

IMAGE PHYLOGENY TREE RECONSTRUCTION BASED ON REGION SELECTION

Paolo Bestagini, Marco Tagliasacchi, Stefano Tubaro

Dipartimento di Elettronica, Informazione e Bioingegneria
Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133, Milano, Italy

ABSTRACT

Nowadays, everyone can download, edit and republish any picture on the web, thus contributing to the diffusion of near-duplicate (ND) images. In order to gain an interesting insight on the way NDs are distributed online, recent works have focused on the reconstruction of the image phylogeny tree (IPT), i.e., an acyclic graph describing the genealogical relationship between ND image pairs. IPT reconstruction methods typically leverage the possibility of reconstructing one image from another one only if they are in parent-child relationship. However, as estimating the possible parent-child transformation is computationally expensive, usually a limited set of global editing operations is considered (i.e., compression, geometric and colour transformations applied to the whole image). However, in a real-world scenario it is customary to edit images also using local operations (e.g., logo insertion, object removal, splicing, etc.), which hinder the possibility of correctly estimating the parent-child relationship. In this paper, we propose an algorithm for IPT reconstruction that deals with the presence of local editing operations.

Index Terms— Image phylogeny, near-duplicates

1. INTRODUCTION

Thanks to the increasing diffusion of multimedia sharing platforms, as well as the high availability of image capturing devices and editing software, the amount of user-generated content available online is growing over time. In order to regulate the spread of illegal or tampered with images, the forensic community has deeply studied image forensics during the years [1]. As an example, many algorithms have been proposed to either verify image integrity [2], localize forgeries [3, 4, 5], or detect the use of a specific processing operation [6, 7, 8].

However, many images uploaded on the web are re-published duplicates of material already available online [9]. For example, different users may post the same pictures referred to the trend of the moment on social networks. Alternatively, newscasts from different broadcasters may publish the same pictures relative to an important event or person. However, re-published images are seldom identical copies of the original ones. Conversely, they often undergo editing operations to either artistically enhance their quality (e.g., colour correction, filtering, etc.), or to make them compliant with a broadcaster standard (e.g., resize, crop, logo insertion, etc.). Images obtained editing the same original content are usually referred to as near-duplicates (NDs) [10, 11].

The possibility of exploiting a set of near-duplicates rather than a object image for forensic analysis has paved the way to the development of novel forensic algorithms that synergistically take into account the information coming from all of them to perform even deeper forensic analyses [12, 13, 14]. As a matter of fact, recent works have shown that it is possible to study the phylogeny of near-duplicate objects (i.e., their genealogical relationships) in order to

shed a very interesting insight on the way content has been modified and re-distributed [15, 16, 17].

A common way to carry out phylogenetic investigations consists in reconstructing the image phylogeny tree (IPT), which is a directional acyclic graph describing the parent-child relationship between each ND image pair [18, 19]. By analysing the IPT, it is then possible to infer which image (the root of the tree) has been used to generate the others (the remaining nodes of the tree). This allows to trace back an image to its owner to either solve authenticity or copyright issues.

In this paper, we focus on IPT reconstruction starting from a pool of near-duplicate images. In order to reconstruct the IPT, state-of-the-art algorithms approximately share a common pipeline [20]: i) a dissimilarity value (proportional to how likely two images are in parent-child relationship) is computed for each image pair; ii) the set of dissimilarity measures is analysed to build the IPT. The proposed approaches differ in the way dissimilarity is computed, or dissimilarity values are aggregated. In this work, we focus on the first problem, i.e., how to compute dissimilarity comparing image pairs.

More specifically, the rationale behind dissimilarity computation is that a child image can be reconstructed applying a set of editing operations to the parent one, whereas the parent can be obtained starting from the children only in a very distorted way. Therefore, it is possible to map an image into another one through a set of transformations if and only if they are in parent-child relationship. The key for a robust dissimilarity computation is the correct estimation of the transformations that possibly map an image into another. As, searching for these transformations is computationally expensive, commonly used methods assume that only a limited number of global transformations are considered (i.e., compression, geometric and colour transformations applied to the whole image). However, this assumption hinders the possibility of correctly compute dissimilarity and estimating the parent-child relationship in presence of local transformations commonly used in a real-world scenario (e.g., logo addition, inpainting, object insertion, etc.).

