Proceedings of the 5th International Symposium on Communications, Control and Signal Processing,
ISCCSP 2012, Rome, Italy, 2-4 May 2012

SALIENCY DETECTION AS A SUPPORT FOR IMAGE FORENSICS

Oleg Muratov, Duc-Tien Dang-Nguyen, Giulia Boato, Francesco G.B. De Natale

Department of Information and Computer Science - University of Trento
via Sommarive 5, 38123 Trento - Italy
{muratov, dangnguyen, boato}@disi.unitn.it, denatale@ing.unitn.it

ABSTRACT

In this work we introduce a new application of saliency detection as a support for forensic analysis in digital images. Indeed, although there are some attempts to achieve both detection and localization of image tampering, the majority of forensics tools require hypothesis about the modified area. Since we claim that salient objects/subjects, conveying the semantic content of the image, are the regions whose integrity is more critical, we propose to combine salient object detection and forensic analysis on the corresponding output. To this aim we introduce here an improved version of a saliency map extractor based on segmentation and describe the application of a tampering detection method allowing digital composite detection. Experimental results are presented and discussed.

Index Terms— Visual saliency, digital image forensics, splicing detection

1. INTRODUCTION

Nowadays the use of media contents is widespread and everyday increasing in various fields like for example journalism and social networks, where an enormous quantity of user generated contents are available. The diffusion of very flexible and easy-to-use photo retouching tools allows users to improve and/or modify their content before uploading and sharing it. Thus, the availability of tools able to verify the integrity of media contents is becoming everyday more critical.

To this aim digital watermarking techniques and more recently digital forensics methods were developed to protect multimedia data applying active and passive analysis allowing authenticity and integrity verification. In the last decade, different tampering detection algorithms were developed relying on pixel-based, format-based or geometric-based analysis (see [1] and reference therein) or more recently on quality metrics [2] to pick out possible forgeries.

Forensic analysis is usually applied on the whole image or rely on some hypothesis about the tampered area. Only recently a few techniques were proposed to automatically localize forgeries. In particular, aligned and non-aligned double JPEG compression artifacts are revealed and localized in [3] (improved version of [4]) and [5], respectively. Although such attempts to allow both detection and localization of tampering, the majority of algorithms still require hints about the to-be-tested region in order to avoid brute force analysis.

Therefore, we propose here to use saliency detection as a support to forensic analysis. Indeed, we assume salient objects/subjects to be the regions of an image whose integrity is more critical and important to be verified, since the semantic content of the photo is highly connected to them. We believe that modification of salient regions of the image can strongly modify the message conveyed by it and thus possibly bias the viewer [6][7].

In this paper, we combine salient object detection and forensic analysis on it. To this aim we introduce here an improved version of a saliency map extractor based on segmentation which is able to support object highlight [8]. This tool automatically detects important objects/subjects of the image where tampering detection can then be applied. As a proof of concept, we use here the method proposed by Farid in [9] which allows digital composite detection via analysis of double JPEG compression traces.

The rest of the paper is organized as follows: Section 2 is devoted to the detailed description of our method for saliency map extraction, in Section 3 we describe the forensic analysis applied on salient objects, in Section 4 we show some experimental results, while Section 5 draws some concluding remarks.

2. SALIENCY MAP EXTRACTION

In this paper we use the bottom-up saliency detection method described in [8] with some modification. This method is based on the analysis of low-level features extracted from a segmented image. The use of segmented image representation allows the classification on a segment-wise level. This results in an output map containing regions that to some extend are consistent with objects on the image. In addition, it reduces the amount of noisy estimation due to the "averaging" effect within segments described in the following. A schema is depicted in Figure 1 and described in the following.

Since visual saliency is commonly caused by irregularities (as can be verified using eye-trackers [10]), the features used in the method are aimed at detecting these irregularities.

978-1-4673-0276-0/12/$31.00 ©2012 IEEE

The features employed in the method can be divided into two groups. One group consists of features extracted from the entire image; the other group in contrast stands for features extracted directly from segments. Here, we introduce some new features with respect to [8].

