

U-PHYLOGENY: UNDIRECTED PROVENANCE GRAPH CONSTRUCTION IN THE WILD

A. Bharati¹, D. Moreira¹, A. Pinto^{1,2}, J. Brogan¹, K. Bowyer¹, P. Flynn¹, W. Scheirer¹ and A. Rocha^{1,2}

¹Department of Computer Science and Engineering, Univ. of Notre Dame, IN, U.S.A.

²Institute of Computing, Univ. of Campinas, SP, Brazil

ABSTRACT

Deriving relationships between images and tracing back their history of modifications are at the core of Multimedia Phylogeny solutions, which aim to combat misinformation through doctored visual media. Nonetheless, most recent image phylogeny solutions cannot properly address cases of forged composite images with multiple donors, an area known as multiple parenting phylogeny (MPP). This paper presents a preliminary undirected graph construction solution for MPP, without any strict assumptions. The algorithm is underpinned by robust image representative keypoints and different geometric consistency checks among matching regions in both images to provide regions of interest for direct comparison. The paper introduces a novel technique to geometrically filter the most promising matches as well as to aid in the shared region localization task. The strength of the approach is corroborated by experiments with real-world cases, with and without image distractors (unrelated cases).

Index Terms— Image Phylogeny, Media Forensics, Undirected Phylogeny Graph.

1. INTRODUCTION

One key concern in digital forensics nowadays is how to fight propaganda and misinformation through visual media. With online visual content easily accessible, their reuse and iterative upload and download naturally lead to the presence of multiple copies of a single object. These copies can be generated through a series of transformations solely from an original image r (so-called near duplicates), from different originals r_i and r_j depicting nearly the same scene but each of which with its own chain of modifications (so-called semantically-similar images) or be combined with various other image donors d_i (generating composite images). Dias et al. [1] studied near-duplicate images and proposed a method to find their kinship relationships or the directions of modifications (and transformations) over time, terming such analysis as image phylogeny. Semantically similar images were studied in a follow-up work [2].

Extending those works, Oliveira et. al [3] formalized cases of image forgeries and compositions with what they called Multiple Parenting Phylogeny (MPP). In an MPP setup, an image can be derived from multiple donors and thus its content might have common pieces with all those donors. Moreover, each composite and donor image might have its own chain of near duplicates and semantically similar images. However, the best MPP solution that exists to date works only with bi-composite images (images that have two parents — a host and a donor image), which solves only a restricted

This material is based on research sponsored by DARPA and Air Force Research Laboratory (AFRL) under agreement number FA8750-16-2-0173. Hardware support was generously provided by the NVIDIA Corporation. We also thank the financial support of FAPESP (Grant #2015/19222-9), CAPES (DeepEyes Grant) and CNPq (Grant #304472 /2015-8).

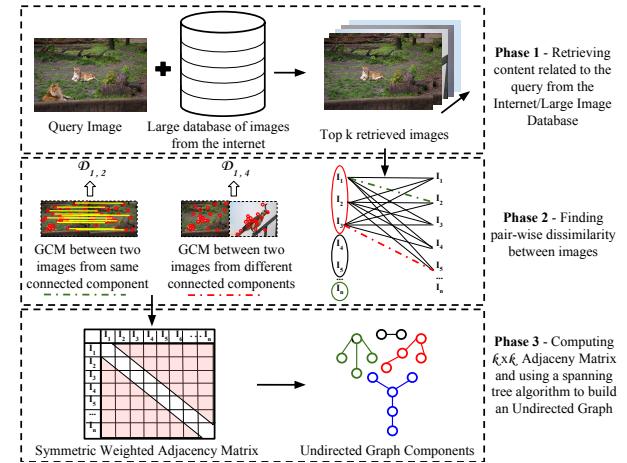


Fig. 1. Different stages of the proposed pipeline for provenance graph building of a given query and its possible donors. Upon searching an image collection and retrieving possible donors, we compare candidates pairwise through the matching of representative keypoints, geometrically check the consistency among possible matches and build a weighted dissimilarity matrix representing all possible pairwise relationships. Ultimately, we use a spanning tree to find the connected components related (and unrelated) to the query. **Note:** GCM stands for geometrically consistent matches.

case of MPP, leaving the more difficult general problem still largely untouched. With donors from multiple images, the information required to estimate the correct transformations to map an image onto each of its possible donors, might not be present, rendering existing MPP solutions inadequate for use in such cases.

