



HAL
open science

News Archive Exploration Combining Face Detection and Tracking with Network Visual Analytics

Benjamin Renoust, Thanh Duc Ngo, Duy-Dinh Le, Shin'Ichi Satoh

► **To cite this version:**

Benjamin Renoust, Thanh Duc Ngo, Duy-Dinh Le, Shin'Ichi Satoh. News Archive Exploration Combining Face Detection and Tracking with Network Visual Analytics. ACM MultiMedia, 2016, Amsterdam, Netherlands. pp.728 - 730, 10.1145/2964284.2973823 . hal-01454918

HAL Id: hal-01454918

<https://hal.science/hal-01454918v1>

Submitted on 3 Feb 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

News Archive Exploration Combining Face Detection and Tracking with Network Visual Analytics

Benjamin Renoust
National Institute of
Informatics
& JFLI CNRS UMI 3527
renoust@nii.ac.jp

Thanh Duc Ngo
University of Information
Technology
VNU-HCM
thanhd@uit.edu.vn

Duy-Dinh Le and
Shin'Ichi Satoh
National Institute of
Informatics
{leddy, satoh}@nii.ac.jp

ABSTRACT

Visual analytics helps analytical reasoning and exploration of complex systems, for which it combines the means of interactive visualization, with the power of data analytics. The recent progress in computer vision techniques opens wide applications in real world video archives. Particularly, recent advances in face detection and recognition have been put under the spotlight. The applications of such techniques are often concern intelligence, or peer recognition in photo posted in social networks. We propose to combine those two domains by demonstrating a visual exploration of over a decade of the news program from the Japanese broadcaster NHK News 7. We derive social networks from face detection and tracking of this large dataset. With the help of a little domain knowledge, we monitor the activity of political public figures and explore the archive. This allows understanding and comparison of the politico-media scene presented by NHK under different Prime Minister's governance. The social networks are interactive, and to explore the multimedia database and its video content.

Keywords

Face detection and tracking; visual analytics; media analytics; social networks analysis; politics; NHK; Japan

1. INTRODUCTION

Media monitoring is a discipline which holds a strategic importance for media professionals, politicians and other public figures. This monitoring can help people manage their public image, broaden the scope of a news story, survey the impact of decision, or even predict market behavior. It requires then to mine the ever producing sources of information that constitutes the media. A traditional way is to look at the semantics behind the stories and Visual analytics often tackled news media analysis by combining topic analysis, named entity recognition, and time series analysis to show trends, reveal important individuals, summarize

stories, or do predictions. However, news media are often conveyed through video and the use of the images became as strategic as the language that tells a story.

Because computer vision tools are becoming mature enough, we propose to go beyond these traditional approaches and to work solely on the visual content conveyed by the news. Through face detection and tracking, we want to monitor the activity of political figures. Since 2001, we have archived thousands of hours of news from NHK, and we propose a visual exploration of these news. We explore the news from a social networks perspective which gives an understanding the politico-media scene presented by the NHK over the different cabinets led by different Prime Ministers. Users can then explore the social networks over different time frames, compare the activity of the different political actors and finally access the video segments in which politicians have been detected.

2. DESCRIPTION OF THE SYSTEM

Our system first combines news segments with a large scale face detection and tracking framework as illustrated

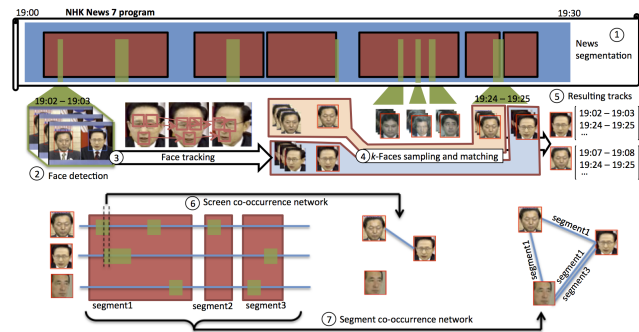


Figure 1: Network construction from face-tracking. 1- The video data is segmented given topical distribution, a topic roughly corresponding to a news segment. 2- Faces are detected in the video with Viola-Jones. 3- Faces are tracked using the KLT point tracker. 4- We use GreedyRSC to cluster the faces which are later manually annotated, resulting in face-tracks (5). This allows us to generate two forms of links between annotated people: we can connect people when they simultaneously co-appear on screen (6) and when they appear during the same news segment (7).