In this paper, we propose an algorithm for IPT reconstruction that specifically takes into account the possible use of local editing operations, still keeping computational complexity at bay. In particular, we propose a method to automatically localize areas that have been locally manipulated by spotting the differences between near-duplicate image pairs. This allows to exclude the locally manipulated regions from dissimilarity computation, considering only the area where global transformations have actually been applied. Notice that, even if this technique is presented for still images, it can be readily used on video frames as well. Once dissimilarity has been computed, IPT is reconstructed using a state-of-the-art dissimilarity aggregation algorithm. The experimental campaign shows that we are able to accurately estimate the locally forged region and that the derived dissimilarity value improves the accuracy of IPT reconstruction.

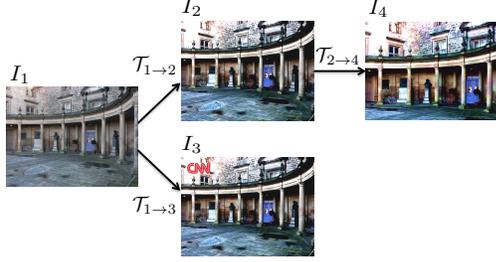


Fig. 1: Image phylogeny tree whose root is I_1 . Each node represents an image and the directionality of the graph represents parent-child relationships.

2. BACKGROUND AND PROBLEM FORMULATION

Two images are considered near-duplicates if they have been generated from a common image applying any number of non-invertible transformations within a set of admissible operations [10, 20]. More formally, let us define a $X \times Y$ resolution image as $I(x, y)$. Dropping the (x, y) indexes where not strictly necessary in favour of a more compact notation, N images I_n , $n \in \{1, \dots, N\}$ are defined as near-duplicates if $I_n = \mathcal{T}_{1 \rightarrow n}(I_1)$, where $\mathcal{T}_{1 \rightarrow n} \in \mathbb{T}$ is a combination of transformations within the set \mathbb{T} of possible ones. Notice that, as transformations are in most cases non-invertible, the parent-image can be mapped into the children (i.e., $\mathcal{T}_{1 \rightarrow n} \in \mathbb{T}$), but not the vice versa (i.e., $\mathcal{T}_{n \rightarrow 1} \notin \mathbb{T}$). Given a pool of ND images, it is possible to describe the parental relationships between them by means of an acyclic directed graph called image phylogeny tree (IPT). Fig. 1 shows an example of IPT, where each node represents an image and directional edges link images in parent-child relationship.

State-of-the-art algorithms for IPT reconstruction share a common pipeline: i) images are pair-wise compared to compute the dissimilarity matrix D that measures the likelihood of images to be in parent-child relationship; ii) dissimilarity matrix D is analysed to reconstruct the IPT.

A thoroughly validated way to compare images and build dissimilarity matrix is reported in [10, 20]. The rationale is that an image can be reconstructed applying a set of transformations to another image if and only if the latter is parent of the former. Therefore, given two images I_n and I_m , first the transformation that best maps I_m into I_n is computed as

$$\hat{\mathcal{T}}_{m \rightarrow n} = \arg \min_{\mathcal{T}} \mathcal{L}(I_n, \mathcal{T}(I_m)), \quad (1)$$

where \mathcal{L} is any distance metric, and $\hat{\mathcal{T}}_{m \rightarrow n} \in \mathbb{T}$. Dissimilarity is then computed as

$$D_{m,n} = \mathcal{L}(I_n, I_{m \rightarrow n}), \quad (2)$$

where $I_{m \rightarrow n} = \hat{\mathcal{T}}_{m \rightarrow n}(I_m)$ is the best estimate of I_n obtained from I_m . If I_m is parent of I_n , a transformation that maps I_m into I_n exists (ideally $\hat{\mathcal{T}}_{m \rightarrow n} = \mathcal{T}_{m \rightarrow n}$), thus leading to a small dissimilarity value $D_{m,n}$. Conversely, as a transformation that maps I_n into I_m does not exist (i.e., a parent cannot be reconstructed by the children), dissimilarity $D_{n,m}$ will be higher.