Multimedia phylogeny solutions can be useful in detecting visual media frauds, preventing propaganda and misinformation dissemination, and resolving news media controversies. Therefore, it is important to devise methods that generalize to different image transformations as well as to any number of possible donors. Updating some terminology used in [3] and [4], for this work, rather than differentiating a donor as a possible host (donor of the background of a composite) or alien (additional donors), all images contributing to a composite image are simply referred to as donor images (DIs). The composite image is called a multi-composite image (MCI) (See Fig. 2).

Inferring directions to connections in the phylogeny graph of an unrelated set of images is difficult as the irrelevant images add noise to the process. Building an undirected graph first helps to reduce the uncertainty in direction finding and is more efficient (in terms of number of image comparisons) for large sets of images than directed phylogeny graph construction. Once the connections are obtained, localized techniques can be devised for pointing out the directions;

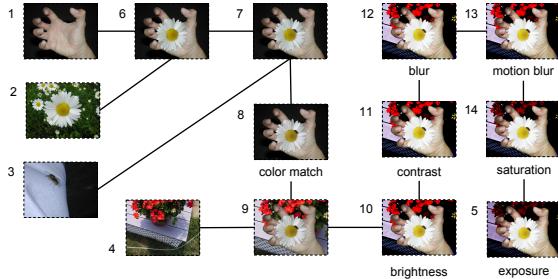


Fig. 2. Undirected phylogeny graph for the case of bi-composite (e.g., node 6) and multi-composite images (e.g., node 7). Examples from NIST NC2017 dataset [5].

human experts may also be an option.

In this paper, we propose an algorithm for Phylogeny Graph Construction that aims at building an undirected graph showing the relationships among images using spatial information provided by representative keypoints and the consistency of their matches. The method eliminates some assumptions made in prior work and generalizes to any set of images, hence "in the wild". The paper also presents results on a difficult dataset recently released by the National Institute of Standards and Technology (NIST) and proposes new metrics for evaluation of Image Phylogeny tasks. Instead of replacing all components of the existing work [4], we build upon them with generalization in mind and extend specific pieces of the method to deal with multiple donors and any kind of image transformation. Fig. 1 shows the end-to-end pipeline we propose. For this paper, we assume an efficient system for Phase 1 (retrieving images from a collection) and use its result (top- k related images to the query) as input for our algorithm.

2. RELATED WORK

Finding parental relationships between images was first explored by Kennedy and Chang [6] with Visual Migration Maps (VMMs) used to select images of interest from a given set of candidates. The main problem with VMMs, however, is the need of detectors for each image transformation, constraining the representation power of the graph to the detectors in place. In turn, De Rosa et. al [7] compared pairs of images using both the image content and the noise information to find possible dependencies.

The Multimedia Phylogeny term was introduced by Dias et al. [8] in their work with image phylogeny tree reconstruction with near-duplicate images. Subsequent work introduced solutions for multiple trees [1] and multiple trees with semantically-similar images [2]. None of the works, however, considered images with multiple parents, and so their solutions were in the form of trees.

Phylogeny (or provenance) graph construction for a more prevalent case of forgery in which objects from one image are spliced into another, was addressed by Oliveira et al. in their recent papers on multiple parenting phylogeny [3, 4], which are the most relevant to our work herein. The authors propose a solution with a strict assumption of two parents (one host and one alien) of a composite image, and also with no unrelated images in the set. Moreover, the authors assume a fixed set of possible transformations that allow an image to be considered as the near duplicate of another resampling, cropping and affine transformation, contrast, brightness, gamma correction, and compression (as defined in [1]).

