Real Time Source Camera Identification using Advanced Blind Source Separation Method

Sneha K A
M Tech scholar
SCET Kerala
sneha1992anilkumar@gmail.com

Delna K V
M Tech scholar
SCET Kerala
delnakv@gmail.com

Aneesh R P
Department of ECE
Regional Center IHRD, Trivandrum
aneeshprakkulam@gmail.com

Abstract—In today's digitalizing world, multimedia devices and software have an important role in our lives. Particularly multi-media devices like digital camera and smart phones. Since people use them for illegal purpoe, a safety measure is essential against them. An efficient digital camera identification algorithm based on Sensor Pattern Noise (SPN) is proposed in this paper. Sensor pattern noise is a unique noise found in digital cameras. The huge information that a sensor pattern noise holds in terms of the frequency makes it unique, because of this uniqueness the SPN is suitable for source camera identification .ICA(Independent Component Analysis) algorithm is used for separating the sensor pattern from the image. Cameras are classified according to the standard deviation of sensor pattern noises of camera. The system is implemented in Raspberry Pi with python and openCV. The algorithm is tested with the imaging dataset containing 250 images of 10 different camera models. The system is successfully implemented in real time application with Pi cam and Web cam.

Index Terms—Sensor Pattern Noise, ICA (Independent Com-ponent Analysis), Source camera detection, Raspberry Pi.

I. INTRODUCTION

Usually the pictures are taken using any of the available digital cameras. As the cost of these devices falls, its func-tionality increased. It became a common object in day to day life. However the credibility of digital images lose because of the availability of the digital image manipulation tools and software. This may mislead one innocent person. Thus the problem of digital image authentication arises. The problem of digital image authentication can be approached in several ways. The simplest way is to check the electronic file itself which will contain the header information. Different image file formats like BMP, PCX, JPEG, PNG, TIFF and AVI files includes headers. These headers define the image size, number of colors and other information needed to display the image. Different headers contain different bytes of information. The exchangeable image file format (Exif) header contains the de-tails about camera type, time, date and exposure etc. Sufficient information will get by referring these headers. This relevant information may not be available if the given image is resaved into another file format or recompressed.

Another method is to embed an invisible or visible fragile watermark into the images. This may increase the credibility to the digital images. For example Epson Photo PC 700/750Z, 800/800Z and Kodak DC290 carry such watermark informa-tion about the digital camera or time stamp. But only relatively expensive cameras like DSLR support this functionality. Thus these approaches have only a limited applicability unless all cameras either insert watermarks or embed secure hashes in images. One of the previously investigated methods is identification of the cameras by examining the pixel defects of its images. The images may contain dead pixels (Defective pixels). Defective pixels are pixels in images that are not performing as expected; these pixels fail to sense light levels correctly. These defective pixels can be used as a fingerprint for identification of cameras. However there are cameras that eliminate these dead pixels by pre-processing.

The area of digital image authentication or source camera identification has been considering a major research work because of its importance of applications. It can be used as evidence in court for reliable identification purpose. It also provides authenticity of digital images ensuring the correct source device, in the same manner the bullet scratches match the bullet in a barrel. The researches have done in different ways for the concept of digital image authenticity using different methods like authenticity using sensor dust pattern, using traditional sensor pattern estimation techniques etc. But these methods are not efficient enough to provide accurate identification of source cameras even if the number of camera increases. Finally the studies showed that the sensor pattern noise estimation based on blind source separation provides better result for the camera identification. .The pro-posed system uses sensor pattern noise as a fingerprint for the camera identification purpose. De-noising is also done with the help of a systematic algorithm known as ICA (Independent Component Analysis).

The traditional approach for camera identification is based on the reference pattern noise [1]. The reference pattern noise of the camera under investigation is found by averaging the noises obtained from different images of the same camera. This reference pattern acts as a unique identification finger print. The presence of the reference noise pattern is identified by the correlation detector. In this case an unsatisfactory identification rate has found unless the large images are used. This identification technique causes desynchronization and incorrect camera detection. And it requires proper synchronization, cropping, resizing, rotation and digital zoom.

Another method proposes an approach for attenuating the influence of details from SPN pattern [2]. The hypothesis

Received the manuscript on march 2016, reviewed on December 2016,April 2017,Dec 2017 . Finalised on April 2018

used for this SPN enhancement method is that in SPN if the signal component is stronger, it is less trustworthy the component should be and thus the component must be at-tenuated. This hypothesis improves the device detection rate. The enhanced SPN can be obtained by assigning weighting factors which are inversely proportional to the magnitude of the SPN components. Digital single lens cameras have sensor dust problem due to interchangable lenses that they use [3] . These dust particles infront of imaging sensors make a pattern in all images taken by that camera. Camera identification technique has described based on this pattern. Dust spots are detected based on intensity loss model and shape properties. This method can not assure the uniqueness of dust pattern, it may lead false camera identification rate.

