proportional to the surface normal and the direction to a single light source. With knowledge of 3-D surface normals, the direction to the light source can be estimated. Because 3-D surface normals usually cannot be determined from a single image, only the 2-D surface normals at occluding boundaries were considered. In return, only two of the three components of the light source direction were estimated. We can estimate the 3-D direction to a light source from the light's reflection in the human eye. The required 3-D surface normals were determined by leveraging a 3-D model of the human eye. In these earlier works, a simplified lighting model consisting of a single dominant light source was assumed.

In practice, however, the lighting of a scene can be complex: any number of lights can be placed in any number of positions, creating different lighting environments. Because 3-D surface normals usually cannot be determined from a single image, we considered the 2-D surface normals at occluding boundaries, from which only five of the nine model parameters could be estimated. We can also estimate the full 3-D lighting environment in images of people. In order to extract the required 3-D surface normals, we fit 3-D models to an image of a person's head and automatically align this model to an arbitrary head pose. We describe how to model and estimate lighting environments using this approach and show its efficiency in detecting photographic composites. This 3-D approach removes the ambiguities in the earlier 2-D lighting techniques, and hence allows for a more powerful forensic analysis. For this we first express an arbitrary lighting environment as a non-negative function on the sphere, specifying the intensity of the incident light along the unit vector direction. Then we will go for 3d model estimation and registration [1].

### 2.1. 3-D Model Estimation

A 3-D morphable model for the analysis and synthesis of human faces was derived by collecting a set of 3-D laser scanned faces and projecting them into a lower-dimensional linear subspace. New faces (geometry, texture/color, and expressions) are modeled as linear combinations of the resulting linear basis. The 3-D model parameters can be estimated from a paired frontal and profile view or from only a single frontal view. This estimation requires the manual selection of several fiducial points on the face (11 on the frontal view and 9 on the profile view), from which the 3-D model is then automatically estimated [1].

### 2.2. 3-D Model Registration

Once estimated, the 3-D model is registered to the face being analyzed. This is done by maximizing an objective function over the camera intrinsic and extrinsic parameters that aligns the 3-D model to the image of the face. The 3-D model is first manually rotated to approximately align it with the image. At least three corresponding points are selected on the model and image (e.g., the center of each eye and base of the nose), from which the optimal translation is estimated using standard least-squares. With this initial alignment as a starting configuration, a brute force search is performed over the three rotation parameters, focal length, and camera center. On each iteration of this search, the translation vector is estimated. In order to reduce the effects of lighting, a high-pass filter is applied to both the image and rendered model. Once the model has been estimated and registered, 3-D surface normals and corresponding intensities are used to estimate the lighting environment.

We can accurately estimate the 3-D lighting environment even at low resolution. But it would have been difficult to estimate 2-D normals in low-resolution image. When creating a composite of two or more people, it is often difficult to match the lighting. Lighting environments can be approximated with a nine dimensional model consisting of a linear combination of spherical harmonics. Lighting inconsistencies across an image are then used as evidence of tampering [1]. This technique extends earlier 2-D lighting approaches that, due to the lack of 3-D surface normals, were only able to estimate a subset of the full lighting model. This new approach removes the ambiguities in these earlier techniques, and hence allows for a more powerful forensic analysis.

## 3. Exposing Digital Forgeries By Detecting Duplicated Image Regions

A common manipulation in tampering with an image is to copy and paste portions of the image to conceal a person or object in the scene. If the splicing is imperceptible, little concern is typically given to the fact that identical or virtually identical regions are present in the image. This is a technique that can efficiently detect and localize duplicated regions in an image by applying a principal component analysis (PCA) on small fixed size image blocks to yield a reduced dimension representation. This representation is robust to minor variations in the image due to additive noise or lossy compression. Duplicated regions are then detected by lexicographically sorting all of the image blocks. The data-driven PCA basis may better capture discriminating features. We show the efficacy of this technique on credible forgeries, and quantify its robustness and sensitivity to additive noise and lossy JPEG compression.

### 3.1 Detecting Duplicated Regions

Given an image with N pixels, our task is to determine if it contains duplicated regions of unknown location and shape. An exhaustive approach that would examine every possible pair of regions would have an exponential complexity in the number of image pixels. Such an approach is obviously computationally prohibitive. A more efficient algorithm might look for duplication of small fixed-sized blocks. By stringing each such block into a vector and lexicographically sorting all image blocks, identical blocks correspond to adjacent pairs in the sorted list. The primary cost of this algorithm would be the lexicographic sorting, yielding a complexity of $O(N \log N)$, since the number of image blocks is proportional to the number of image pixels, N. Note that this is a significant improvement over the brute-force exponential algorithm. The drawback of this approach, however, is that it is sensitive to small variations between duplicated regions due to, for example, additive noise or lossy compression [2]. We describe next an algorithm that overcomes this limitation while retaining its efficiency.