Existing image phylogeny methods mainly focus on two steps to find the provenance graph: (i) computing the dissimilarity ma-

trix for all images in a collection; and (ii) building a directed graph using a spanning tree algorithm. Step (i) is further divided in (a) detecting matching keypoints for every pair of images; (b) estimating the best geometric transformation between those sets of points and warping one image onto the other; (c) matching color and compression parameters between the pairs of images; and (d) computing a pixel-wise difference between the mapped images.

While Step (ii) has been "the" subject of research in prior work, Step (i) has been overlooked, which has streamlined the research in the field by far. The many constraints with the existing solutions obfuscate the difficulties of the general MPP problem. Firstly, the set of transformations is not exhaustive and there can be transformations in real-world cases that have not been accounted for in the existing literature. Secondly, the methods used for estimating the transformations might have some limitations since they are based on local pixel information. Mapping the color distribution of content-related regions, for instance, can be similar in both directions (e.g., with reverse-prone transformations), thus not proving helpful or discriminatory for kinship direction finding. In addition, the information required to perform compression mappings, such as the compression table, might not be available for non-JPEG lossless-compressed images. Finally, there can be images completely unrelated (in terms of sharing one scene but with similar color distribution) to the query, as part of the result of the retrieval or the effect of the semantic gap [9].

3. PROPOSED SOLUTION

The proposed solution points out the undirected binary relations that might exist among the elements of a given set of images, based on their visual content. These relations aim at supporting the revelation of the phylogeny of the images. As explained in Sec. 2, prior work made strong assumptions regarding the probable phylogeny of the images, advancing the state of the art up to the particular case of donor-host composites. We extend the literature [4] toward the direction of analyzing more general composite cases, here called MCIs.

Fig. 1 outlines the main steps of the proposed solution. An end-to-end implementation starts with a query image q of interest, whose donors (if any) are to be discovered, together with possible near-duplicates of the query and the donors. The first step involves querying a large collection of images, for finding and sorting a list of the top- k potentially related items, according to their similarity to the query (c.f., *Phase 1* in Fig. 1, and Sec. 3.1 for details). Once the top- k related images to the query are retrieved, in *Phase 2*, we calculate the dissimilarity of each pair of images, including the query. For this step, we introduce novel strategies for computing dissimilarities and constitute *symmetric weighted adjacency matrices*, which are robust to varied image transformations, and rely upon image-pairwise keypoint detection, description, and geometrically consistent matching (GCM). This method is termed as *U-Phylogeny* (Undirected Phylogeny). We also propose an extended *U-Phylogeny* by computing dissimilarity values using pixel-wise local dissimilarity computations after the GCM, as this might be useful in some situations, but at the cost of an increased runtime (Sec. 3.2). This method can improve the results since it has more knowledge about forgery through transformations. Ultimately, in *Phase 3*, we estimate the query's provenance graph as a minimally connected undirected subgraph, which can be presented to an expert, or be fed to a further forensic image provenance oracle tool (Sec. 3.3).

3.1. Image Retrieval

Retrieving images related to the query image is the first step in the end-to-end pipeline of generating a phylogenetic graph. Different

Table 1. Experiments without distractors using *U*-Phylogeny (proposed algorithm) and its different forms of calculating the dissimilarity matrix. *MSE* denotes Mean Squared Error and *GCM*, Geometrically Consistent Matches.

Dissimilarity Metric	Performance Without Distractor Images								
	<i>Precision</i> _{edges}			<i>Recall</i> _{edges}			<i>VEO</i>		
Small	Medium	Large	Small	Medium	Large	Small	Medium	Large	
Avg. Distance of <i>GCM</i>	0.62 \pm 0.20	0.47 \pm 0.08	0.31 \pm 0.16	0.62 \pm 0.20	0.48 \pm 0.08	0.32 \pm 0.16	0.82 \pm 0.09	0.75 \pm 0.04	0.66 \pm 0.08
Number of <i>GCM</i>	0.75 \pm 0.19	0.61 \pm 0.13	0.52 \pm 0.15	0.75 \pm 0.19	0.61 \pm 0.12	0.54 \pm 0.15	0.88 \pm 0.09	0.81 \pm 0.06	0.77 \pm 0.07
<i>MSE</i>	0.73 \pm 0.19	0.56 \pm 0.10	0.42 \pm 0.04	0.73 \pm 0.19	0.56 \pm 0.10	0.43 \pm 0.03	0.87 \pm 0.09	0.79 \pm 0.05	0.72 \pm 0.02
Mutual Information	0.76 \pm 0.17	0.64 \pm 0.16	0.57 \pm 0.12	0.76 \pm 0.17	0.65 \pm 0.16	0.58 \pm 0.11	0.89 \pm 0.08	0.83 \pm 0.08	0.79 \pm 0.06