Dissimilarity matrix D describes a complete directed graph, where each node is an image, and edges are dissimilarity values. Once it has been computed, it is possible to reconstruct the IPT as the minimum spanning tree of this graph using different state-of-the-art algorithms such as Oriented Kruskal (OK) [11] or Optimum Branchings (OB) [21].



Fig. 2: With reference to the IPT of Fig. 1, I_3 (a) cannot be reconstructed from I_1 (b) due to the presence of the logo that appears in their difference (c).

The most computationally expensive task is $\hat{\mathcal{T}}_{m \rightarrow n}$ estimation between every image pair (i.e., $N(N-1)$ times), achieved by minimising (1). For this reason, it is common to adopt a heuristic algorithm that works under the assumption that \mathbb{T} is limited to a set of global operations, i.e., resizing, rotation, cropping, colour transformations and JPEG compression [10, 20]. As a matter of fact, $\hat{\mathcal{T}}_{m \rightarrow n}$ is estimated as the combination of three transformations: i) $\hat{\mathcal{T}}_{m \rightarrow n}^{\text{geo}}$ geometrically aligns I_m to I_n using keypoints correspondences (e.g., using SURF features [22]); ii) $\hat{\mathcal{T}}_{m \rightarrow n}^{\text{col}}$ matches the colour histogram of I_m to I_n , and; iii) $\hat{\mathcal{T}}_{m \rightarrow n}^{\text{JPEG}}$ compresses I_m with the JPEG quality factor used for I_n .

However, as in a real-world scenario it is common to use local image editing operations (e.g., object insertion), an estimate of $\hat{\mathcal{T}}_{m \rightarrow n}$ based on global operations only does not approximate well $\mathcal{T}_{m \rightarrow n}$ on the locally forged region. An example is given in Fig. 2. The presence of a logo in I_3 prevents it to be well reconstructed from I_1 , thus leading to a high dissimilarity value that does not well represent the parent-child relationship between the two images. In the following we formally analyse this issue and show how to solve it.

3. THE EFFECT OF LOCAL MANIPULATION

In this section we analyse the effect of local manipulation on dissimilarity. To this purpose, let us consider two images I_p (parent) and I_c (children) in parent-child relationship and the mean squared error as distance metric \mathcal{L} for dissimilarity computation as in [10, 20]. The approach can be easily extended to other dissimilarity measures. Defining the image residual as $R_{p,c} = I_{p \rightarrow c} - I_c$, dissimilarity between I_p and I_c can be written as

$$D_{p,c} = \frac{1}{XY} \sum_{(x,y)} |R_{p,c}(x,y)|^2. \quad (3)$$

With these definitions at hand, let us analyse dissimilarity behaviour in presence of either global or local transformations.

Global transformations. Let us now consider that $I_c = \mathcal{T}_{p \rightarrow c}(I_p)$ has been generated from I_p using only global transformations (as I_2 was generated from I_1 in Fig. 1). The transformation $\hat{\mathcal{T}}_{p \rightarrow c} = (\hat{\mathcal{T}}_{p \rightarrow c}^{\text{JPEG}} \circ \hat{\mathcal{T}}_{p \rightarrow c}^{\text{col}} \circ \hat{\mathcal{T}}_{p \rightarrow c}^{\text{geo}})$, where \circ represents function composition, can be estimated according to [10, 20] as outlined in Section 2 (and $\hat{\mathcal{T}}_{c \rightarrow p}$ as well). Image residuals become

$$R_{p,c} = I_{p \rightarrow c} - I_c = \hat{\mathcal{T}}_{p \rightarrow c}(I_p) - \mathcal{T}_{p \rightarrow c}(I_p) \simeq 0, \quad (4)$$

$$R_{c,p} = I_{c \rightarrow p} - I_p = \hat{\mathcal{T}}_{c \rightarrow p}(I_c) - I_p, \quad (5)$$

where the approximation to zero in (4) holds as $\hat{\mathcal{T}}_{p \rightarrow c} \simeq \mathcal{T}_{p \rightarrow c}$. As expected $D_{p,c}$ will be lower than $D_{c,p}$.