The detection algorithm proceeds as follows. First, to further reduce minor variations due to corrupting noise, the reduced dimension of each image block, is component-wise quantized. A matrix is constructed whose rows contain these quantized coefficients. Let the matrix be the result of lexicographically sorting the rows of this matrix in column order. Let $\vec{s}_i$ denote the $i^{th}$ row of this sorted matrix, and let the tuple $(x_i, y_i)$ denote the block's image coordinates (top-left corner) that corresponds to $\vec{s}_i$. Consider next all pairs of rows $\vec{s}_i$ and $\vec{s}_j$, whose row distance, $|i - j|$, in the sorted matrix S is less than a specified threshold. The offset, in the image, of all such pairs is given by:

$$(x_i - x_j ,\ y_i - y_j) \text{ if } x_i - x_j > 0$$
$$(x_j - x_i,\ y_i - y_j) \text{ if } x_i - x_j < 0$$
$$(0,\ |y_i - y_j|) \text{ if } x_i = x_j$$

From a list of all such offsets, duplicated regions in the image are detected by noting the offsets with high occurrence [2]. For example a large duplicated region will consist of many smaller blocks, each of these blocks will appear in close proximity to each other in the lexicographically sorted matrix, and will have the same offset. In order to avoid false hits due to uniform intensity areas, offset magnitudes below a specified threshold are ignored. The results of this detection can be visualized by constructing a duplication map- a zero image of the same size as the original is created, and all pixels in a region believed to be duplicated are assigned a unique gray scale value.

The complexity of this algorithm is dominated by the lexicographic sorting. There are at least two ways in which this algorithm can be extended to color images. The simplest approach is to independently process each color channel (e.g., RGB) to yield three duplication maps. The second approach is to apply PCA to color blocks of size 3b, and proceed in the same way as described above.

We have presented an efficient and robust technique that automatically detects duplicated regions in an image. This technique works by first applying a principal component analysis (PCA) on small fixed-size image blocks to yield a reduced dimension representation that is robust to minor variations in the image due to additive noise or lossy compression. Duplicated regions are then detected by lexicographically sorting all of the image blocks. The technique is effective on plausible forgeries, and has quantified its sensitivity to JPEG lossy compression and additive noise. Detection is possible even in the presence of significant amounts of corrupting noise.

## 4. Exposing Digital Forgeries Through Specular Highlights On The Eye

When creating a digital composite of two people, it is difficult to exactly match the lighting conditions under which each individual was originally photographed. In many situations, the light source in a scene gives rise to a specular highlight on the eyes. We show how the direction to a light source can be estimated from this highlight. Inconsistencies in lighting across an image are then used to reveal traces of digital tampering. In this work, we show how the location of a specular highlight can be used to determine the direction to the light source. Inconsistencies in the estimates from different eyes, as well as differences in the shape and color of the highlights, can be used to reveal traces of digital tampering.

### 4.1 Methods

The position of a specular highlight is determined by the relative positions of the light source, the reflective surface and the viewer (or camera). The light direction can be estimated from the surface normal and view direction at a specular highlight. In the following sections, we describe how to estimate these two 3-D vectors from a single image. Note that the light direction is specified with respect to the eye, and not the camera. In practice, all of these vectors will be placed in a common coordinate system, allowing us to compare light directions across the image.

#### a) Camera Calibration

In order to estimate the surface normal and view direction in a common coordinate system, we need to estimate the projective transform that describes the transformation from world to image coordinates. With only a single image, this calibration is generally an under-constrained problem. In our case, however, the known geometry of the eye can be exploited to estimate this required transform. The limbus, the boundary between the sclera (white part of the eye) and the iris (colored part of the eye), can be well modeled as a circle. The image of the limbus, however, will be an ellipse except when the eye is directly facing the camera. Intuitively, the distortion of the ellipse away from a circle will be related to the pose and position of the eye relative to the camera [3]. We therefore seek the transform that aligns the image of the limbus to a circle.