Table 2. Experiments with distractors. *MSE* stands for Mean Squared Error and *GCM* stands for Geometrically Consistent Matches.

Performance With Distractor Images					
Dissimilarity Metric	<i>Precision</i> _{nodes}	<i>Recall</i> _{nodes}	<i>Precision</i> _{edges}	<i>Recall</i> _{edges}	<i>VEO</i>
Avg. Distance of <i>GCM</i>	0.98 \pm 0.05	1.00 \pm 0.00	0.56 \pm 0.16	0.55 \pm 0.18	0.79 \pm 0.07
Number of <i>GCM</i>	0.98 \pm 0.05	1.00 \pm 0.00	0.72 \pm 0.15	0.69 \pm 0.16	0.85 \pm 0.07
<i>MSE</i>	1.00 \pm 0.00	1.00 \pm 0.00	0.69 \pm 0.14	0.64 \pm 0.11	0.84 \pm 0.06
Mutual Information	1.00 \pm 0.00	1.00 \pm 0.00	0.78 \pm 0.15	0.72 \pm 0.12	0.88 \pm 0.06

context-based information retrieval algorithms [10, 11, 12] can be adapted to tackle this part of the problem given large collections. In this paper, we do not focus on this particular task and assume to have been provided with a set of images after retrieval. The top-retrieved images may or may not be related to the query. Prior work on multiple parenting phylogeny [3, 4] did not consider unrelated images thoroughly.

3.2. Computation of the Dissimilarity Matrix

Given a set of k images, a dissimilarity matrix D is a $k \times k$ matrix with the value of dissimilarity between every pair of images from the set. The matrix can be considered as a weighted adjacency matrix of a graph, in which each image is a node and the values in the matrix correspond to weights of edges between any two nodes. Each edge weight is computed using the following steps:

Detection and Description of Points of Interest. Speeded Up Robust Features (SURF) [13] keypoints are detected on both images. The SURF keypoints highlight the important regions within the image content, and provide a description process and representation that are robust to transformations [14].

Keypoint Matching. To find correspondences between the detected keypoints in the two images, we compute the matches M between the two sets of descriptors D_1 and D_2 that are obtained from the previously detected keypoints. The first set is treated as the query set and the other is treated as the gallery set. For each descriptor $d \in D_1$, the best matching descriptor is found inside D_2 using L2 distance. In addition, inspired by the Nearest Neighbor Distance Ratio (NNDR) matching quality [15], we ignore all the keypoints whose ratio of distances to their first and second matched descriptors is smaller than an NNDR threshold t , implying that the keypoint might be of poor distinctive quality.

Keypoint Match Filtering. Once the matches are established, upon using NNDR, it is not uncommon to gather *geometrically inconsistent matches*, thus there is the need to remove spurious matches that are not truly representing the real transformations of one image onto the other (e.g., crossing matches among two images). A contribution of this paper is solving this problem with a filter of matched keypoints that keeps only the matches whose spatial dispositions are geometrically consistent in both images, say M_g . To obtain M_g , rather than relying on the value of the matched image pixels, we rely upon the spatial positions of the two best matched points $p_1(x_{p1}, y_{p1}) - p_2(x_{p2}, y_{p2})$ in I_1 , and $q_1(x_{q1}, y_{q1}) - q_2(x_{q2}, y_{q2})$ in I_2 . By taking the positions, the distance l_1 , and the angle a_1 between p_1 and q_1 (all from image I_1), as well as the positions, the distance l_2 , and

the angle a_2 between p_2 and q_2 (all from image I_2), we estimate the constraints with respect to scale, translation, and rotation transformations, from I_1 onto I_2 , thus applying them on all the keypoints of I_1 . We then compare the new positions of I_1 and the positions of I_2 , and remove the keypoints (and respective matchings) that do not follow the estimated constraints.