The first group includes luminance contrast and center-surround histograms features. Center-surround histograms measure the distance between histograms of the inner and outer windows moved through the image and is based on the observation that the histogram of the foreground region usually has wider peak regions in the spectrum compared to that of the background region. Since the features in the first group are computed on the global level their values for each segment are obtained by averaging value of the pixels that belong to the segment.

The second group includes the following features: dominant color, spatial location, size, geometric class and occlusion. The dominant color feature is computed in a 12 tone colorspace. This feature is used in two fashions: i) in conjunction with psychological studies on the impact of certain colors for human-attention, and ii) for detecting segments with colors different from others after normalization over all segments. The spatial location feature is based on the observation that amateurs tend to place the most interesting part of the image into the center. It is computed as the location of segment's center of mass. The size feature plays a role of a filter not allowing for too big (likely to be a part of the background) or too small regions being misclassified as salient. In order to improve accuracy in the map extraction, we introduce also the detection of geometric class for each segment, by exploiting the classifier described in [11] that provides classes such as ground, sky, diagonal and vertical planes. This feature allows for discrimination between planes that are usually belong to background (sky, ground) and planes that with equal probability can either belong to background or foreground regions. Finally, occlusions are also considered and detected by comparing locations of centers of mass of neighbor regions, their shapes and sizes.

A Naive Bayesian classifier is used after training on a database with over 700 images with ground-truth data. The input of the classifier is a vector of features values assigned to a segment. The output is the probability of a segment to be salient. Combining probabilities and the segmentation map the output saliency map is constructed.

3. FORENSIC ANALYSIS

In our work we do not use the map directly, but instead we generate a bounding box from the corresponding map. This allows for more generalization of our method, due to the relatively high number of region-based forgery detection methods that require (or simply prefer) a bounding box as an input.

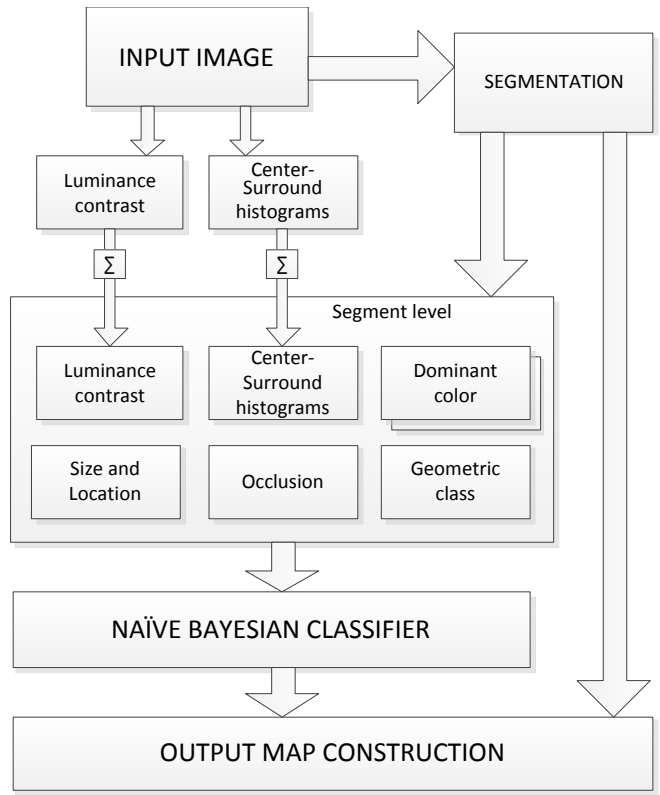

Fig. 1. The schema of the saliency detector.

3.1. Bounding box definition

Usually bounding boxes are made in such a way that they cover all the pixels belonging to the test region. The drawback of such boxes is that for shapes far from rectangular a relatively high number of pixels from neighbor regions are included. If we assume that we have a region that is in fact a forgery, and its neighbors are not, then the pixels included from its neighbors could lead to decrease in the accuracy of the estimation. In addition, saliency maps may include a part of an unmodified region labelled as salient, that leads to the same effect.