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

MM '16 October 15-19, 2016, Amsterdam, Netherlands

© 2016 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-3603-1/16/10.

DOI: <http://dx.doi.org/10.1145/2964284.2973823>

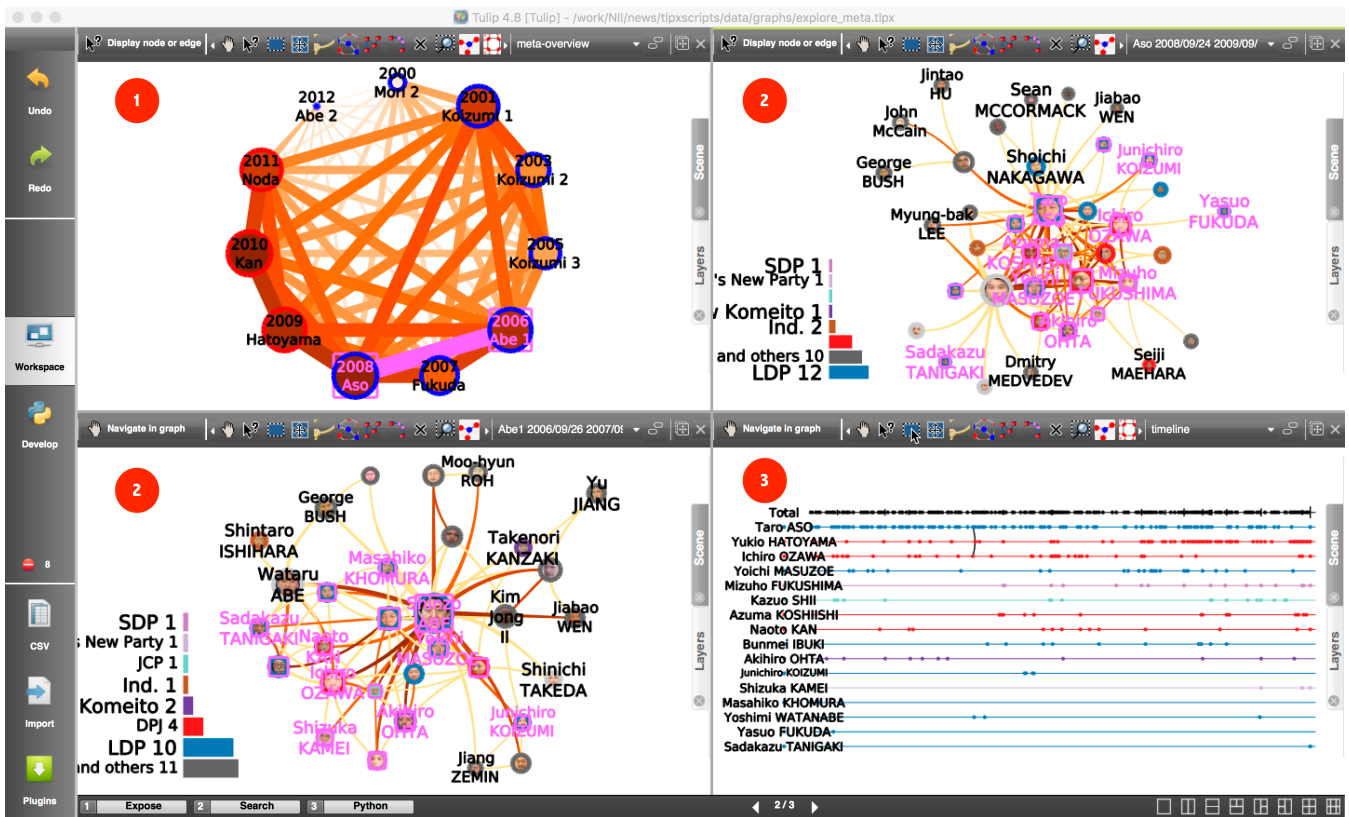


Figure 2: Network visualization. 1- The visualization offers first an overview of the cabinets ordered in time, with links showing similarity among politico-media scenes. 2- Selecting two cabinets in the previous view opens two network views corresponding to each selected cabinet. The common political figures are highlighted in pink. 3- The timelines of appearance of the selected political figure during Aso’s cabinet. Political leaning colors the line, and crossing connections (in black) correspond to screen co-occurrence.

in Fig. 1, so we can identify in which stories are the different political actors participating in, in order to construct relevant social networks.