Once estimated, the projective transform can be decomposed in terms of intrinsic and extrinsic camera parameters. The intrinsic parameters consist of the camera focal length, camera center, skew and aspect ratio. The extrinsic parameters consist of a rotation matrix and a translation vector that define the transformation between the world and camera coordinate systems [3]. Since the world points lie on a single plane, the projective transform can be decomposed in terms of the intrinsic and extrinsic parameters.

#### b) Surface Normal

The 3-D surface normal N at a specular highlight is estimated from a 3-D model of the human eye. The model consists of a pair of spheres. The larger sphere represents the sclera and the smaller sphere represents the cornea. The surface normal depends on the view direction [3]. The surface normal is determined by intersecting the ray leaving specular highlight along the view direction with the edge of the sphere.

**c) Light Direction**

The position of the specular highlight is then used to determine the surface normal. Combined with the estimate of the view direction V, the light source direction L can be estimated. In order to compare light source estimates in the image, the light source estimate is converted to camera coordinates [3].

When creating a composite of two or more people it is often difficult to match the lighting conditions under which each person was originally photographed. Specular highlights that appear on the eye are a powerful cue as to the shape, color and location of the light source(s). Inconsistencies in these properties of the light can be used as evidence of tampering. We can also measure the 3-D direction to a light source from the position of the highlight on the eye. The shape and color of a highlight are relatively easy to quantify and measure and should also prove helpful in exposing digital forgeries. Since specular highlights tend to be relatively small on the eye, it is possible to manipulate them to conceal traces of tampering. To do so, the shape, color and location of the highlight would have to be constructed so as to be globally consistent with the lighting in other parts of the image. Also working in our favor is that even small artifacts on the eyes are visually salient. Nevertheless, as with all forensic tools, it is still possible to circumvent this technique.

## 5. Exposing Digital Image Forgeries Using Color Illumination

Methods based on illumination inconsistencies have two main characteristics that make them potentially effective in splicing detection. Firstly, from the viewpoint of a manipulator, a perfect adjustment of illumination conditions is very difficult to achieve when creating a composite photograph. Secondly, this class of methods can also be used to analyze analog pictures [9]. Illumination color analysis is a promising cue to expose image composites. In earlier work, Riess and Angelopoulou proposed to analyze illuminant color estimates from local image regions to detect spliced images [6]. Unfortunately, the authors leave the interpretation of the so-called illuminant maps to human experts. In practice, it turns out that it is very challenging to decide whether or not an image is tampered with based just on illuminant maps. Moreover, we cannot simply rely solely on a subject's or expert's opinion, as the human visual system can be quite inept at judging inconsistencies in photographs, especially when it involves lighting and shadows [4].

We propose a new semi-automatic method that is considerably easier to use and more reliable than earlier approaches. We make use of the fact that local illuminant estimates are most discriminative when comparing objects that are made of the same (or similar) material. Thus, we focus on the automated comparison of regions of human skin, and more specifically, faces [4]. We classify the illumination on two faces as either consistent or inconsistent. The only interaction that is required by the user is to select image regions that contain objects of similar materials. Specifically, we restrict the required user interaction to marking bounding boxes around the faces in an image under investigation.

### 5.1 Core of the System: Estimation of the Locally Dominant Illuminant and Interpretation of Illuminant Maps

Here we propose a new approach that minimizes user dependence and improves the state-of-the-art. We classify the illumination for each pair of faces in the image as either consistent or inconsistent. First, we present an overview of the algorithm. Then, we present the algorithmic details for every step. The proposed method consists of five main components:

- **Dense Local Illuminant Estimation (IE):** The input image is segmented into homogeneous regions. Per illuminant estimator, a new image is created where each region is colored with the estimated illuminant color. This resulting intermediate representation is called illuminant map (IM).
- **Face Extraction:** This is the only step that may require human interaction. An operator sets a bounding box around each face (e. g., by clicking on two corners of the bounding box) in the image that should be investigated. Alternatively, an automated face detector can be employed. We then crop every bounding box out of each illuminant map, such that only the illuminant estimates of the face regions remain.
- **Computation of Illuminant Features:** For all face regions, texture-based and gradient based features are computed on the IM values. Further analysis is performed on these features.
- **Paired Face Features:** Our goal is to assess whether two faces in an image are consistently illuminated. For that, we combine the feature vectors from each pair of faces in the image creating a pair-of-faces feature vector.
- **Classification:** We use a machine learning approach to automatically classify the feature vectors. Given an image with f faces, we consider an image as a forgery if at least one pair of faces (represented by one feature vector) is classified as inconsistently illuminated [5].