Many of the forensic tasks has been using a unique identifier called sensor photoresponse nonuniformity [4]. The problem identified in this method is that, when a person estimates this fingerprint from a group of images of a camera and combine it onto an image from another camera, it will effect innocent victim. This paper propose a reliable method for identifying such fake fingerprints. Triangle test is used to find fake finger print. But planting a sensor fingerprint in an image without leaving a trace is significantly more difficult. In another literature paper the camera is identified based on heteroscedastic noise model [5]. The hetero scedastic noise model describes a natural raw image. The approach contains the theory of statistical hypothesis testing. The heteroscedastic approach describes the noise variance. This noise variance varies linearly with the pixel's mathematical expectation. This approach mainly focus on raw images. The main limitation is that the raw images may not be available in practice. A camera raw image contains minimally processed data from the sensor and they are not ready to be printed. These difficulties can overcome using the proposed system with the real time implementation of camera detection.

II. DESIGN AND DEVELOPMENT

In some literature papers they ave selected different types of digital cameras available in the market and classified accord-ingly. This paper mainly focus on the real time implementation of the source camera detection system. Using both Pi cam and Web cam (one at a time), it take any picture from the current location independent of the environmental conditions like light variations under various focusing conditions of the camera. These pictures are stored in folders to create a dataset. By giving a test image as input to the system it can decide from which camera the image has taken. The database created for training the camera include more than eighty pictures. The proposed system setup is shown in the figure given below.

The block diagram shows the logic used in source camera identification system. An image from a Pi camera or Web camera is taken as test input for the system. Before estimating the noise pattern the images are subjected to pre-processing which will convert the image into gray and then filtered using median filter. The total noise in the image is calculated using the following equation.

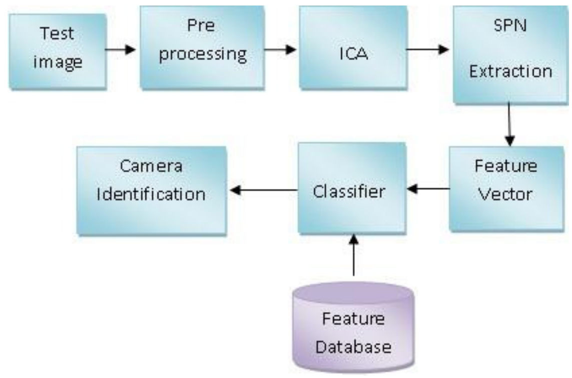

Fig. 1. The proposed system

$$N = I \quad F(I) \qquad (1)$$

N represents the overall noise in the image. I is the gray colored image and F(I) is the filtered image. The proposed algorithm called ICA (Independent Component Analysis) is then applied to this noise component. It will result three dif-ferent layers/components of noises. It is assuming that one of the component is the residual noise/sensor pattern noise of test image. A training vector is then formed for the classification of the camera. The training vector is a text file which contains the variance, standard deviation and the mean of the SPN of each reference images from the dataset. The features of SPN of both cameras are stored in a single text file(feature vector), which will further use for the classification of cameras. The sensor pattern noise of the reference camera is calculated in the same way as SPN of test image has calculated. The residual noise parameters (variance, standard deviation, mean)is saved to different variables. These variables are then compared with the estimated reference parameters which are already stored in the text file. The camera with lowest difference in standard deviation with that of the test image is chosen to be the source camera. In this paper the standard deviation is used for the classification purpose. According to the matching of standard deviation of SPN with that of test image the classification of the cameras can be done.

A. Signal processing in digital cameras

Simply converting an analog image that is captured by the CCD into digital data does not create a digital image. Only after the image processing engine performs a variety of calculations on a huge amount of digital image data one can see a completed color image. Since this processing is directly related to color accuracy, image detail and shooting performance, it ultimately determines the digital camera's performance level.

This section includes the processing of signals in a digital camera. The digital camera comprises of camera lenses, anti-aliasing filter, color filter array (CFA) and imaging sensor. The heart of digital camera is the imaging sensors. Commonly used imaging sensors in digital cameras are CCD and CMOS.

