

Image Phylogeny Forests Datasets

1 Overview

In this material, we present more details regarding the datasets used in Image Phylogeny Forests approaches. Two types of datasets are described: a *controlled dataset*, in which we know a priori the ground-truth of the image relationships, and a *real cases dataset*, with semantically similar images collected from the Internet, and from which we often do not have the knowledge about their relationship. The following papers use some or all the datasets presented in this document:

- F. de O. Costa, M. A. Oikawa, Z. Dias, S. Goldenstein, and A. Rocha, “Image Phylogeny Forests Reconstruction”, *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 10, pp. 1533-1546, 2014 (DOI: 10.1109/TIFS.2014.2340017).
- M. A. Oikawa, Z. Dias, A. Rocha, and S. Goldenstein, “Manifold Learning and Spectral Clustering for Image Phylogeny Forests”, *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 1, pp. 5–18, 2016 (DOI: 10.1109/TIFS.2015.2442527).

All datasets are available at <http://www.recod.ic.unicamp.br/~oikawa/datasets.html>.

2 Controlled Datasets

When dealing with image phylogeny forests, instead of just near-duplicate images, the images under analysis are all semantically similar, i.e., they have the same content but may differ in the vantage point, zoom, or other camera parameters. To show that our approach is not biased on how the images were acquired, the dataset comprises two scenarios: one with all images coming from the same camera, but with changes in the camera parameters and slight changes in the camera pose (*single camera*), and another in which all images come from different cameras (*multiple cameras*).

2.1 Training Dataset

It represents a small exploratory set with $2 \times 3^3 \times 4 \times 1 \times 10 = 2,160$ forests, containing images from single and multiple camera scenarios, taken from three different cameras, three different scenes, three images per camera, four forest sizes $|F| = \{2..5\}$, one topology, and ten random variation of parameters for creating the near-duplicate images.

2.2 Dataset A

It comprises images from single and multiple camera scenarios, three different scenes, three different cameras, three images per camera, four forest sizes $|F| = \{2..5\}$, four different tree topologies, and ten random variations of parameters for creating the near duplicate images. Therefore, the dataset comprises $2 \times 3^3 \times 4 \times 4 \times 10 = 8,640$ forests.

2.3 Dataset B

It comprises images randomly selected from a set of 20 different scenes, 10 different cameras, 10 images per camera, 10 different tree topologies, 10 random variations of parameters for creating the near-duplicate images, and forests with 10 trees each. For each of the cases, single and multiple cameras, a total of 2,000 forests within this set were randomly selected with forests of size $|F| = \{2..10\}$. Therefore, this set comprises

$2 \times 2,000 \times 9 = 36,000$ test cases. Furthermore, in this data set, the family of image transformations \mathcal{T} considered are: re-sampling, cropping, affine warping, brightness and contrast adjustments, and lossy compression using the standard lossy JPEG algorithm.

In particular, this dataset comprises 10 indoor scenes and 10 outdoor scenes, taken with the following cameras:

- Canon EOS50D
- Canon PowerShot G12
- Canon PowerShot SX1-IS
- Fujifilm Finepix S4500
- Kodak EasyShare z981
- Nikon D5000
- Olympus SP-800UZ
- Panasonic Lumix
- Sony Alpha
- Sony Cybershot DSC-HX1

In the single camera scenario, images are grouped per camera for each scenario (indoor / outdoor). In the multiple cameras scenario, the images for each scenario are mixed with images coming from each of the 10 cameras. One sample image from each of the groups of indoor and outdoor scenes is depicted on Figure 1 and Figure 2, respectively.



Figure 1: Indoor scenes.



Figure 2: Outdoor scenes.

2.4 Real cases dataset

The datasets comprising real cases scenarios are described below:

- (a) **10 different target image groups from the Internet:** This dataset was originally used to evaluate the reconstruction of IPTs [1]. It comprises images from ten different target image groups from the Internet that became viral at the time of their publishing, as described below:
- **TG₁** - Iranian missiles - This image is the four-missile version of the image released by Sepah News, the media arm of Iran’s Revolutionary Guard. It was later revealed that the image was modified, with the third rocket from the left being the same down to the exhaust trail as the second rocket, and the cloud of smoke beneath it was copy-pasted from the smoke on the right.
 - **TG₂** - Bush reading - The original version of the photo was taken in the Summer of 2002 while the former USA President Bush was visiting George Sanchez Charter School in Houston. In the modified images that became viral, he appears holding his book upside down.
 - **TG₃** - WTC tourist - It features a photograph of a tourist standing on the roof of WTC, who appeared in many manipulated pictures after 9/11 attacks.
 - **TG₄** - BP oil spill - Picture from BP’s website showing workers monitoring underwater images on video screens that were actually blank in the unaltered image.
 - **TG₅** - Israeli-Palestinian peace talks - Egyptian newspaper Al-Ahram doctored a photo of a group of world leaders including Barack Obama, Mahmoud Abbas, King Abdullah II, and Benjamin Netanyahu, moving the Egyptian President Hosni Mubarak to the front of the group.
 - **TG₆** - Criminal record - The image portrays a criminal record that supposedly belonged to the current Brazilian President Dilma Rousseff published by a popular Brazilian newspaper Folha de Sao Paulo in 2009. It was also proved to be a fake, being acknowledged by the newspaper two months after it was published.
 - **TG₇** - Palin and rifle - The image shows the former Alaska Governor Sarah Palin posing in an American flag bikini while holding a gun. Palin’s head was removed from another picture of her marching in the Chugiak July 4 Parade, and pasted onto the body of a bikini-wearing model.
 - **TG₈** - Beatles rubber - Cover of the The Beatles’ Rubber Soul album.
 - **TG₉** - Kerry and Fonda - Image showing John Kerry and Jane Fonda speaking together at an anti-Vietnam war protest during the 2004 presidential election campaign. The image turned out to be a forgery, composed by two different images, and further altered to appear as if it was from an old newspaper clipping.
 - **TG₁₀** - O.J. Simpson - Image of the former football star O.J. Simpson that was charged with the murder of his ex-wife Nicole Brown Simpson and her friend Ron Goldman. A darkened version of this image with the size of the prisoner ID number reduced became the cover of the Time magazine and also the source of several controversial images and discussions in the Internet.