A news program traditionally presents different topics in a succession of news stories that we call news segments. We first enrich our videos with meta-data and automatically extract news segment boundaries (Fig. 1-1), kindly provided by the authors of [2].

We then proceed with a large scale face detection and tracking. We first detect faces in each frame using Viola-Jones[6] (Fig. 1-2). We create face-tracks by regrouping faces instances matching through the KLT point tracker [5] (Fig. 1-3). Each face-track is temporally sampled using k -faces [4], and we extract their average point in the OpenFace 128-dimensional feature space [1]. Face-tracks are finally clustered using GreedyRSC [3] (Fig. 1-4), and manually annotated. We finally obtained about 150 different public figures (Fig. 1-5) over 15k facetracks in 3370 clusters.

From this process we can extract two types of relationships. Relationships appear when two individuals co-appear on screen (Fig. 1-6). Although strong, the meaning of this relationship is ambivalent: it often means that people appeared on the same place at the same time, however, news often place multiple people on screen to express an opposition of ideas. This links are sparse, and because our videos often focus on one individual rather than on a group.

The second type of relationship is derived from the news segment boundaries (Fig. 1-7). Indeed, different politicians appear during a same news segment when they took part in a same news event. We thus create a link between two people each time they appeared during a same news segment.

From the 150 public figures annotated, we recognized 75 people involved in politics, including the different Prime Ministers leading Japan at the moment of capture. We can segment time, creating a time frame for each of these Prime Ministers’ cabinets, each corresponding to a sub-network. We additionally gathered party affiliation for each of the cabinets to enrich the visualization.

3. A MULTI-LEVEL VISUALIZATION

We first¹ offer a meta-view where each node corresponds to a cabinet’s time-frame (Fig. 2-1). Links between nodes correspond to the (Jaccard) proximity between cabinets in terms of political actors we have detected, mapped to size and color. Similarly, the size and color of nodes correspond to the number of politicians detected during the cabinet.

With interaction, the user can explore these cabinets, for which the social network is displayed (Fig. 2-2), a node corresponding to a politician. The user can select political parties and politicians, and linked highlighting (Fig. 2-2