Received the manuscript on march 2016, reviewed on December 2016,April 2017,Dec 2017 . Finalised on April 2018

IJIIE/2018041002

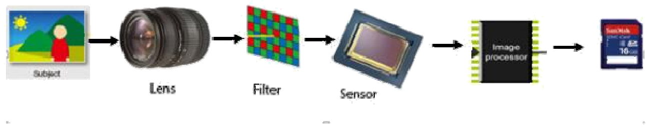

Fig. 2. Parts of a digital camera

The sensor is devided into tiny minimal addressable picture elements, which accepts photons and converts into voltage. This voltage is converted into a digital value with the help of a A/D converter. Before reaching into the sensor, the light passes through camera lenses, antialiasing filter and colour filter array. An image sensor or imaging sensor is a sensor that senses the information that constitutes an image. Image sensors are used in electronic imaging devices of both analog and digital types, which include digital cameras, camera modules, medical imaging equipment, night vision equipment such as thermal imaging devices, radar, sonar, and others.

A camera lens (also known as photographic lens or photographic objective) is an optical lens or assembly of lenses used in conjunction with a camera body and mechanism to make images of objects either on photographic film or on other media capable of storing an image chemically or electronically. In photography, a color filter array (CFA), or color filter mosaic (CFM), is a mosaic of tiny color filters placed over the pixel sensors of an image sensor to capture color information. Color filters are needed because the typical photosensors detect light intensity with little or no wavelength specificity, and therefore cannot separate color information. Since sensors are made of semiconductors they obey solid-state physics. The color filters filter the light by wavelength range, such that the separate filtered intensities include information about the color of light. The Foveon X3 sensor uses a different structure such that a pixel utilizes properties of multi-junctions to stack blue, green, and red sensors on top of each other. This arrangement does not require a demosaicing algorithm because each pixel has information about each color. An anti-aliasing filter (or low pass filter) helps to reduce distortion in fine patternsyou often see this in a suit or tie with a very tight pattern getting distorted,where stripes don't look like stripes anymore but start to swirl together.

B. Sensor pattern noise

The sensor pattern noise comes from the photon detector imperfections against light. The sensor pattern noise is esti-mated from the reference images of a particular camera. It is tested with unknown images of different cameras. In this paper the idea is to identify whether the unknown image is acquired using the reference camera presented. Sensor pattern noise is extracted with the help of fixed pattern noise (FPN). FPN is a noise found in digital imaging sensors. It is a general term that identifies a non-uniformity (forming a constant pattern) in an imaging system. It provides the same pattern of brighter and darker pixels present in images taken under same illumination conditions.

The digital cameras used in smart phones, camcorders and the scanners contain imaging sensors. This sensor contains many photon detectors, which sense the photons and converts it into electrical signal. However these photon detectors have imperfections in their manufacturing process or in the sensitiv-ity of photons, which produces a unique sensor pattern noise. This non-uniformity can be used to identify the reference camera presented.The way to extract SPN is shown in the figure given below.

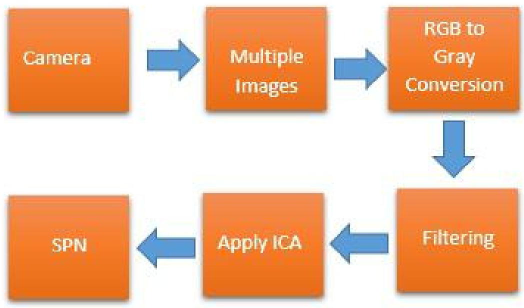

Fig. 3. SPN Extraction Process

For estimating the reference sensor pattern of a camera, it need multiple images taken by that camera. After some preprocessing ICA is applied to multiple images. The pre-processing include RGB to gray conversion and filtering and finally will get SPN. All these processes are done in python with help of openCV modules. Python development is much quicker than other languages. It supports a programming style that uses simple functions and variables without engaging in class definitions. It also supports writing much larger programs and better code reuse through a true object-oriented programming style. OpenCV (Open Source Computer Vision) is a library of programming functions mainly aimed at real-time computer vision, originally developed by Intel's research center in Nizhny Novgorod(Russia), later supported by Willow Garage and now maintained by Itseez.

C. Independent Component Analysis (ICA)

Independent component analysis (ICA) is a statistical and computational technique for revealing hidden factors that underlie sets of random variables, measurements, or signals. Imagine that you are in a room where two people are playing different instruments simultaneously. Two microphones are placed in different locations. The output of the microphones can be denoted as x1(t) and x2(t). The recorded signals are the combination of weighted sum of music signals emitted from two different instruments, which is denoted by s1(t) and s2(t). It could express this as a linear equation:

Received the manuscript on march 2016, reviewed on December 2016,April 2017,Dec 2017 . Finalised on April 2018

$$x_1(t) = a_{11}s_1 + a_{12}s_2 \quad (2)$$

$$x_2(t) = a_{21}s_1 + a_{22}s_2 \quad (3)$$

Where a11, a12, a21 and a22 are some parameters and whose value depends on the distance of microphones from players. Using only the recorded signals x1(t) and x2(t), it can now estimate the two original music signals s1(t) and s2(t). this is called the cocktail party problem.