Local transformations. Let us now consider the children I_c^{loc} , generated as I_c with the addition of local manipulation on a given area (as I_3 was generated from I_1 in Fig. 1). More formally, let us describe the locally manipulated pixels of I_c^{loc} by means of a mask $M_c(x, y)$ the same size of I_c^{loc} defined as

$$M_c(x, y) = \begin{cases} 1, & \text{if } I_c^{\text{loc}}(x, y) = I_c(x, y), \\ 0, & \text{if } I_c^{\text{loc}}(x, y) \neq I_c(x, y), \end{cases} \quad (6)$$

where the inequality is due to the presence of a local operation. If the locally manipulated area is sufficiently small, we can still estimate the global transformation $\hat{T}_{p \rightarrow c}$ by the sole knowledge of I_c^{loc} and I_p . However, the residual in the parent-to-child case becomes

$$\begin{cases} R_{p,c}^{\text{loc}}(x, y) = R_{p,c}(x, y) \simeq 0, & \text{if } M_c(x, y) = 1, \\ R_{p,c}^{\text{loc}}(x, y) \neq R_{p,c}(x, y), & \text{if } M_c(x, y) = 0, \end{cases} \quad (7)$$

This means that local manipulation yields high local dissimilarity, thus $D_{p,c}^{\text{loc}} > D_{p,c}$. This is in contrast to our goal of keeping $D_{p,c}^{\text{loc}}$ as low as possible to well describe the parent-child relationship.

A solution to the aforementioned problem is to modify dissimilarity in (3) as

$$D_{p,c}^{\text{loc}} = \frac{1}{|\mathcal{M}_c|} \sum_{(x,y) \in \mathcal{M}_c} |R_{p,c}^{\text{loc}}(x, y)|^2, \quad (8)$$

where $\mathcal{M}_c = \{(x, y) | M_c(x, y) = 1\}$ defines the non-locally manipulated region and $|\mathcal{M}_c|$ is the cardinality of \mathcal{M}_c . In the following we show how this is possible.

4. PROPOSED ALGORITHM

In this section we propose our algorithm to localise the locally manipulated region, and show how to exploit it for IPT reconstruction.

Region selection. In order to estimate M_c , we exploit the possibility of estimating I_c from the sole knowledge of I_p and I_c^{loc} . More specifically, we estimate the global transformation $\hat{T}_{p \rightarrow c}$ from I_p and I_c^{loc} , then approximate I_c as

$$\hat{I}_c = \hat{T}_{p \rightarrow c}(I_p) \simeq I_c. \quad (9)$$

Notice that, since we only applied global transformations, \hat{I}_c does not present any traces of local manipulation. It is then possible to compute an estimation of $I_{c \rightarrow p}$ as

$$\hat{I}_{c \rightarrow p} = \hat{T}_{c \rightarrow p}(\hat{I}_c), \quad (10)$$

and the derived residual become

$$\begin{aligned} \hat{R}_{c,p} &= \hat{I}_{c \rightarrow p} - I_p = \hat{T}_{c \rightarrow p}(\hat{I}_c) - I_p = \\ &= \hat{T}_{c \rightarrow p}(\hat{T}_{p \rightarrow c}(I_p)) - I_p \simeq \hat{T}_{c \rightarrow p}(I_c) - I_p. \end{aligned} \quad (11)$$

It is important to notice that $\hat{R}_{c,p} \simeq R_{c,p}$ does not depend on I_c^{loc} , thus it does not present traces left by local manipulation on any pixel. Therefore we can compare $\hat{R}_{c,p}$ and $R_{c,p}^{\text{loc}}$ to infer which pixels of $R_{c,p}^{\text{loc}}$ are affected by the presence of local manipulations (i.e., $\{(x, y) | M_p(x, y) = 0\}$).

This is done following a probabilistic approach. More specifically, for pixel locations $\{(x, y) | M_p(x, y) = 1\}$, we expect $|R_{c,p}^{\text{loc}}(x, y)|$ to assume low values as $|\hat{R}_{c,p}(x, y)|$. Conversely, for pixel locations $\{(x, y) | M_p(x, y) = 0\}$, we expect $|R_{c,p}^{\text{loc}}(x, y)|$ to assume high values due to the presence of local manipulations.

