Illumination Color Classification Based Image Forgery Detection: A Review

Nitish L. Nirmalkar
Department of Computer Technology,
Yeshwantrao Chavan College of Engg.
Nagpur, India
nitishnirmalkar@gmail.com

Shailesh D. Kamble
Department of Computer Technology,
Yeshwantrao Chavan College of Engg.
Nagpur, India
shailesh_2kin@rediffmail.com

***Abstract:*-Over the years, photographs have been used as document of events as well as proof in legal proceedings. Although some of photographers are able to manipulate the images and make composites of original images, this process requires expert knowledge and demands lot of time. In today's world, a lot of skillfully developed image editing softwares are available which makes image manipulation and modifications an easy task. This reduces trust in photographs and presents an open question on authenticity of images and photographs. This paper, aims at focusing on the most common forms of image manipulation, which is called as image splicing. Here we propose a forgery detection method which uses inconsistencies in illumination of images. For achieving this, we make use of physics and statistics based illuminant estimators. These illuminant estimates are used to extract features which are then given to a machine-learning technique which supports us in decision-making.**

Keywords: Illuminant estimates ,machine learning.

I. INTRODUCTION

A large number of digital and analog images are captured and produced by a wider range of electronic devices and gadgets such as cameras handheld multimedia devices, computers. Image and photographs are distributed across the globe via various means of communication. World Wide Web plays a major role in establishing communication and distributing the documents among the masses. Nowadays, it is very easy to use image processing techniques to manipulate images, image editing software makes the modification and manipulation an easier job. However, before considering an image inappropriate or taking actions on any image, one must confirm whether the image has been altered or is a composite image [2]. Image splicing is image manipulation technique which is a commonly used. To determine whether the image is authentic or not, forensic investigators use a number of sources for assessment of alteration in an image.

II. LITERATURE REVIEW

Illumination-based techniques in broadly classified fall in two categories .The methods based on color examine for inconsistencies in the relationship between object color and light color. The techniques based on Geometry-based methods aims at identifying inconsistencies in positions among particular objects in the image and light source used.

Some of the techniques that are suggested for image forgery detection makes use of the direction of the light source from which the light is incident on objects in an image . Johnson and Farid [6] suggested a technique which estimates a low-dimensional descriptor of the lighting environment within the image in 2-D. It estimates the illumination direction from the intensity distribution on object boundaries of similar color. Kee and Farid [8] extended this approach by exploiting known 3-D surface within the case of faces, a mesh of 3-D normal increase the estimate of the illumination direction, to achieve this, a 3-D face model is registered with the 2-D image facial landmarks. Fan et al.[9]proposed a way for estimating 3-D illumination which makes use of shape-from-shading methods. This technique doesn't requires 3-D model of the object, which also reduces dependableness of the method.

Johnson and Farid [7] suggested by utilizing specular highlights in the eyes the splicing can be detected. Saboia et. al.[12],extended this by classifying the images by exploring some additional extra options. The pertinence of each of these approaches is limited by the very fact that images should be available in high resolution and eye should be clearly captured while taking the photographs.

Gijsenij et.al. [15] suggested a pixel wise illuminant estimator. This estimator lets the segmentation of image into regions illuminated by various illuminants. These illuminated regions which are distinct have sharp transitions from different illuminant shades. The drawback of this method is that, a single estimator may not be always accurate and there is chance it may not always succeed in estimating the illumination properly. A new technique is needed that collaborates and combines the outcome from multiple illuminant estimators. Kawakami[33] suggested another approach which is based on physics a that is made for specific images for discriminating illuminants from differently illuminated regions.

III. METHODOLOGY

The method [1] is broadly classified into five components:

A. Dense Local Illuminant Estimation

The input image is segmented into a specific type of regions, which are the regions of similar constant hue, by using algorithm suggested by by Felzenszwalb and Huttenlocher [17] . For each such regions mentioned above, illuminant color of the regions is calculated. The generalized grey world estimates is one color estimator and an estimate based on physics which is called as inverse-intensity chromaticity space estimator are used for obtaining the illumination estimates. The two illuminant maps are obtained by recoloring the regions with similar hue. The both illuminant maps obtained are then examined individually afterwards.