#### 5.1.1 Dense Local Illuminant Estimation

To detect inconsistencies in the illumination color, we need a dense set of localized estimates. We segment the input image into regions of approximately constant chromaticity (so-called super-pixels). Then we estimate the color of the illuminant per super-pixel. By recoloring the super-pixels with the estimated illuminant chromaticity, we obtain an illuminant map. We use two separate methods to obtain a version of this map:

**a) Generalized Gray World Estimates**

We follow the generalized gray world approach by van de Weijer et al [10]. Let $f = (R, G, B)^T$ denote the observed color of a pixel. Van de Weijer et al. assume a Lambertian scene (i. e., objects of purely diffuse reflectance) and linear camera response. Then, f is formed by

$$f = \int_{\Omega} e(\lambda)s(\lambda)c(\lambda)\,d\lambda$$

where λ denotes the wavelength of the light, e(λ) denotes the spectrum of the illuminant, s(λ) the surface reflectance of an object, and c(λ) the sensitivity of the camera as a vector for each color channel.

**b) Inverse Intensity-Chromaticity Estimates**

The second illuminant estimator is the inverse intensity-chromaticity (IIC) color space. The observed image intensities are assumed to exhibit a mixture of diffuse (i. e., Lambertian) and specular reflectance. Pure specularities are assumed to consist of only the color of the illuminant. Let f = $(R, G, B)^T$ be the observed colors of a pixel. Then, using the same notation as for the generalized gray world model, f is modeled as

$$f = \int_{\Omega} (e(\lambda)s(\lambda) + e(\lambda))c(\lambda)d\lambda$$

### 5.1.2 Face Extraction

Unconstrained estimation of the illuminant color can be error-prone and affected by the reflectance properties of the materials in the scene. However, it is possible to improve the accuracy of the relative error between two estimates by focusing only on objects of approximately the same material. For this work, we limit our examination of illumination consistency to human skin and, in particular, to faces. Pigmentation is the most obvious difference in skin characteristics between different ethnicities. This pigmentation difference depends on many factors as quantity of melanin, amount of UV exposure, genetics, melanosome content and type of pigments found in the skin. However, this intramaterial variation is typically smaller than that of all materials possibly occurring in a scene.

All faces in the image that should be part of the investigation have to be annotated with a bounding box. In principle, this can be done automatically, through the use of a face detector. However, we prefer a human operator for this task for two main reasons: a) this minimizes false detections or misses of faces; b) scene context is important when judging the lighting situation. For instance, consider an image where all persons of interest are illuminated by ashlight. The illuminants are expected to agree with one another. Conversely, assume that a person in the foreground is illuminated by ashlight, and a person in the background is illuminated by ambient light. Then, a difference in the color of the illuminants is expected. Such differences are hard to distinguish in a fully automated manner.

### 5.1.3 Interpreting Illuminant Maps as Texture Maps

From an image processing perspective, we can interpret the illuminant maps from face regions as texture maps. Many different texture descriptors have been proposed in the literature thus far. One of the most effective methods is the Statistical Analysis of Structural Information (SASI) descriptor [11]. The most important advantage of SASI application is its remarkable capability of capturing small granularities and discontinuities which are present in texture patterns. These patterns appear mainly in sharp corners and abrupt changes such as the ones present in illuminant maps, especially in the face region of composite images.

### 5.1.4 Face Pair

To compare two faces, we combine the same descriptors for each of the two faces. For instance, we can concatenate the SASI-descriptors that were computed on gray world. The idea is that a feature concatenation from two faces is different when one of the faces is an original and one is spliced. For an image containing f faces ($f \geq 2$), the number of face pairs is (f(f -1))=2.

### 5.1.5 Classification

We classify the illumination for each pair of faces in an image as either consistent or inconsistent. Assuming all selected faces are illuminated by the same light source, we tag an image as manipulated if one pair is classified as inconsistent. Individual feature vectors, i. e., texture on either gray world or IIC-based illuminant maps, are classified using a Support Vector Machine (SVM) classifier with a Radial Basis Function (RBF) kernel.

In this work, we presented a new method for detecting forged images of people using the color of the illuminant. We estimate the illuminant color with a statistical gray edge method and a physics-based method using the inverse intensity-chromaticity color space. We interpret these illuminant maps as texture maps and also extract edge information from them. Although the proposed method is tailored to detect splicing on images containing faces, there is no principal hindrance in applying it to other, problem-specific materials in the scene. The proposed method requires only a minimum of human interaction and provides a crisp statement on the authenticity of the image. Additionally, it is an important leap ahead to exploit color as a forensic cue. Prior color-based work either assumes complex user interaction or imposes very limiting assumptions. Although promising as forensic cues, methods that operate on illuminant color are inherently prone to estimation errors.