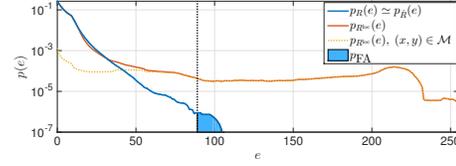


Fig. 3: $p_R(e)$ (blue line) assumes mostly low values, whereas $p_{R^{\text{loc}}}(e)$ (orange line) assumes also higher values. Values of $p_{R^{\text{loc}}}(e)$ in the local tampered region (yellow dashed line) dominate the tail of $p_{R^{\text{loc}}}(e)$, proving that tail values well describe the local edited area. Choosing a threshold Γ (black dashed line) it is possible to fix a p_{FA} and separate the most part of locally edited pixels.

Therefore, the probability mass function (pmf) $p_{\hat{R}}(e)$ of $|\hat{R}_{c,p}|$ approximately follows the pmf $p_R(e)$ of $|R_{c,p}|$, which is concentrated around low values. On the other hand the pmf $p_{R^{\text{loc}}}(e)$ of $|R_{c,p}^{\text{loc}}|$ will present higher values corresponding to the manipulated pixels. An example is shown in Fig. 3. The goal is to detect from which value of the $p_{R^{\text{loc}}}(e)$ tail belong pixels in the locally edited area. This can be done by setting a threshold Γ that fixes a given probability of false alarm $p_{FA} = \sum_{e=\Gamma}^{\infty} p_R(e) \simeq \sum_{e=\Gamma}^{\infty} p_{R^{\text{loc}}}(e)$. We then estimate the mask as

$$\hat{M}_p(x, y) = \begin{cases} 0, & \text{if } R_{c,p}^{\text{loc}}(x, y) \geq \Gamma, \\ 1, & \text{otherwise.} \end{cases} \quad (12)$$

The mask \hat{M}_c in the I_c domain is then estimated from \hat{M}_p by simply warping its domain according to the transformation $\hat{T}_{p \rightarrow c}^{\text{geo}}$.

In order to reduce noise in the estimated mask (i.e., small areas of isolated pixels not due to local tampering), some morphological operations can be applied to \hat{M}_c . In this work we decided to apply closing (with a squared structuring element of 5×5 pixels) followed by erosion (with a squared structuring element of 11×11 pixels).

IPT reconstruction. With the ability of estimating the mask \hat{M}_c , we can reconstruct the IPT modifying the common pipeline in [11]. More specifically, given an image pair I_p and I_c^{loc} , we compute: i) the global transformation $\hat{T}_{p \rightarrow c} = (\hat{T}_{p \rightarrow c}^{\text{JPEG}} \circ \hat{T}_{p \rightarrow c}^{\text{loc}} \circ \hat{T}_{p \rightarrow c}^{\text{geo}})$ using the approach in [11]; ii) the mask \hat{M}_p on I_p domain according to (12); iii) the mask \hat{M}_c on I_c^{loc} domain as $\hat{M}_c = \hat{T}_{p \rightarrow c}^{\text{geo}}(\hat{M}_p)$; iv) the residual $R_{p,c}^{\text{loc}} = \hat{T}_{p \rightarrow c}(I_p) - I_c^{\text{loc}}$; v) dissimilarity value $D_{p,c}^{\text{loc}}$ according to (8). Once dissimilarity matrix D has been computed, we apply the OB algorithm [21] to reconstruct the IPT.

5. EXPERIMENTAL RESULTS

In order to validate the proposed method, we tested it on a pool of IPTs with different sizes and topologies. Each tree was generated randomly selecting as root a 512×384 resolution image from the UCID database [23]. For each tree, the root was JPEG compressed. All the other images were generated randomly selecting a node of the tree as parent and applying: i) a combination of up to five randomly selected global transformations among contrast enhancement, brightness adjustment, resizing and cropping); ii) optionally (i.e., with probability p_{loc}) a local manipulation; and iii) JPEG compression with randomly selected quality between 50% and 100%. As local manipulation we considered object insertion, which is a very commonly used technique of photo manipulation. More specifically we inserted a logo randomly chosen from a selection of nine different ones in a random position within the image. Each logo approximately covers the 8% of the pixels of an image.