For this reason we use spread-based boxes that are aimed at diminishing of this effect by means of exclusion of boundary zones. Let us consider a salient region with $X_{min}, X_{max}, Y_{min}$ and Y_{max} denoting its most left, right, top and bottom coordinates, respectively. The spread-based box, with coordinates $R_{Left}, R_{Top}, R_{Width}$ and R_{Height}, is computed as follows:

$$\begin{aligned} R_{Left} &= (X_{min} + X_{CoM} - SP_{nx})/2 \\ R_{Top} &= (Y_{min} + Y_{CoM} - SP_{ny})/2 \\ R_{Width} &= (X_{min} + X_{max} + SP_{nx} + SP_{px})/2 \\ R_{Height} &= (Y_{min} + Y_{max} + SP_{ny} + SP_{py})/2 \end{aligned}$$

where X_{CoM} and Y_{CoM} are the coordinates of the center of

mass, while $SP_{nx}, SP_{px}, SP_{ny}$ and SP_{py} are the spreads of the region computed as follows:

$$SP_{nx} = S(Reg_{nx})/height(Reg_{nx})$$
$$SP_{px} = S(Reg_{px})/height(Reg_{px})$$
$$SP_{ny} = S(Reg_{ny})/width(Reg_{ny})$$
$$SP_{py} = S(Reg_{py})/width(Reg_{py})$$

where Reg_{nx} and Reg_{py} denote parts of the region that are to the left and to the right from its center of mass, respectively, Reg_{ny} and Reg_{py} are parts above and below the center of mass, and $S(Reg_i)$ is the area of the region.

Boxes computed in the proposed way may be considered as a downscaled version of a usual bounding box, although our approach allows for better shape aware results.

3.2. Splicing detection

Since forgeries are often created by combining several images, we focus on this kind of tampering usually called splicing. Moreover, if images used for the fake creation have different JPEG compression qualities, then inconsistencies within and outside the tampered region can be easily detected. This is a reasonable assumption since usually splicing are done exploiting images coming from different cameras, thus very often presenting different quality factors.

We adopt the method presented in [9], which allows to detect whether a specific region has a different compression quality than the rest of the image. The basic idea behind this technique is the following: tampered region with different JPEG qualities can be highlighted by analyzing the differences between the original image and its JPEG-compressed versions and by revealing local minima.

More in details, following [9], differences from the input image with each counterpart compressed with quality factor q are calculated and averaged across a $b \times b$ pixels region:

$$\delta(x,y,q) = \frac{1}{3}\sum_{i=1}^{3}\frac{1}{b^2}\sum_{b_x=0}^{b-1}\sum_{b_y=0}^{b-1} \left[f(x+b_x, y+b_y, i) - f_q(x+b_x, y+b_y, i)\right]^2$$

where (x, y) is the current pixel position, i is the color channel index, and $f(\cdot)$ represents the analyzed pixel value. All differences are then normalized as follows:

$$d(x,y,q) = \delta(x,y,q) - \frac{min_q\left[\delta(x,y,q)\right]}{max_q\left[\delta(x,y,q)\right] - min_q\left[\delta(x,y,q)\right]}$$

Finally, the Kolmogorov-Smirnov (K-S) statistic is used to determine whether the original image and the compressed counterpart at quality factor q are similar or distinct. The K-S statistic is defined as:

$$K = max_u \left|C_1(u) - C_2(u)\right|$$

where $C_1(u)$ and $C_2(u)$ are the cumulative probability distributions of two regions in $d(x, y, q)$. Values of the K-S statistic above a threshold T are considered as hints of tampering.

It can be noticed that the method requires an hypothesis on the region C_1 to be analyzed. Here, we apply the described procedure on the bounding box extracted from the saliency map as detailed in Section 3.1, considering C_1 with coordinates $R_{Left}, R_{Top}, R_{Width}$ and R_{Height} and C_2 the region of the image outside such a box.

4. EXPERIMENTAL RESULTS

In order to test the proposed approach we have created a set of test images. Forgery images were created by combining different images, cutting a foreground part from one image and pasting it onto the background of another image. Since the forgery detection method used is sensitive to re-sampling, if that was required than necessary statistics were recreated by compressing the original image at required scale and only after that it was used as a source. For localization of forgery region a map was extracted from an image followed by computation of the corresponding bounding box. The forgery was detected by measuring the difference of the output of JPEG ghost inside and outside the retrieved bounding box. We performed analysis with quality range from 15 to 90 with step 5 and the threshold T equal to 0.35.