Dissimilarity. Finally, we compute the dissimilarity $D_{1,2}$ between the two images. For *U*-Phylogeny, we use the number of filtered (geometrically consistent) matches, M_g and the average match score of these matches as dissimilarity. Upon filtering the keypoint matches, for the extended (more expensive) algorithm, a few more steps are involved before computing the dissimilarity. The keypoints corresponding to the filtered matches are used to estimate homography between the two images. The images are registered based on the estimated parameters. Localized regions of interest (ROIs) are cropped from the registered images by computing the bounding box of the convex hull around the filtered keypoints.

The pixel value distribution of the two ROIs is matched using a frequency-based histogram-matching approach. The method involves computing the cumulative distribution for the pixel values for both source and target images (in each image pair). Each pixel of the source image is mapped onto the closest pixel value from the same quantile of the target histogram. Since our dissimilarity matrix is symmetric, the transformation only needs to be performed once (either of the two images can be a source image). Then, we compute the pixel-wise dissimilarity, in the form of Mutual Information and MSE, between the mapped source image and the target image.

3.3. Phylogeny Graph Construction

Upon obtaining the complete dissimilarity matrix, Kruskal’s Minimum Spanning Tree (MST) algorithm [16] is used to build an undirected graph connecting all images. The method requires two inputs – the number of retrieved images and the weighted adjacency matrix containing real-valued finite weight for each edge. The output graph is a binary adjacency matrix (*BAM*) for which BAM_{ij} is set to 1 whenever there is an edge (i.e., $edge(i, j) \in MST$).

4. EXPERIMENTAL SETUP

Datasets Used. For evaluation, we use NC2017-DEV2 dataset provided by the National Institute of Standards and Technology (NIST) as part of the Nimble 2017 Challenge [5]. The dataset is divided into query set (59 images) and gallery set (10446 images). Images from the gallery set may or may not be related to the query set. The

dataset has 59 phylogeny cases with 750 images in total. The average graph order (i.e., the average number of related images) for such cases is 12.7. The range of number of related images is [3, 82]. We organized such cases into three categories based on the number of nodes — small (≤ 12 nodes), medium (13-20 nodes), and large graphs (> 20 nodes).

To evaluate the robustness of the methods, we consider building the graphs under the presence of unrelated images. For this particular experiment, we sample 20 cases with graph order ≤ 25 . We process 25 images each time, regardless of the variable size of the provenance graphs. For instance, if we have a test case with 5 nodes, we complete this case with 20 randomly selected distractors and perform the analysis considering 25 nodes in total. The materials for reproducing this work are available at <https://gitlab.com/notredame-provenance/u-phylogeny>.

Evaluation Metrics. Existing metrics for evaluating image phylogeny generally focus on the notion of image phylogeny trees [1] and do not conform with undirected graphs as there is no notion of roots, leaves or ancestors therein. Hence, we rely on more general graph comparison metrics to evaluate results. We use precision and recall of nodes and edges and a combined metric, Vertex and Edge Overlap (VEO) as discussed in [17] in the context of web graphs. For each provenance case, the values for these metrics are obtained by comparing the output graph G' of a method with the ground truth graph G using the formulae in Eqs. 1 and 2, where P and R stand for precision and recall, respectively. Here, $nd_{G'}$ denotes the set of nodes (images) in graph G' while nd_G denotes the same for graph G . Precision and recall of edges is computed similarly.

$$P(nd) = \frac{|nd_{G'} \cap nd_G|}{|nd_{G'}|} \quad R(nd) = \frac{|nd_{G'} \cap nd_G|}{|nd_G|} \quad (1)$$

$$VEO(G', G) = 2 * \frac{|nd_{G'} \cap nd_G| + |edges_{G'} \cap edges_G|}{|nd_{G'}| + |nd_G| + |edges_{G'}| + |edges_G|} \quad (2)$$

The metrics take values in the range of 0 to 1. Higher values indicate better performance. The overall values reported for these metrics have been averaged over the three categories in the dataset.