The total number of images across the ten target groups is 187. As this is a real-world dataset, we do not know, with 100% certainty, the correct roots and parenting relationships in the sets. However, we can introduce some controlled descendants in the sets and evaluate if they always appear at the same positions preserving the structure of the trees. Therefore, the evaluation here is performed by creating a direct descendant for each image of the dataset using five variations of parameters (considering the same transformations used to construct the controlled dataset), and comparing the reconstructed phylogeny before and after the insertion of these artificial nodes. Therefore, the total number of images in this dataset is $187 \times 6 = 1122$ images. Figure 3 depicts examples of the target groups.



Figure 3: 10 different target image groups from the Internet.

- (b) **The situation room:** This dataset is introduced in the first work of IPFs [?], and comprises the image taken by the White House photographer Peter Souza on May 1st, 2011, and its variants collected from the Internet (this episode is also known as *The Situation Room*). For this experiment, 98 near-duplicate images were collected through Google Images and with a manual analysis, these images were divided in different patterns considering: cases of inserting the Italian soccer player Marion Balotelli (ID a*), text overlay (ID b*), watermarking (ID c*), face swapping (ID d*), insertion of a joystick (ID e*) and hats (ID g*), and changes in the image size without splicing (ID n*). Qualitative evaluation is performed by checking whether the groups were correctly separated in the reconstructed tree.

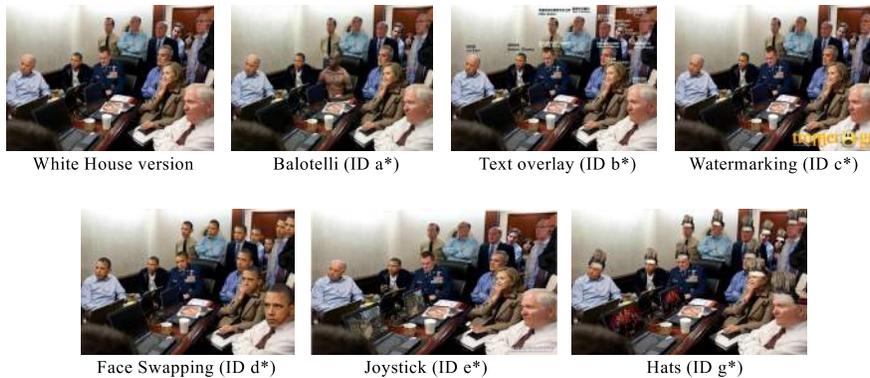


Figure 4: *The Situation Room* dataset.

- (c) **The Ellen DeGeneres' selfie taken at the Oscar 2014:** In this paper, we introduce a new dataset composed by images related to the *selfie* taken by the host Ellen DeGeneres and some famous actors during the 86th Academy Awards held on March 2nd, 2014. The original image became increasingly popular right after it was posted on her Twitter account. To this day, it has been retweeted more than three million times, and in addition to the retweets, several edited versions of this picture appeared on the Internet, with cases of text overlay, face swap, insertion of other people and animals in the picture, among others. This dataset is a clear example of a phylogeny forest, as the images are semantically similar but they were either taken with different cameras or in different points in time. We manually collected 44 pictures from this event posted on Twitter, blogs, and news websites, and divided it in five groups:

- **Group a*:** Edited versions of the picture posted at DeGeneres' Twitter account (@TheEllenShow).

- **Group b***: The moment the selfie was being taken but from the viewpoint of another camera.
- **Group c***: Similar to group b*, but with slight differences on the posture of the people in the picture. For instance, Angelina Jolie moves her arms and Brad Pitt straightens his back.
- **Group d***: Similar to group b* and c*, but the main differences are on their facial expressions.
- **Group e***: The moment before the selfie was taken, when the people on the picture starts gathering.



Figure 5: The *DeGeneres' Oscar 2014 Selfie* dataset.

References

- [1] Z. Dias, A. Rocha, and S. Goldenstein, "Image phylogeny by minimal spanning trees, *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 2, pp. 774–788, 2012.