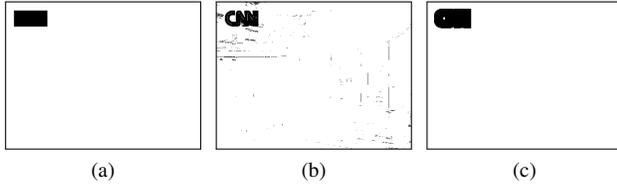


Fig. 4: With reference to I_1 and I_3 of Fig. 1, the ground truth M_p (a), the estimated M_p with $p_{FA} = 0.02$ (b) and M_p after morphological operations.

Table 1: Mask estimation results evaluated on $\mathcal{D}_{10}^{p=0.7}$. The target p_{FA} is reached using morphological operations (reported in bold).

p_{FA}	Morph.	TPR	TNR	FPR	FNR
0.02	<input type="checkbox"/>	0.957	0.961	0.039	0.043
	<input checked="" type="checkbox"/>	0.978	0.978	0.022	0.022
0.04	<input type="checkbox"/>	0.966	0.932	0.068	0.034
	<input checked="" type="checkbox"/>	0.984	0.962	0.038	0.016
0.06	<input type="checkbox"/>	0.970	0.908	0.092	0.030
	<input checked="" type="checkbox"/>	0.987	0.947	0.053	0.013

To analyse many scenarios, we generated different datasets of IPTs changing the number of nodes N per tree (i.e., 10 or 20) and the probability p_{loc} of applying local manipulation from one node to the other (i.e., 0, 0.3, 0.5 and 0.7). We denote each dataset as $\mathcal{D}_N^{p=p_{loc}}$. For each combination of parameters, we generated 100 different trees, for a total number of 12000 images. Notice that $p_{loc} = 0$ means that no local manipulations are applied, which is the working condition of the state-of-the-art method [21].

Mask estimation. To evaluate the accuracy of the mask estimation algorithm, we considered the 1000 images of dataset $\mathcal{D}_{10}^{p=0.7}$. For each parent-child image pair we estimated the local manipulation mask \hat{M}_p and compared it with the ground truth M_p . We tested different Γ obtained for different probabilities of false alarm p_{FA} imposed. We also considered masks obtained either with or without the use of morphological operations. An example of obtained masks is reported in Fig. 4.

The goodness of the obtained masks is evaluated using as metrics: i) true positive rate (TPR), i.e., the percentage of pixels correctly estimated as locally edited, and; ii) true negative rate (TNR), i.e., the percentage of pixels correctly estimated as not locally edited. For convenience, false negative rate (FNR) and false positive rate (FPR) are also reported.

Table 1 reports the values of the evaluation metrics averaged on all the considered image pairs for all the performed tests. Both TPR and TNR always exceed 90%. Moreover, if we compare the reference p_{FA} with the achieved FPR, we notice that the trend is always verified (i.e., increasing p_{FA} also FPR grows). Additionally, using morphological operations to remove some spurious pixels from the estimated masks we are able to predict the threshold Γ that actually ensures a given p_{FA} .

IPT reconstruction. We evaluated the IPT reconstruction algorithm based on mask estimation comparing our results to those obtained with the baseline state-of-the-art method [21]. Our method was used considering estimated masks \hat{M} (fixing $p_{FA} = 0.04$) for the removal of the local edited area. For each tree of each dataset we then reconstructed the IPT, and evaluated the performance of each algorithm according to the metrics commonly used in image phylogeny works [16, 20]: i) *root* counts whether the root of the tree is correctly identified (i.e., 1) or not (i.e., 0); ii) *edges* is the percentage of directed edges correctly identified; iii) *leaves* measures the per-

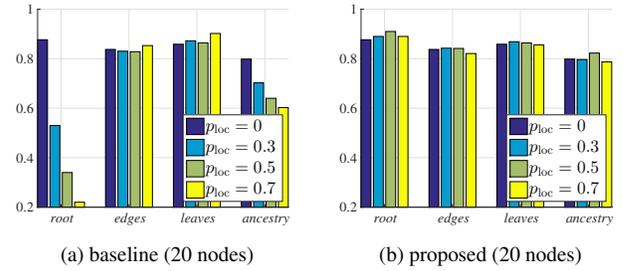


Fig. 5: Results of IPT reconstruction on trees of 20 nodes using the baseline method [21] (a)(b) and our algorithm (c)(d).