Experimental Details. Upon receiving a list of images for phylogeny graph construction, we can have both related and unrelated images and we need to refine the list to create the graph. With this in mind, we divided the experiments into two setups:

1. **Without Distractor Images.** In this setup, all analyzed images for phylogeny graph construction are related to the query image. Differently from prior work [4], there might be multiple donors for each given query.
2. **With Distractor Images.** This scenario comprises possible failures in the retrieval of possibly related images and evaluates the performance of U -Phylogeny and its extensions in the presence of related and unrelated images.

For the dissimilarity matrix, we compute 2000 SURF keypoints for each image. The quality of these keypoints is governed by the hessian threshold (set to 100 to select the most important keypoints) and NNDR (individually computed for each image based on the top two-matched keypoints). Following [15], we use an NNDR threshold $t = 0.8$. The detected keypoints are filtered using the three parameters of rotation, scale and translation, individually computed for each image. For the U -phylogeny version, we use the match distance and count of these keypoints as dissimilarity.

For the extended version, the images are registered using the affine transformation matrix estimated by these keypoints. After

cropping the regions of interest (ROI) and mapping the pixel distribution of one to another, the mutual information and MSE is computed. These become the values of dissimilarity between the images.

5. RESULTS

Table 1 shows the result of Experiment 1. The $Precision_{nodes}$ and $Recall_{nodes}$ values are not valid for this setup as it has no distractors. Observe that the version of our approach using the keypoint information is on par with the extended version for small phylogeny cases and slightly below par for medium and large cases. It is important to note that the extended version of our algorithm matches the pixel color distribution as an additional step for each pair of images. The results show that the transformation mapping and pixel-wise comparison improve upon U -Phylogeny. In addition, mutual information of mapped pixel values outperforms the mean-squared-difference. This result is consistent with the literature [4]. The best performance for cases without any distractors is obtained with mutual information as the dissimilarity metric for the extended U -Phylogeny. Directly comparing our methods with [4] is not possible as [4] assumes a fixed set of transforms and at most two donors.

The ‘in-the-wild’ evaluation of the methods uses the same phylogeny cases with added unrelated images (distractors). Table 2 shows the results for this experiment. In this case, we also report the $Precision_{nodes}$ and $Recall_{nodes}$, which represent the methods’ performance for connecting images that are related and leaving out the unrelated ones. As can be seen from the values, the proposed approach works remarkably well in getting the connected nodes. U -Phylogeny also seems to be robust to distractors in terms of edges and the extended version achieves an 88% of vertex and edge overlap, a significant result for image phylogeny graph construction. As for efficiency, on an Intel(R) i7-5930K CPU 3.50GHz with 64GB of RAM, U -phylogeny takes 3.8s to compare two images whilst the extended version is twice as expensive.

6. DISCUSSION ON DIRECTION IN PHYLOGENY GRAPHS AND CONCLUSIONS

Until now, image phylogeny has been approached as tree-building given an asymmetric dissimilarity matrix. The asymmetry in the values while matching I_i to I_j and I_j to I_i is based on the transformations performed on the source image to match the target image. Assuming the transformations are not symmetric, the pixel-wise dissimilarity values of the two images are different. However, under real-world conditions, more complex (and not necessarily asymmetric) transformations might be present, especially when considering multiple donors to a composite. For such cases, the established concept of finding directions fails.

In this vein, this paper provides a generalized extension to the existing solution for multimedia phylogeny [1] and more specifically to the MPP problem [4]. We introduce methods to construct undirected phylogeny graphs for multi-composite cases with unknown number of donors. The method generalizes well over images in non-JPEG formats (another constraint present in prior work [4]) by not utilizing the loss information from image compression. Moreover, the paper also introduced the usage of new metrics of evaluation for phylogeny, which are adequate for undirected graphs. The proposed methods are reasonably robust to the presence of distractors landing themselves as promising preliminary solutions for phylogeny graph construction in the wild. Future work will be devoted to further refine the edge connections and infer the kinship directionality.

