**State of the Art report for deliverable 2.1**

**MOD4**

**Overall description**

The MOD4 of Fence deals with splicing detection observing the 2D lighting inconsistencies in the image, based on the fact that a perfect illumination adjustment in an image composition is very hard to achieve.

In particular MOD4 aims at detecting if a given image is pristine or not and, in case it is fake, highlight the splicing in a generic image with a blind approach.

**Main sub-modules:**

1. Image illuminant maps estimation
2. Image segmentation
3. Single image segment classification
4. Integrity decision and fake image parts highlighting

**Related State of the Art Review**

Investigating image’s lighting is one of the most common approaches for splicing detection. This approach is particularly robust because of the fact that it's really hard to preserve the consistency of light dealing with forgeries.

In this scenario, there are mainly two types of approaches: the one based on the object-light geometric arrangement and the one based on illuminant colors. We focused our attention on the illuminant-based approach, which assumes that a scene is lit by the same light source. In this condition, single scenes have unique illuminant colors, potential inconsistencies could be used for splicing detection.

Illuminant Maps locally describes the lighting in a small region of the image. In the computer vision literature exists many different approaches for determining the illuminant of an image and they are divided into two groups*: statistical-based* approaches and *physics-based* approaches.

Regarding the first group, the method proposed by Van de Weijer *et al.* [3], called *Generalized Grayworld Estimate* (GGE), is an extension of Gray-World and Gray-Edge algorithms. The main assumption on which this method is based on is that, under a white light source, the average pixel color in a scene is achromatic.

For the latter group, was considered the method proposed by Riess *et al.* [4], called *Inverse Intense Chromaticity* (IIC), which tries to model the illuminants considering the dichromatic reflection model. Images are divided into blocks named *superpixels*, in order to achieve lighting condition suitable with the IIC space.

Carvalho *et al.* [1] then presents a method that relies on a combination of the two approaches for the detection of manipulations on images containing human faces. In addition to maps, a large set of shape and texture descriptors are used together. Human faces are considered because the fact that, from a theoretical viewpoint, it is advantageous to consider only image regions that consist of approximately the same underlying material.

In this article [1] it is also shown that the difference between the two maps, GGE and IIC, increased when fake images are processed. This insight leads to the idea of the ability to discriminate if part of the image has been doctored using a metric defined between the two IMs.

This result is shown in the publication of Schetinger *et al.* [2]. The authors show in the article that even just using a Euclidean distance between two different maps, doctored images highlight a big difference. It is therefore proposed a descriptor of the image which takes account of this fact. The idea of the authors is that the two maps have eigenvalues pointing in the same direction, being extracted from an with almost constant lighting conditions. In the case of spliced images this condition is no longer fulfilled. The image descriptor (or a generic ROI descriptor) will be given then by differences between different eigenvalues of the two maps according to some metric. At this point, using a classifier such as SVM, you can split the pristine images from those manipulated.

Starting from [2], the idea is to propose a blind method for detecting image splicing, using a default segmentation policy (e.g. by using quad-tree or kd-tree segmentation) and classify its content at different scale and resolution basing on a trained SVM model.

References:

[1] T. Carvalho, et al. *Illuminant-Based Transformed Spaces for Image Forensics*. IEEE Transactions on Information Forensics and Security 11.4 (2016): 720-733.

[2] V. Schetinger et al. *Exploring Statistical Differences Between Illuminant Estimation Methods for Exposing Digital Forgeries*; 2016.

[3] J. Van de Weijer, Th. Gevers, A. Gijsenij, *Edge-Based Color Constancy*, IEEE Trans. Image Processing, accepted 2007.

[4] C. Riess and E. Angelopoulou, *Scene illumination as an indicator of image manipulation*. In Proceedings of the 12th international conference on Information hiding, Berlin, Heidelberg, 66-80; 2010