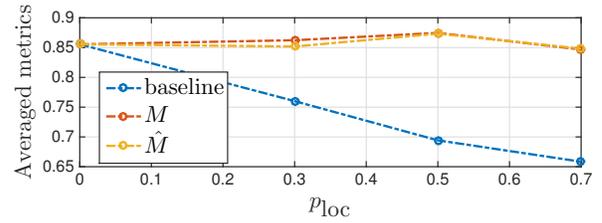


Fig. 6: Averaged metrics values on all the datasets for increasing p_{loc} . Baseline method [21] is compared to our algorithm using either ground truth (M) or estimated (\hat{M}) masks.

centage of leaves correctly detected; iv) *ancestry* is the percentage of ancestral relationships correctly guessed.

Fig. 5 shows the achieved IPT reconstruction results on 20 nodes trees (i.e., $\mathcal{D}_{20}^{p=0}$, $\mathcal{D}_{20}^{p=0.3}$, $\mathcal{D}_{20}^{p=0.5}$ and $\mathcal{D}_{20}^{p=0.7}$) with the baseline method and with our algorithm based on the estimated masks. Results for 10 nodes trees are not reported as the behaviour is exactly the same. It is interesting to notice that our method practically achieves the same results as the baseline when no local manipulations are applied (i.e., $p_{loc} = 0$). However, the baseline method suffers from the presence of local manipulations, specifically concerning the metrics *root* and *ancestry*.

To highlight the accuracy gain of the proposed algorithm, we aggregated the four metrics for trees with the same p_{loc} , considering both 10 and 20 nodes trees. In this case, we also tested our method applying the ground truth masks M . Fig. 6 shows the dependency between the averaged metrics and p_{loc} . When $p_{loc} = 0$ (i.e., no local manipulation), our method reaches the state-of-the-art. However, as p_{loc} increases, the baseline method loses in accuracy, while the proposed one does not suffer any accuracy decrease. Notice that using ground truth or estimated masks does not change the achieved results, thus masks are correctly estimated.

6. CONCLUSIONS

In this paper we proposed a method for IPT reconstruction that improves over state-of-the-art in presence of locally manipulated ND images. The algorithm localises the locally manipulated areas and excludes them from dissimilarity computation. Dissimilarity values are then aggregated using the OB algorithm.

In addition to IPT reconstruction, the proposed region selection method can be fruitfully applied to other scenarios where similar images are under analysis. As an example, it can be applied for tampering detection in ND images [24], or forgery detection exploiting novel camera thumbnails-based techniques [25].

