# Illumination Inconsistency Sleuthing for Exposing Composition Telltales in Digital Images

**Tiago Carvalho**[1,2], **Hélio Pedrini**[2], **and Anderson Rocha**[2]

[1]National Center for Monitoring and Early Warning of Natural Disasters (CEMADEN)
São José dos Campos – SP – Brazil

[2]Institute of Computing – University of Campinas (UNICAMP)
Campinas – SP – Brazil

tiago.carvalho@cemaden.gov.br, {helio,anderson.rocha}@ic.unicamp.br

***Abstract.*** *Images are powerful communication tools. Due to this power it is often worrisome when image manipulations come into play allowing forgers to deceive viewers, change opinions or even affect how people perceive reality. Therefore, it is paramount to devise and deploy efficient and effective forgery detection techniques. From all types of image forgeries, composition images (forgeries using parts of two or more images) are of particular interest. Among different techniques for spotting forgeries, image illumination inconsistencies are the most promising. This work builds upon the hypothesis that "image illumination inconsistencies are strong and powerful evidence of image composition" and presents four original and effective approaches to detecting image forgeries. The first method explores eye specular highlight telltales to estimate the light source and viewer positions in an image. The second and third approaches explore the color phenomenon called metamerism, whereby two colors may appear to match under one light source but appear completely different under another one. Finally, the last approach relies on user's interaction to specify 3-D normals of suspect objects in an image from which the 3-D light source position can be estimated. These approaches represent important advances, which certainly will be a strong tool for forensic analysts against image forgeries.*

## 1. Introduction

In the last years, computer science advances are notorious and have been supporting improvements in different fields of science. Medicine, biology, chemistry, law, national security and many other areas are just some examples of knowledge domains whereby computational methods promoted significant progress. However, such computational advances are a double-edged sword. If on one hand they help make people's lives better, on the other, they can be used to prejudice people. In law, for example, fake evidences might be forged through computational methods, deceiving people and thwarting opinions.

Keeping focus in the law field, we can say that one of the most powerful types of evidences are images. Images are so compelling and influential that, in the past, people used to say: *an image is worth a thousand words*. Unfortunately, nowadays this expression is not a hundred percent true. Taking advantage of powerful and advanced computational tools, as Adobe Photoshop and GIMP packages, people are able to manipulate images in different ways. In special, *image splicing* is a very common and powerful type of image forgery. It consists in using parts of two or more images to compose a new

one, reflecting distorted realities, which can even influence people's memories of past moments [Carvalho et al. 2012, Carvalho et al. 2015c].

Aware of the constant struggle between the forensic community and counterfeiters, and looking for promoting justice with efficient and effective methods at this "arms race", this Ph.D. thesis presents four methods for detecting image splicing forgeries. Using concepts related to computer vision, digital image processing, machine learning and computer graphics, in this work we developed and deployed methods for exploring light inconsistencies to detect image splicing.

Our first proposed method explores how light is reflected on people's eyes and uses this information to provide indicative cues whether some image is an image splicing or not. Our second and third methods are based, roughly speaking, on light color and how this light interacts with people's skin. Finally, the last proposed method seeks for forgery telltales in a more general scenario, allowing users to detect forgeries using an estimation of 3D light source position in a 2D image.

As we will describe in the following sections, each one of the proposed methods provides important scientific contributions to the digital forensic community and together they represent a remarkable step toward a more robust and quantifiable set of tools for assessing digital image evidences.

## 2. Exposing Forgeries Based on Eye Specular Highlights

*The eyes never lie*. This strong sentence, allied with the fact that most of splicing images involve people, is the inspiration to our first method, which has two main contributions: (1) proposition of new image features for forensics not explored before; and (2) deployment of machine learning approaches (single and multiple classifier combination) to the decision-making process instead of relying on simple and limited hypothesis testings. The proposed method redoces the classification error when analyzing an image in more than $20\%$ when compared to the literature. Figure 1 depicts an overview of proposed method [Saboia et al. 2011].

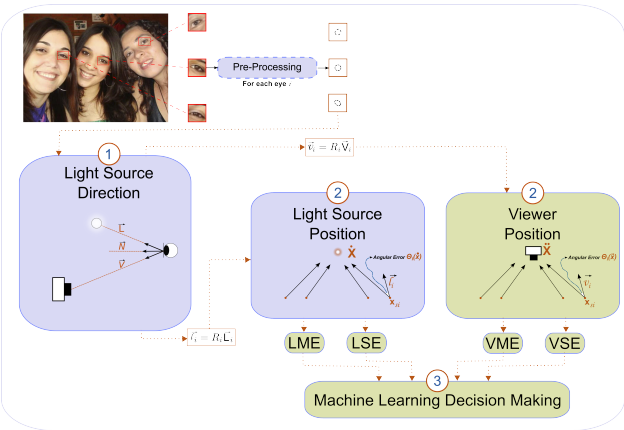



**Figure 1. First proposed method based on eye specular highlights – Overview.**

Given an image where people's eyes are visible, the proposed method consists in using specular highlights, reflected on people's eyes to estimate, for each person in the image, the 2D directions of the light source and the viewer (in this case, the viewer is

the camera). Once estimated, directions of all people in the image are used to estimate a single light source direction and a single viewer direction for the image.

