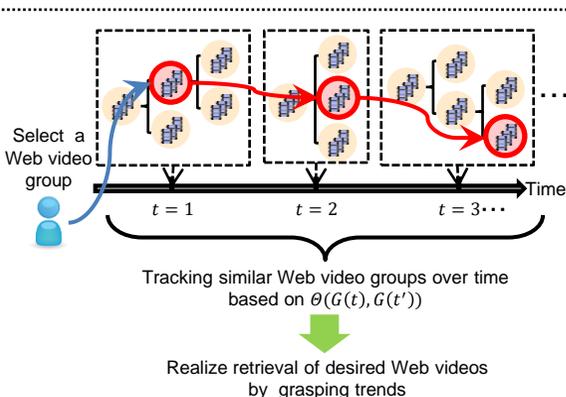


(ii) Estimation of salient keywords



Phase II: Tracking the hierarchical structure of Web video groups over time

Main contribution



Novelty Similarity definition between Web video groups at different time stamps $\theta(G(t), G(t')) = S_k(G(t), G(t')) \times S_b(G(t), G(t'))$

① Similarity based on salient keyword distribution

$$S_k(G(t), G(t')) = \frac{\sum_u \sum_{u'} a(u, G(t)) a(u', G(t')) \delta_{uu'}}{\sqrt{\sum_u a(u, G(t))^2} \sqrt{\sum_{u'} a(u', G(t'))^2}}$$

$a(u, G(t))$: attribution degree of a salient keyword u to a Web video group $G(t)$
 $\delta_{uu'}$: 1 if u and u' are the same and 0 otherwise

Can track Web video groups with similar topics via multimodal features

② Similarity based on semantic broadness

$$S_b(G(t), G(t')) = \exp\left(-\frac{\|H(G(t)) - H(G(t'))\|^2}{2\sigma^2}\right)$$

$H(G(t))$: entropy [19] of salient keywords of $G(t)$
 σ : predefined threshold

Can reduce over-tracking if many salient keywords are the same although semantic broadness is different from each other

4. Experimental Results

I. Purpose

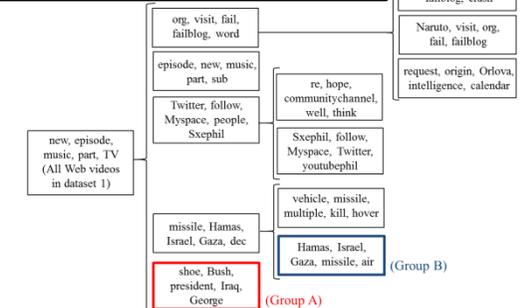
Verify the effectiveness of our new similarity $\theta(G(t), G(t'))$ that consists of salient keyword distribution (novelty ①) and semantic broadness (novelty ②)

II. Datasets

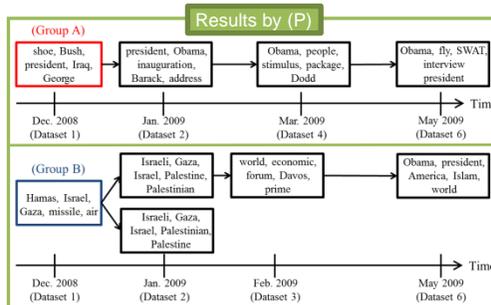
Public datasets MCG-WEBV 1.0 [20] containing "Most Viewed" YouTube videos of each month

	Crawling time stamp	Num. of Web videos
Dataset 1	December 2008	1315
Dataset 2	January 2009	1204
Dataset 3	February 2009	1141
Dataset 4	March 2009	1333
Dataset 5	April 2009	1439
Dataset 6	May 2009	1382

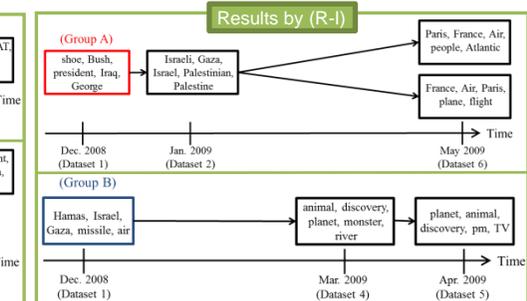
III. Hierarchical structure for dataset 1



IV. Tracking results for groups A and B



(P): Proposed method
 (R-I): Method that calculates $S_k(G(t), G(t'))$ by Doc2Vec [24] features
 (R-II): Method that ignores semantic broadness $S_b(G(t), G(t'))$



(P) ... Can accurately grasp changes of topics and people attention over time
 (R-I) ... Cause mistracking since importance of each word cannot be considered
 (R-II) ... Cause more than 100 unsuitable results including some common salient keywords in any hierarchies