## 6. Conclusion

The authenticity of an image is major research challenge in the field image forensic for real world events. The image integrity verification as well as identifying the areas of tampering on images without need to any expert support or manual process or prior knowledge original image contents is now days becoming the challenging research problem. Thus to solve this problem recently some techniques were presented and new techniques will be developed to make better and harder to detect fakes (for exposing photographic frauds). In this paper we have discussed different methods of detection for digital image forgery. For the future work we suggest to work over improved new method with efficient skin detection methods.

An attempt has been made to introduce various promising techniques that represent reasonable improvements in the forgery detection methods. Still these improvements are far from being perfect and have certain drawbacks that must be eliminated to obtain effective results. There are techniques exhibiting improved detection accuracy, but having high computational complexity. Moreover, most of the methods may not be that responsive to the geometric transformations, such as rotation and scaling. The factor of human perception is also not counted as a factor during the development of these techniques. Therefore there is a need to develop techniques that are automatic, and effective against geometric transformations.

## References

[1] E. Kee and H. Farid, “Exposing digital forgeries from 3-D lighting environments,” in Proc. IEEE Int. Workshop on Inform. Forensics and Security, Dec. 2010, pp. 1–6.
[2] Alin C Popescu and Hany Farid, “Exposing Digital Forgeries by Detecting Duplicated Image Regions,” Dept. Comput. Sci., Dartmouth College, Tech. Rep. TR2004-515, 2004.
[3] J. F. O’Brien and H. Farid, “Exposing photo manipulation with inconsistent reflections,” ACM Trans. Graphics, vol. 31, no. 1, pp. 1–11, Jan. 2012.
[4] M. Johnson and H. Farid, “Exposing digital forgeries through specular highlights on the eye,” in Proc. Int.Workshop on Inform. Hiding, 2007, pp.311–325.
[5] Tiago José de Carvalho, Christian Riess, Elli Angelopoulou, Hélio Pedrini and Anderson de Rezende Rocha, “Exposing Digital Image Forgeries byIllumination Color Classification”, IEEE Transactions On Information Forensics And Security, Vol. 8, No. 7, July 2013.
[6] C. Riess and E.Angelopoulou, “Scene illumination as an indicator of image manipulation,” Inf. Hiding, vol. 6387, pp. 66–80, 2010.
[7] Tanzeela Qazi, Khizar Hayat, Samee U. Khan, Sajjad A. Madani, Imran A. Khan, Joanna Kołodziej, Hongxiang Li, Weiyao Lin, Kin Choong Yow, Cheng-Zhong Xu, “Survey on blind image forgery detection” , The Institution of Engineering and Technology 2013.
[8] H. Farid and M.J. Bravo, “Image forensic analyses that elude the human visual system,” in SPIE Symposium on Electronic Imaging, San Jose,CA, 2010.
[9] A. Rocha, W. Scheirer, T. E. Boult, and S. Goldenstein, “Vision of the unseen: Current trends and challenges in digital image and video forensics,” ACM Comput. Surveys, vol. 43, pp. 1–42, 2011.
[10] J. van de Weijer, T. Gevers, and A. Gijsenij, “Edge-based color constancy,” IEEE Trans. Image Process., vol. 16, no. 9, pp. 2207–2214, Sep. 2007.
[11] A. Carkacioglu and F. T. Yarman-Vural, “Sasi: A generic texture descriptor for image retrieval,” Pattern Recognit., vol. 36, no. 11, pp.2615–2633, 2003.

## Author Profile

**Hemin Jose** received the BE degree in Electronics and communication Engineering from Anna University, Coimbatore in 2011. She is currently pursuing M.Tech Degree in Communication Engineering and Signal processing from Kannur University. Her research area is Digital Image Processing.

**Lekshmy S** received the B.tech degree in Electronics and communication from Kerala University in the year 2007 and received the ME in Applied Electronics from Anna University in the year 2010. Then she worked as Assistant Professor in Sree Narayana Guru College of Engineering, Kannur for the period of January 2010 to May 2011. Currently she is working as Assistant Professor in the Department of Electronics And Communication Engineering at Vimal Jyothi Engineering College, Chemperi, Kerala, India. Her area of interest is in Digital Image Processing.