By calculating the error between directions of each person and directions of the image as a whole, we build a feature vector composed of four features: (1) LME: mean of the angular errors related to the light source; (2) LSE: standard deviation of the angular errors related to the light source; (3) VME: mean of the angular errors, related to the viewer; (4) VSE: standard deviation of the angular errors, related to the viewer. Using a machine learning method for modelling the behavior of fake images and performing tests using a five-fold cross validation test protocol over a database composed of 120 images (60 non-manipulated or pristine images and 60 compositions), we were able to achieve a promising classification rate ($\sim 70\%$ outperforming existing methods using the same principle.

## 3. Exposing Forgeries Based on Illuminant Maps

Despite being a very difficult region to doctoring, without leaving traces, eyes are not always visible on images involving people. Thinking about this limitation, our second proposed method focuses on identifying image splicing through cues found in the face of the people in an image. These cues are detected by using *illuminant maps* (IM) which, in a high level of abstraction, are an estimation of the light color in all image points.

The interaction between illuminant and a specific kind of material (in our work we focused on the skin material), should produce similar answers. In other words, if the image is not fake, the illuminant estimated from two faces should present similar appearance [Carvalho et al. 2013]. However, interpreting illuminant maps is not an easy task. Therefore, the main idea related to this method is to use capture the forgery telltales left behind by manipulation operations through the proper statistical characterization of the illuminant maps. For that, we employ image descriptors to capture intrinsic properties, not always visible to the naked eyes. Our method first estimates an illuminant map from the image using two different approaches: one statistical-based, named Gray-World (GGE), and one physical-based named Inverse of Intensity Chromaticity (IIC). Once illuminant maps are estimated, we characterize each face using two image descriptors, the well-known image descriptor Statistical Analysis of Structural Information (SASI) and a brand new one based on bag of words representation calculated on top of image edges. Together, the descriptors extracted from the illuminant maps are used to characterize pairs of people's faces, which are classified through a meta fusion of a Support Vector Machine (SVM) classifier.

To validate the proposed method, we used two different public datasets: DSO-1 and DSI-1 [Carvalho et al. 2013]. We used a five-fold cross validation in DSO-1 achieving an AUC of $86.3\%$ effectiveness. For testing in DSI-1, we applied a cross-dataset protocol, whereby DSO-1 was used as the training set and DSI-1 was used as the testing set, achieving an AUC of $82.6\%$. Figure 2 depicts an overview of the proposed method.

Among the main contributions of this method, we can set out: (1) interpretation of the illuminant maps as object texture for feature computation; (2) proposition of a novel edge-based characterization method for illuminant maps which explores edge attributes related to the illumination process; (3) the creation of a benchmark dataset comprising 100 skillfully created forgeries and 100 original photographs; (4) quantitative and qualitative

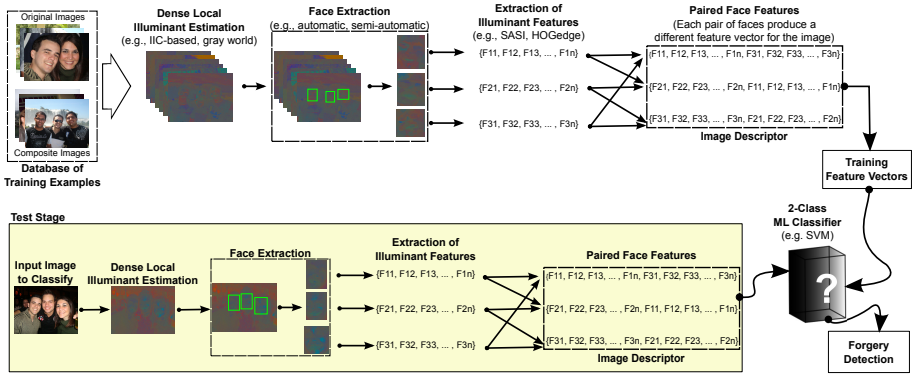



**Figure 2. First method based on illuminant maps – Overview.**

evaluations with users using the Mechanical Turk tool giving us important insights about the difficulty in detecting forgeries in digital images.