7. REFERENCES

- [1] Z. Dias, A. Rocha, and S. Goldenstein, “Image phylogeny by minimal spanning trees,” *IEEE Trans. on Information Forensics and Security*, vol. 7, no. 2, pp. 774–788, April 2012.
- [2] Z. Dias, S. Goldenstein, and A. Rocha, “Toward image phylogeny forests: Automatically recovering semantically similar image relationships,” *Elsevier Forensic Science Intl.*, vol. 231, no. 1, pp. 178–189, 2013.
- [3] A. Oliveira, P. Ferrara, A. De Rosa, A. Piva, M. Barni, S. Goldenstein, Z. Dias, and A. Rocha, “Multiple parenting identification in image phylogeny,” in *IEEE Int. Conference on Image Processing*, 2014, pp. 5347–5351.
- [4] A. de Oliveira, P. Ferrara, A. De Rosa, A. Piva, M. Barni, S. Goldenstein, Z. Dias, and A. Rocha, “Multiple parenting phylogeny relationships in digital images,” *IEEE Trans. on Information Forensics and Security*, vol. 11, no. 2, pp. 328–343, 2016.
- [5] National Institute of Standards and Technology (NIST), “The 2017 nimble challenge evaluation datasets,” <https://www.nist.gov/itl/iad/mig/nimble-challenge>, Jan. 2017.
- [6] L. Kennedy and S-F. Chang, “Internet image archaeology: Automatically tracing the manipulation history of photographs on the web,” in *ACM Intl. Conference on Multimedia*, 2008, pp. 349–358.
- [7] A. De Rosa, F. Uccheddu, A. Costanzo, A. Piva, and M. Barni, “Exploring image dependencies: a new challenge in image forensics,” in *IS&T/SPIE Electronic Imaging*, 2010, pp. 75410X–75410X.
- [8] Z. Dias, A. Rocha, and S. Goldenstein, “First steps toward image phylogeny,” in *IEEE Int. Workshop on Information Forensics and Security*, 2010, pp. 1–6.
- [9] R. Datta, D. Joshi, J. Li, and J.Z. Wang, “Image retrieval: Ideas, influences, and trends of the new age,” *ACM Computing Surveys*, vol. 40, no. 2, pp. 5, 2008.
- [10] Y. Ke, R. Sukthankar, and L. Huston, “Efficient near-duplicate detection and sub-image retrieval,” in *ACM Intl. Conference on Multimedia*, 2004, pp. 869–876.
- [11] W. Dong, Z. Wang, M. Charikar, and K. Li, “High-confidence near-duplicate image detection,” in *ACM Int. Conference on Multimedia Retrieval*, New York, NY, USA, 2012, pp. 1:1–1:8.
- [12] J. Yuan and X. Liu, “Product tree quantization for approximate nearest neighbor search,” in *IEEE Int. Conference on Image Processing*, Sept 2015, pp. 2035–2039.
- [13] H. Bay, A. Ess, T.uytelaars, and L. Van Gool, “Speeded-up robust features (surf),” *Comput. Vis. Image Underst.*, vol. 110, no. 3, pp. 346–359, June 2008.
- [14] T.uytelaars and K. Mikolajczyk, “Local Invariant Feature Detectors: A Survey,” *ACM Foundations and Trends in Computer Graphics and Vision*, vol. 3, no. 3, pp. 177–280, 2008.
- [15] D. Lowe, “Distinctive image features from scale-invariant keypoints,” *Springer Intl. Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.
- [16] J.B. Kruskal, “On the shortest spanning subtree of a graph and the traveling salesman problem,” *Proc. of the American Mathematical Society*, vol. 7, no. 1, pp. 48–50, 1956.
- [17] P. Papadimitriou, A. Dasdan, and H. Garcia-Molina, “Web graph similarity for anomaly detection,” *Springer Journal of Internet Services and Applications*, vol. 1, no. 1, pp. 19–30, 2010.