1) Generalized Gray World Estimates: The gray world is suggested by Buchsbaum [18] which is based on the assumption that the mean color of an image or picture is gray. If there is variation or deflection of the mean of the image intensities from the presumed gray color results because of the illuminance .Van de Weijer [16] extended this scheme of gray world hypothesis by inclusion of three parameters which are Derivative order (n): the assumption that the mean of the illuminants is gray can be extended to the absolute value of the sum of the derivatives of the image. Minkowski norm (p): Rather than summing-up image intensities, better results can be achieved by computing the p-th Minkowski norm .In Gaussian smoothing preprocessing the image is don so as to remove noise and smoothen the image.

2) Inverse Intensity-Chromaticity Estimates [20]: In this approach, the image intensities which obtained are assumed to show a combination of diffuse and specular reflectance. Diffuse reflection is the reflection of light from a surface such that an incident ray is reflected at many angles .Specular reflection is the mirror-like reflection of light , in which light ray coming from solitary direction is reflected into solitary outgoing direction[34]. Crisp specularities are presumed to be composed of only the color of the illuminant. Riess and Angelopoulou [3] suggested a technique to compute these estimates in an image in small regions. In the voting procedure, two conditions are enforced on a small region to increase flexibility with induced noises. If a small region does not fulfill the conditions stated above, then such small regions are decategorized from voting procedure. The final estimate of the illuminant is based on majority votes of the obtained estimates.

B. Face Extraction

It is needed that all the faces in an image or photograph that are needed for consideration should be enclosed with the bounding boxes .The bounding boxes can be procured by using an automated algorithm by Schwartz et. al.[23] or a human operator is chosen for bounding boxes. The advantages using a human operator for bounding boxes are that a human is better at judging the area of face to be enclosed in bouding boxes, reduces missed faces as well as false detection of faces another reason for preferring human over automated algorithm is the scene context is crucial in determining the lighting situation. Given a scenario there is an image where all persons under consideration are illuminated by the single source of light. The illuminants are expected to be similar. On the contrary, assuming that a person in the foreground is illuminated by camera flashlight and a person in the background is illuminated by surroundings light. There will be dissimilarity in the color of the illuminants is presumed. This small dissimilarity are difficult to figure out in a fully-automated technique, but this can be eliminated with the human intervention.

C. Computation of Illuminant Features

1. Texture Description

The SASI Statistical Analysis of Structural Information descriptor suggested by Carkacioglu and Yarman-Vural [24] which is used for obtaining information regarding the image from illuminant maps.Penatti [25] suggested that this descriptor gives good performance. The salient aspect of using this descriptor is its ability to detect discontinuities in patterns of texture and capture minute granular information. Dissimilar illuminant colors interact varying with the surfaces differing in types, which gives dissimilar illumination. This descriptor estimates structural attributes of textures, which is uses the information from autocorrelation of horizontal, vertical and diagonal pixel lines at different scales. Another autocorrelation is calculated on the basis of specific fixed orientation, scale, and shift. The mean and standard deviation of all pixel values gives two feature aspects. This process is repeated for calculating different orientations, scales and shifts and applied till it gives a 128-dimensional feature vector. In the last step the obtained feature vector normalization is done.

2. *Interpretation of Illuminant Edges*

a) Extraction of Edge Points: The Canny edge detector[26] used for obtaining edge points from the illuminant map of the face region, which gives edge points that are close in distance to each others. The number of edge point obtained may be large for minimizing the number of edge points; an approach is applied on the output of the Canny edge detector. A starting point is selected as a seed point, all other edge pixels in a region of interest which are centered around the seed point are then discarded from the process. The edge points nearer to region of interest are elected as seed points for the next step. This operation is repeated on the whole image till number of points are minimized and it is assured that every face region has similar or approximate density of points.

b)Point Description: The Histograms of Oriented Gradients[34] is computed which gives description about the distribution of the chosen edge points. This Histograms of Oriented Gradients uses the concept of normalized local histograms as a basis for obtaining image gradient orientations. The HOG descriptor is built neighboring every edge points. The area around of this type of an edge point is called as cell. Each cell gives localized 1-D histogram of quantized gradient directions. The feature vector for this obtained feature descriptor is created by combining the histograms of all cells which are distributed in spatially larger region, of which histogram contrast is normalized. The output of this Histograms of Oriented Gradients is used as feature vector for the next steps.

c) Visual Vocabulary:

The number HOG vectors that are obtained differs, which is contingent upon the size and structure of the face which is being analyzed. The visual dictionaries as suggested by Csurka et. al.[28] are created to get feature vectors of specified length. This Visual dictionaries form a powerful representation, in which every face is assumed as a group of region descriptors. The spatial location of each of this region is eliminated [29]. The visual dictionary is built by subdividing the data given for training into feature vectors of original and altered images. Every group is then clustered into clusters by applying the k-means algorithm [30].After applying k-means a visual dictionary is built using the visual words, where every word is characterized by the center of the cluster. By this, the most representative feature vectors of the training set are epitomized by the visual dictionary.

d)Quantization Using the Precomputed Visual Dictionary:

The HOG feature vectors obtained are then mapped to the visual dictionary. Every feature vector in an image is characterized by the nearest word in the dictionary which is the distance of Euclidean norm. A histogram is obtained for the word counts in the dictionary which in turn represents the distribution of HOG feature vectors in a face of the image.

D.Paired Faces:

The same descriptors for each of the two faces are merged for analyzing and comparing the two faces. The SASI-descriptors obtained on gray world can be coupled together for similarity .The scheme behind this is a feature concatenation from two faces is distinct when one of the faces is an original and another is doctored one[1]. The Statistical Analysis of Structural Information descriptor and Histograms of Oriented Gradients edge descriptors seize two varying attributes of the face regions. The both descriptors can be imagined are digital signatures with distinct operations in digital signal processing context. The average value and standard deviation for each feature dimension is calculated. The feature dimensions with the max difference in the average values for are considered.

E.Classification:

The illumination for every pair of faces in an photograph or image can be categorized as consistent or inconsistent. By supposing that all chosen faces are illuminated by the same lighting conditions, the image labeled as doctored if a pair is categorized as inconsistent, the discrete feature vectors are categorized by utilizing support vector machine classifier which makes use of radial basis function kernel. The SASI feature and HOG edge features impart information that are supportive of each other .A technique based on machine learning is used for improving the detection performance. Ludwig et al. [21] suggested a technique that can be used which a late fusion technique is named SVM-Meta Fusion. Every combination of illuminant map and feature type are separately classified by making use of 2-class SVM classifier to for getting the distance between the image's feature vectors and the classifier decision boundary. The marginal distances provided by all individual classifiers to build a new feature vectors are combined by SVM-Meta Fusion.

IV. CONCLUSION

In this paper, we presented review of methods for detecting altered or doctored images of people using the illuminant color classification mechanism. The illuminant color is estimated by using a statistical based gray edge method and a method which is based on physics which uses the inverse intensity-chromaticity color space. The illuminant maps obtained can be supposed as maps of texture .Based the distribution of edges on these maps information can be obtained. The two descriptors considered were HOG descriptors and SASI descriptors. The method reviewed above needs minimum human intervention only for bounding the boxes around faces. It provides good idea about authenticity of the image and also eliminates dependency on human visual system for analyzing and detecting the forged images.

REFERENCES

[1] Tiago José de Carvalho, Christian Riess,Elli Angelopoulou, , Hélio Pedrini, , and Anderson de Rezende Rocha," Exposing Digital Image Forgeries by Illumination Color Classification", IEEE Trans. Info. Forensics and Security,vol. 8,July 2013

[2] 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.

[3] C. Riess and E. Angelopoulou, "Scene illumination as an indicator of image manipulation," Inf. Hiding, vol. 6387, pp. 66–80, 2010.

[4] Y. Ostrovsky, P. Cavanagh, and P. Sinha, "Perceiving illumination inconsistencies in scenes," Perception, vol. 34, no. 11, pp. 1301–1314, 2005.

[5] M. Johnson and H. Farid, "Exposing digital forgeries by detecting inconsistencies in lighting," in Proc. ACM Workshop on Multimedia and Security, New York, NY, USA, 2005, pp. 1–10.

[6] M. Johnson and H. Farid, "Exposing digital forgeries in complex lighting environments," IEEE Trans. Inf. Forensics Security, vol. 3, no. 2, pp. 450–461, Jun. 2007.

[7] 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.

[8] E. Kee and H. Farid, "Exposing digital forgeries from 3-D lighting environments," in Proc. IEEE Int. Workshop on Inform. Forensics and Security (WIFS), Dec. 2010, pp. 1–6.

[9] W. Fan, K. Wang, F. Cayre, and Z. Xiong, "3D lighting-based image forgery detection using shape-from-shading," in Proc. Eur. Signal Processing Conf. (EUSIPCO), Aug. 2012, pp. 1777–1781.

[10] 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.