## 4. Exposing Forgeries Based On Illuminant Map Fusion

The second method proposed in this work represents a great step forward in the digital forensic field. The use of forensic methods in real life scenarios requires improvements in terms of confidence provided by this kind of methods. Inspired by this important requirement, the third method employs complementary features combined with a machine learning fusion framework to improve the fake image detection, achieving an accuracy of $94\%$ [Carvalho et al. 2015a].

Given an IM (estimated using IIC or GGE), this method converts the IM into different color spaces (RGB, YCbCr, HSV) generating IM'. Then, we characterize IM' using different image descriptors that capture properties as color, shape and texture. In total, we obtain 54 different descriptors for each image, each one focusing on one image property. Relying on a custom-tailored machine learning framework for decision-model selection, we select the best combination of descriptors, color spaces and classifiers to detect image splicing. Additionally, if an arbitrary image is classified as fake, we are also able to pinpoint the face with highest probability to be the fake face. Figure 3 depicts an overview of the proposed method.

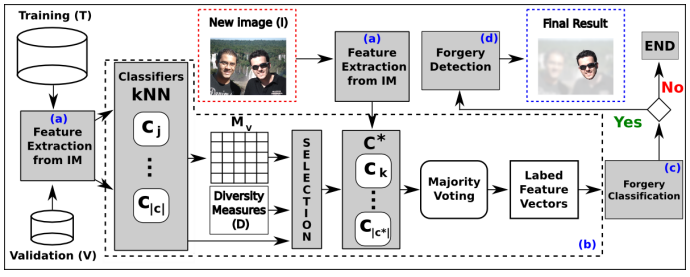



**Figure 3. Second method based on illuminant maps – Overview.**

To validate the proposed methods, we performed different experiments with distinct validation protocols including five-fold cross validation, cross-dataset validation and a qualitative analysis of questionable images downloaded from the Internet.

The most important contributions of this method are: (1) the exploration of illumination maps represented in different color spaces; (2) the incorporation of complementary

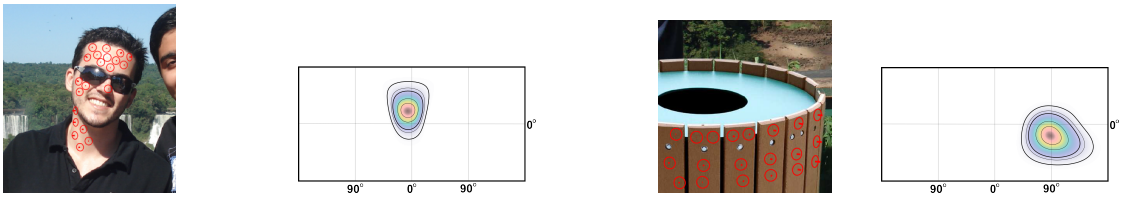
color descriptors, which showed to be very effective when characterizing IMs; (3) a full study of the effectiveness and complementarity of many different image descriptors applied to IMs to detect image illumination inconsistencies; (4) a quantitative evaluation of the differences among illuminant estimation methods, assessing the behavior of normal and fake images relative to these differences; (5) the adoption of a machine learning framework aiming at automatically selecting the best combination of all the factors of interest (e. g., IMs, color spaces, descriptors, classifiers); (6) the introduction of a new approach to detecting the most likely doctored part in fake images; (7) an improvement of 15 percentage points in the classification accuracy and the possibility of providing a confidence degree associated with the classified image when compared with our previously proposed method (Section 3).

## 5. Exposing Forgeries Based on 3D Light Source Positions

When analyzing splicing images broadcasted on the Internet, it is easy to realize that most of them involve people. However, it is possible that other types of objects can be introduced into the scene. Inspired by the necessity of producing a method able to deal with splicing images involving different kinds of objects, we proposed a method for estimating the 3D position of light source in an image [Carvalho et al. 2015b].

The user provides, through a custom-tailored interface, the normal orientations of some objects of the image. The orientation of each normal is improved by using a statistic model and then, for each object, the 3D position of the light source is estimated multiples times, generating a possible light source region. When analyzing objects that originally belong to the image, their 3D light source regions should be in a similar location with intersection. If any object presents a very different position of light source region, this image contains traces of forgery by splicing. To validate the method, we performed tests with more than 20 users in synthetic images and in different real images.

Figure 4 depicts examples of two objects and their respectively light source region. The object (a), which originally belongs to the image, presents a central light source region, depicted in (b). On the other hand, the object (c), included through splicing into the same image, depicts a very different light source region, which is incompatible with other objects in the same image.