Some results can be seen in Figure 2, where we show a few examples of tampered images, the corresponding ground truths, the saliency maps and the bounding boxes extracted with the proposed method, and finally a set of JPEG ghost maps used for splicing detection. In addition, to complete the analysis of the method we have performed tests on untampered images proving that the approach proposed does not interfere with false alarms rate.

In order to evaluate to what extend the estimation of the tampered region is precise, the agreement between bounding boxes extracted from a map and mask was used. The performance on the test dataset achieved 70% precision with 60% recall. This result can be considered as a quite good inasmuch as saliency detection by itself is an unsolved problem. Though the dataset was created with images containing a relatively heterogeneous content allowing for discrimination between background and foreground (tampered) parts, the background region still is subject to cause some amount of attention. For this reason for images with complex content the precision of foreground extraction decreases leading to the decrease in the overall performance. In addition, one could tamper a small portion of a visually unmeaningful region, that cannot be detected by the approach proposed. However, we do believe that in most cases one desires to change both semantically and visually meaningful part of the content and as it was shown above for such cases our method allows for localization and detection of tampered regions.

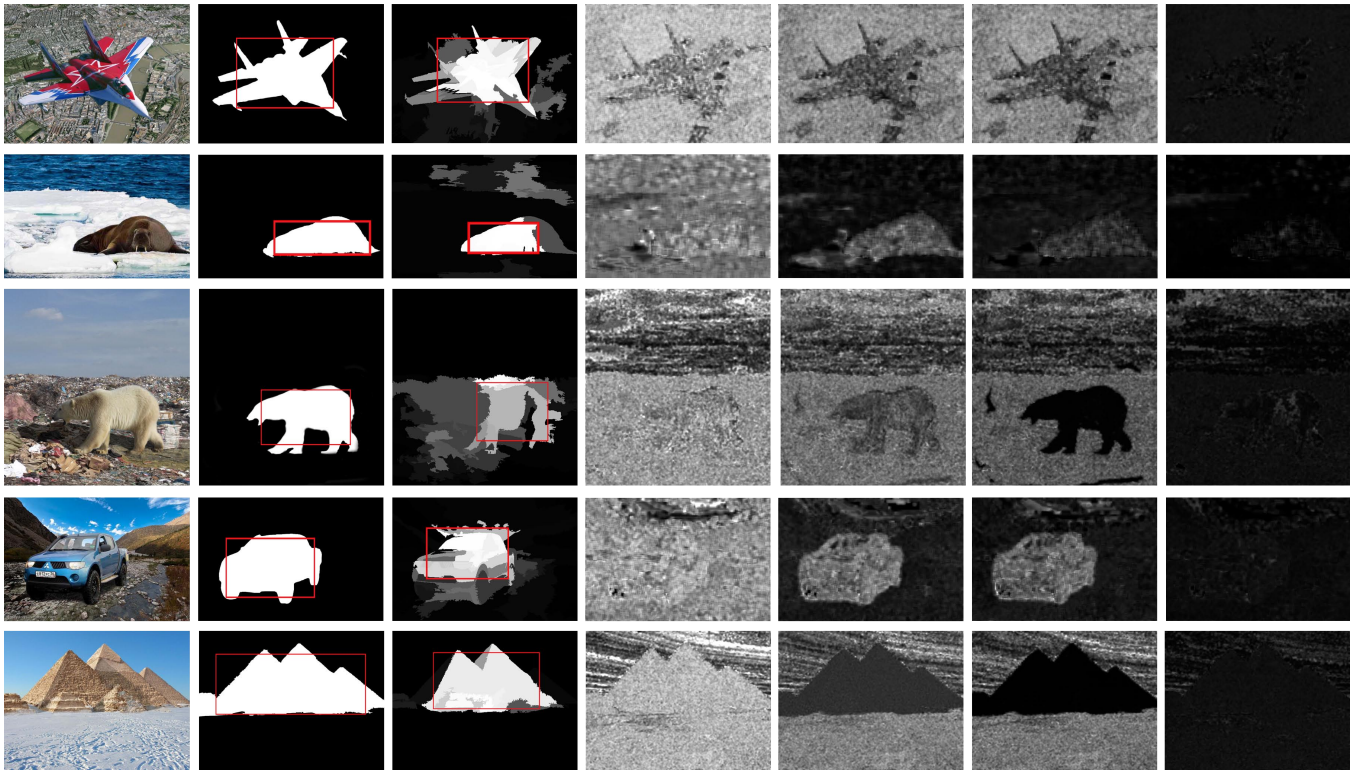

Fig. 2. Some examples of tampered images and their results. From left to right: first column: tampered images; second column: tampered region maps with the corresponding boxes; third column: saliency maps with the corresponding boxes extracted using the proposed method; and the last columns: JPEG ghost maps at compression qualities of 30, 45, 60, and 75.

5. CONCLUSION

We have proposed here a novel idea to apply forensics analysis on salient objects/subjects in images. To the best of our knowledge this is the first attempt to combine saliency map extraction and tampering detection. We assume salient objects/subjects to be the regions of an image whose integrity is more critical and important to be verified, since the semantic content of the photo is highly connected to them. Moreover, we believe that modification of salient regions of the image can strongly modify the message conveyed by it and thus possibly bias the viewer.

To this aim we have presented an improved version of a saliency map extractor based on segmentation which allows in the output map to highlight regions that are consistent with objects on the image. Thus, we have detailed the application of a tampering detection method allowing digital composite detection performed on salient objects.