[11] X.Wu and Z. Fang, "Image splicing detection using illuminant color inconsistency," in Proc. IEEE Int. Conf. Multimedia Inform. Networking and Security, Nov. 2011, pp. 600–603.

[12] P. Saboia, T. Carvalho, and A. Rocha, "Eye specular highlights telltales for digital forensics: A machine learning approach," in Proc. IEEE Int. Conf. Image Processing (ICIP), 2011, pp. 1937–1940.

[13] C. Riess and E. Angelopoulou, "Physics-based illuminant color estimation as an image semantics clue," in Proc. IEEE Int. Conf. Image Processing, Nov. 2009, pp. 689–692.

[14] A. Gijsenij, T. Gevers, and J. van de Weijer, "Computational color constancy: Survey and experiments," IEEE Trans. Image Process., vol. 20, no. 9, pp. 2475–2489, Sep. 2011.

[15] A. Gijsenij, R. Lu, and T. Gevers, “Color constancy for multiple light sources,” IEEE Trans. Image Process., vol. 21, no. 2, pp. 697–707, Feb. 2012.

[16] 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.

[17] P. F. Felzenszwalb and D. P. Huttenlocher, “Efficient graph-based image segmentation,” Int. J. Comput. Vis., vol. 59, no. 2, pp. 167–181, 2004.

[18] G. Buchsbaum, “A spatial processor model for color perception,” J. Franklin Inst., vol. 310, no. 1, pp. 1–26, Jul. 1980.

[19] A. Gijsenij, T. Gevers, and J. van de Weijer, “Improving color constancy by photometric edge weighting,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 34, no. 5, pp. 918–929, May 2012.

[20] R. Tan, K. Nishino, and K. Ikeuchi, “Color constancy through inverse- intensity chromaticity space,” J. Opt. Soc. Amer. A, vol. 21, pp. 321–334, 2004.

[21] O. Ludwig, D. Delgado, V. Goncalves, and U. Nunes, “Trainable classifier-fusion schemes: An application to pedestrian detection,” in Proc. IEEE Int. Conf. Intell. Transportation Syst., 2009, pp. 1–6.

[22] C. Riess, E. Eibenberger, and E. Angelopoulou, “Illuminant color estimation for real-world mixed-illuminant scenes,” in Proc. IEEE Color and Photometry in Comput. Vision Workshop, Barcelona, Spain, Nov. 2011.

[23] W. R. Schwartz, A. Kembhavi, D. Harwood, and L. S. Davis, “Human detection using partial least squares analysis,” in Proc. IEEE Int. Conf. Comput. Vision (ICCV), 2009, pp. 24–31.

[24] 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.

[25] O. A. B. Penatti, E. Valle, and R. S. Torres, “Comparative study of global color and texture descriptors for web image retrieval,” J. Visual Commun. Image Representat., vol. 23, no. 2, pp. 359–380, 2012.

[26] J. Canny, “A computational approach to edge detection,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 8, no. 6, pp. 679–698, Jun. 1986.

[27] N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection,” in Proc. IEEE Conf. Comput. Vision and Pattern Recognition, 2005, pp. 886–893.

[28] G. Csurka, C. R. Dance, L. Fan, J. Willamowski, and C. Bray, “Visual categorization with bags of keypoints,” in Proc. Workshop on Statistical Learning in Comput. Vision, 2004, pp. 1–8.

[29] J.Winn, A. Criminisi, and T. Minka, “Object categorization by learned universal visual dictionary,” in Proc. IEEE Int. Conf. Comput. Vision (ICCV), 2005, pp. 1800–1807.

[30] C. M. Bishop, Pattern Recognition and Machine Learning (Information Science and Statistics). Secaucus, NJ, USA: Springer-Verlag New York, Inc, 2006.

[31] M. Kirchner, “Linear row and column predictors for the analysis of resized images,” in Proc. ACM SIGMM Multimedia Security Workshop, Sep. 2010, pp. 13–18.

[32] A. Rocha, T. Carvalho, H. Jelinek, S. K. Goldenstein, and J. Wainer, “Points of interest and visual dictionaries for automatic retinal lesion detection,” IEEE Trans. Biomed. Eng., vol. 59, no. 8, pp. 2244–2253,Aug. 2012.

[33] R. Kawakami, K. Ikeuchi, and R. T. Tan, “Consistent surface color for texturing large objects in outdoor scenes,” in Proc. IEEE Int. Conf. Comput. Vision, 2005, pp. 1200–1207.

[34] http://en.wikipedia.org

.