(a) (b) (c) (d)

**Figure 4. Different objects (a, c) and their respective light source regions (b, d) extracted from a fake image (a). The light source region (d) estimated for the fake object (c) is totally different from the light source region (b) provided by the other object (a) originally present in the image.**

## 6. Conclusions and Future Research Directions

This Ph.D. thesis presented four techniques for detecting image splicing using inconsistencies in lighting. Each one of the proposed methods has its scientific contributions, as

well as its own limitations. Furthermore, during this work, we realized that digital forensic methods are in constant progress and there is no *silver bullet* able to deal with all kinds of images/scenarios. Therefore, it is important to develop and combine complementary solutions in order to better capture the image telltales left behind by a forger.

For future research directions, we suggest further improvements on detection of doctored faces in illuminant methods and the use of a better correction models for normals in 3D light source approach.

## 7. Publications, Awards and Relevant Production

1. **Best Ph.D. Thesis** (2014). In XXVII Conference on Graphics, Patterns and Images (SIBGRAPI), Rio de Janeiro, Brazil.
2. Carvalho, T., Riess, C., Angelopoulou, E., Pedrini, H., and Rocha, A. (2013). *Exposing Digital Image Forgeries by Illumination Color Classification*. IEEE T.IFS, 8(7):1182–1194.
3. Carvalho, T., Faria, F., Torres, R., Pedrini, H., and Rocha., A. (2015). *Image Composition Detection through Illuminant Map Fusion*. Second round of reviews in IEEE T.IFS.
4. Carvalho, T., Pedrini, H., and Rocha, A. (2015) *Visual Computing and Machine Learning Techniques for Digital Forensics*. RITA.
5. Saboia, P., Carvalho, T., and Rocha, A. (2011). *Eye Specular Highlights Telltales for Digital Forensics: A Machine Learning Approach*. 18th IEEE ICIP, Brussels, Belgium, pages 1937–1940.
6. Carvalho, T., Farid, H., and Kee, E. (2015). Exposing Photo Manipulation From User- Guided 3-D Lighting Analysis. In SPIE Electronic Imaging , San Francisco, CA, USA.
7. Carvalho, T., Pinto, A., Silva, E., da Costa, F., Pinheiro, G., and Rocha, A. (2012). Chapter *Crime Scene Investigation (CSI): da Ficção à Realidade*. Escola Regional de Informática de Minas Gerais, UFJF, pp. 1–23.
8. Carvalho, T., Pedrini, H., and Rocha, A. (2014). *Visual Computing and Machine Learning Techniques for Digital Forensics*. In Tutorials of XXVII SIBGRAPI, Rio de Janeiro, Brazil.
9. Carvalho, T., Pedrini, H., Rocha, A. (2014) *Illumination Inconsistency Sleuthing for Exposing Fauxtography and Uncovering Composition Telltales in Digital Images*. In CTD of SBSeg.
10. TV Interviews: 2; Interviews on Newspapers: 3; Talks: 6.

## 8. Acknowledgments

The authors thank the financial support of Unicamp, Brazilian National Research Counsel – CNPq (Grants #140916/2012-1, #477662/2013-7, #307113/2012-4, and #304352/2012-8), São Paulo Research Foundation – FAPESP, (Grant #2010/14910-0, #2010/05647-4, and #2011/22749-8), Coordination for the Improvement of Higher Education Personnel – CAPES (Grants #0214-13-2), CAPES DeepEyes Project, IF Sudeste MG, and Microsoft Research.

## References

Carvalho, T., Faria, F., Torres, R., Pedrini, H., and Rocha., A. (2015a). Image composition detection through illuminant map fusion. In second round of reviews in IEEE T.IFS.

Carvalho, T., Farid, H., and Kee, E. (2015b). Exposing Photo Manipulation From User-Guided 3-D Lighting Analysis. In *SPIE Symposium on Electronic Imaging*, San Francisco, CA, USA.

Carvalho, T., Pedrini, H., and Rocha, A. (To appear in 2015c). Visual computing and machine learning techniques for digital forensics. *Revista de Informática Teórica e Aplicada (RITA)*.

Carvalho, T., Pinto, A., Silva, E., da Costa, F., Pinheiro, G., and Rocha, A. (2012). *Escola Regional de Informática de Minas Gerais*, chapter Crime Scene Investigation (CSI): da Ficção à Realidade. UFJF.

Carvalho, T., Riess, C., Angelopoulou, E., Pedrini, H., and Rocha, A. (2013). Exposing Digital Image Forgeries by Illumination Color Classification. *IEEE T.IFS*, 8(7):1182–1194.

Saboia, P., Carvalho, T., and Rocha, A. (2011). Eye Specular Highlights Telltales for Digital Forensics: A Machine Learning Approach. In *IEEE ICIP*, pages 1937–1940.