The idea presented here can be applied with various tampering detection methodologies. Furthermore, we suggest to exploit a very nice and novel framework for application of different forensic tools and fusion of their results [12], which allows to automatically summarize all outputs to get both a binary and a soft interpretation of the global result.

6. ACKNOWLEDGMENTS

This work was supported by the European Union within the Seventh Framework project LivingKnowledge (IST-FP7-231126).

7. REFERENCES

[1] H. Farid, "A survey of image forgery detection," *IEEE Signal Processing Magazine*, vol. 2, no. 26, pp. 16–25, 2009.

[2] F. Battisti, M. Carli, and A. Neri, "Image forgery detection by means of no-reference quality metrics," *SPIE Electronic Imaging: Media Watermarking, Security, and Forensics*, 2012.

[3] T. Bianchi, A. De Rosa, and A. Piva, "Improved dct coefficient analysis for forgery localization in jpeg images," *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2011.

[4] Z. Lina, J. He, X. Tang, and C.-K. Tang, "Fast, automatic and fine-grained tampered jpeg image detection via dct coefficients analysis," *Pattern Recognition*, vol. 42, no. 11, pp. 2492–2501, November 2009.

[5] T. Bianchi and A. Piva, "Analysis of non-aligned double jpeg artifacts for the localization of image forgeries," *IEEE International Workshop on Information Forensics and Security*, 2011.

[6] P. Zontone, G. Boato, F.G.B. De Natale, A. De Rosa, M. Barni, A. Piva, J. Hare, D. Dupplaw, and P. Lewis, "Image diversity analysis: Context, opinion and bias," *International Workshop Living Web: Making Web Diversity a True Asset*, 2009.

[7] G. Boato, F.G.B. De Natale, and P. Zontone, "How digital forensics may help assessing the perceptual impact of image formation and manipulation," *International Workshop on Video Processing and Quality Metrics for Consumer Electronics*, 2010.

[8] O. Muratov, P. Zontone, G. Boato, and F.G.B. De Natale, "A segment-based image saliency detection," *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2011.

[9] H. Farid, "Exposing digital forgeries from jpeg ghosts," *IEEE Transactions on Information Forensics and Security*, vol. 4, no. 1, pp. 154–160, March 2009.

[10] A. Armanini and N. Conci, "Eye tracking as an accessible assistive tool," *International Workshop on Image Analysis for Multimedia Interactive Services*, 2010.

[11] D. Hoiem, A.A. Efros, and M. Hebert, "Geometric context from a single image," *IEEE International Conference of Computer Vision*, 2005.

[12] M. Fontani, T. Bianchi, A. De Rosa, A. Piva, and M. Barni, "A dempster-shafer framework for decision fusion in image forensics," *IEEE International Workshop on Information Forensics and Security*, 2011.