¹Work implemented using the Tulip framework tulip.labri.fr

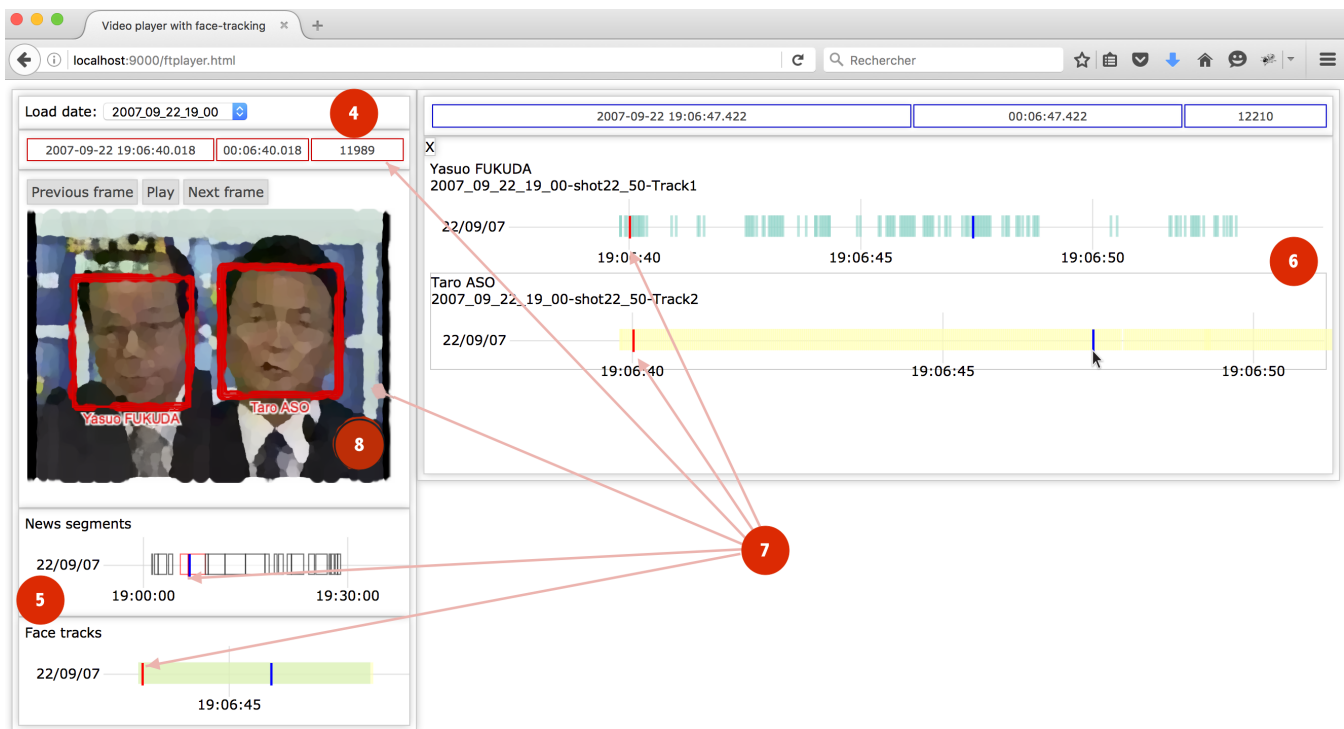


Figure 3: The face-track video player. Selecting sections in the timeline (Fig. 2-3) loads a playlist corresponding to the dates of detection (4). For a given date, we display timelines corresponding to topic segmentations (5) and individual face-tracks (6), all interactively linked highlighted independently to scales (7). Upon interaction, the video is played and face bounding boxes are displayed accordingly (8).

in pink) displays the selection in other views, to compare cabinets between one another.

Interaction also allows the user to compare the activity of the different politicians using timelines (Fig. 2-3). Each bin of the timeline corresponds to a day of detection of a politician, its height being mapped to the number of news segments appearances. Eventually, we also displays co-appearance links when available.

Finally, another interaction in the timeline allows for exploration of individual news programs. It opens a player² and loads a playlist corresponding to the selection (Fig. 3-4). The player displays multiple timelines corresponding to news segments (Fig. 3-5) and face-tracks (Fig. 3-6). The different timelines are all linked highlighted over hovering and video playing (Fig. 3-7) for a precise face-track exploration. Face-track boundaries are then displayed over the video (Fig. 3-8), allowing users to identify holes and duplicates in the face detection.

4. CONCLUSION

This system allows for a didactic discovery of the politico-media scene presented by NHK during our period of capture. Through exploration, we can overview different phenomena, such as: the changes of network topology between when Japan’s leading party changes; the rise of (future Prime Minister) Shinzo Abe during Jun’ichiro Koizumi’s cabinets; and the identification the very peculiar Ichiro Ozawa.

²Implemented over Canvas and D3 *d3js.org*

Acknowledgements

This material is based on work supported by the Japanese Society for the Promotion of Science (JSPS) under grant P14773. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect those of the JSPS.

5. REFERENCES

- [1] B. Amos *et al.* OpenFace: Face Recognition with Deep Neural Networks. <http://github.com/cmusatyalab/openface>. Accessed: 2016-01-11.
- [2] I. Ide *et al.* Topic threading for structuring a large-scale news video archive. *Image and Video Retrieval*, 1(1):123–131, 2004.
- [3] D. D. Le and S. Satoh. Indexing faces in broadcast news video archives. In *2011 IEEE 11th ICDM Workshops*, pages 519–526, Dec 2011.
- [4] T. D. Ngo *et al.* Face retrieval in large-scale news video datasets. *IEICE TRANSACTIONS on Information and Systems*, 96(8):1811–1825, 2013.
- [5] J. Shi and C. Tomasi. Good features to track. In *Proceedings CVPR’94.*, pages 593–600. IEEE, 1994.
- [6] P. Viola and M. J. Jones. Robust real-time face detection. *International journal of computer vision*, 57(2):137–154, 2004.