7. REFERENCES

- [1] A. Piva, "An overview on image forensics," *ISRN Signal Processing*, vol. 2013, pp. 22, 2013.
- [2] D. Cozzolino, D. Gragnaniello, and L. Verdoliva, "Image forgery detection through residual-based local descriptors and block-matching," in *IEEE International Conference on Image Processing (ICIP)*, 2014.
- [3] P. Ferrara, T. Bianchi, A. De Rosa, and A. Piva, "Image forgery localization via fine-grained analysis of CFA artifacts," *IEEE Transactions on Information Forensics and Security*, vol. 7, pp. 1566–1577, 2012.
- [4] D. Cozzolino, D. Gragnaniello, and L. Verdoliva, "Image forgery localization through the fusion of camera-based, feature-based and pixel-based techniques," in *IEEE International Conference on Image Processing (ICIP)*, 2014.
- [5] I. Amerini, R. Becarelli, R. Caldelli, and A. Del Mastio, "Splicing forgeries localization through the use of first digit features," in *IEEE International Workshop on Information Forensics and Security (WIFS)*, 2014.
- [6] M.C. Stamm and K.J.R. Liu, "Forensic detection of image manipulation using statistical intrinsic fingerprints," *IEEE Transactions on Information Forensics and Security (TIFS)*, vol. 5, pp. 492–506, 2010.
- [7] V. Conotter, P. Comesana, and F. Perez-Gonzalez, "Joint detection of full-frame linear filtering and JPEG compression in digital images," in *IEEE International Workshop on Information Forensics and Security (WIFS)*, 2013.
- [8] S. Milani, P. Bestagini, M. Tagliasacchi, and S. Tubaro, "Demosaicing strategy identification via eigenalgorithms," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2014.
- [9] L. Kennedy and S.-F. Chang, "Internet image archaeology: automatically tracing the manipulation history of photographs on the web," in *ACM International Conference on Multimedia (ACM-MM)*, 2008.
- [10] Z. Dias, A. Rocha, and S. Goldenstein, "First steps toward image phylogeny," in *IEEE International Workshop on Information Forensics and Security (WIFS)*, 2010.
- [11] Z. Dias, A. Rocha, and S. Goldenstein, "Image phylogeny by minimal spanning trees," *IEEE Transactions on Information Forensics and Security (TIFS)*, vol. 7, pp. 774–788, 2012.
- [12] S. Lameri, P. Bestagini, A. Melloni, S. Milani, A. Rocha, M. Tagliasacchi, and S. Tubaro, "Who is my parent? reconstructing video sequences from partially matching shots," in *IEEE International Conference on Image Processing (ICIP)*, 2014.
- [13] S. Milani, M. Fontana, P. Bestagini, and S. Tubaro, "Phylogenetic analysis of near-duplicate images using processing age metrics," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2016.
- [14] A. Melloni, S. Lameri, P. Bestagini, M. Tagliasacchi, and S. Tubaro, "Near-duplicate detection and alignment for multi-view videos," in *IEEE International Conference on Image Processing (ICIP)*, 2015.
- [15] A. De Rosa, F. Uccheddu, A. Piva, M. Barni, and A. Costanzo, "Exploring image dependencies: A new challenge in image forensics," in *SPIE Conference on Media Forensics and Security (MFS)*, 2010.
- [16] Z. Dias, S. Goldenstein, and A. Rocha, "Large-scale image phylogeny: Tracing image ancestral relationships," *IEEE MultiMedia (MM)*, vol. 20, pp. 58–70, 2013.
- [17] F. O. Costa, S. Lameri, P. Bestagini, S. Dias, A. Rocha, M. Tagliasacchi, and S. Tubaro, "Phylogeny reconstruction for misaligned and compressed video sequences," in *IEEE International Conference on Image Processing (ICIP)*, 2015.
- [18] Z. Dias, S. Goldenstein, and A. Rocha, "Toward image phylogeny forests: Automatically recovering semantically similar image relationships," *Forensic Science International (FSI)*, vol. 231, pp. 178–189, 2013.
- [19] F. O. Costa, M. Oikawa, Z. Dias, S. Goldenstein, and A. Rocha, "Image phylogeny forests reconstruction," *IEEE Transactions on Information Forensics and Security (TIFS)*, vol. 9, pp. 1533–1546, 2014.
- [20] A. Melloni, P. Bestagini, S. Milani, M. Tagliasacchi, A. Rocha, and S. Tubaro, "Image phylogeny through dissimilarity metrics fusion," in *European Workshop on Visual Information Processing (EUVIP)*, 2014.
- [21] D. Zanoni, G. Siome, and R. Anderson, "Exploring heuristic and optimum branching algorithms for image phylogeny," *Journal of Visual Communication and Image Representation (JVCIR)*, vol. 24, pp. 1124–1134, 2013.
- [22] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (SURF)," *Computer Vision and Image Understanding (CVIU)*, vol. 110, pp. 346–359, 2008.
- [23] G. Schaefer and M. Stich, "UCID - an uncompressed colour image database," in *SPIE Storage and Retrieval Methods and Applications for Multimedia (SPIE-SRMAM)*, 2004.
- [24] L. Gaborini, P. Bestagini, S. Milani, M. Tagliasacchi, and S. Tubaro, "Multi-clue image tampering localization," in *IEEE International Workshop on Information Forensics and Security (WIFS)*, 2014.
- [25] M. Kirchner, P. Winkler, and H. Farid, "Impeding forgers at photo inception," in *SPIE Conference on Media Watermarking, Security, and Forensics*, 2013.