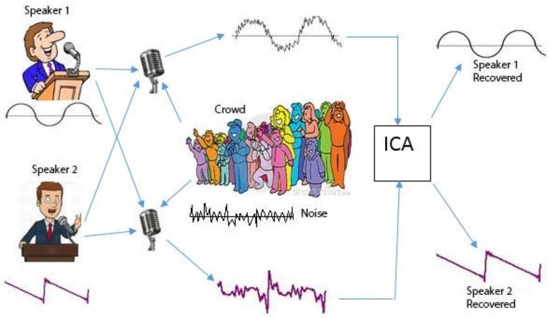

Fig. 4. The Cocktail Party Problem

For solving this problem, use some information on the statis-tical properties of the signals s1(t) to estimate aij. The recently developed technique of independent component analysis or ICA can be used to estimate aij, which allows us to separate the two original signals s1(t) and s2(t) from their mixtures. Assume that there are n linear mixture observed signals x1,x2.....xn which are mutually independent.

$$x_j = a_{j1}s_1 + a_{j2}s_2 + ::: + a_{jn}s_n \quad F \text{ or all } j \quad (4)$$

Now assume that both the mixture variables and the inde-pendent component have zero mean. For the ease of represen-tation vector-matrix notation is used here. Let us denote by x the random vector whose elements are the mixtures x1.....xn and s be the random vector with elements s1.....sn. Let us denote by A the matrix with element aij. Using this vector-matrix notation, the above mixing model is written as

$$X = AS \quad (5)$$

This model can also be written as

$$x = \sum_{i=1}^{n} a_i s_i \quad (6)$$

There are two main pre-processing stages in ICA, called centering and whitening. The primary reasons for preprocess-ing include simplification of algorithms, reduction of dimen-tionality of the problem and reduction of number of parameters to be estimated. In centering the mean is subtracted from each signal. Subtracting the mean is equivalent to translating the center of coordinates to the origin. The mean can always be re-added to the result at the end. Whitening is a transformation that converts the data such that it has an identity covariance matrix. The input data matrix x must be prewhitened or centered and whitened before applying the ica algorithm. PCA (Principal Component Analysis) is a technique used to the data dimension reduction. ICA can only separate linearly mixed sources. Since ICA separates sources by maximizing their non-Gaussianity, perfect Gaussian sources can not be separated. Even when the sources are not independent, ICA finds a space where they are maximally independents. ICA can be seen as an extension to principal component analysis and factor analysis. ICA is a much more powerful technique, however, capable of finding the underlying factors or sources when the classic methods fail completely.

III. EXPERIMENTS AND RESULTS

The system is implemented with Raspberry Pi and it is a series of credit card-sized single-board computers and it is based on a Broadcom SoC (System of Chip) with an ARM processor [700 MHz], a GPU and 256 to 512 MB RAM. The operating system used is Raspbian. Raspbian is a free operating system based on Debian optimized for the Raspberry Pi hardware. Although Raspberry Pi is as small as the size of a credit card, it works as if a normal computer at a relatively low price. By this system the camera will be detected as soon as it is taken.

In this paper two cameras are considered for the performance analysis of the system. The cameras are one pi cam and one web cam(iball face2face), without any special settings for each cameras more than eighty images from each camera were taken as reference. Images are taken in the different environmental conditions and light variations. These two cameras are connected to the Raspberry Pi board, which is then connected to a monitor. Pi cam is a high definition camera module compatible with the Raspberry Pi model A and model B. It includes omnivision 5647 CMOS image sensor in a fixed-focus module with integral IR filter and the resolution is 5-megapixel. Iball face 2face web cam is having high quality CMOS sensor and 4 LEDs for night vision, with brightness controller.

Fig. 5. Pi cam and Web cam(iBall face2face)

Iball face 2face web cam is having high quality CMOS sensor and 4 LEDs for night vision, with brightness controller.

ICA is the method also used in this paper for denoising images. It is a method for determining the components from multidimensional statistical data. What distinguishes ICA from other methods is that it finds a components that are statistically independent and nongaussian. Well-known algorithms for ICA include infomax, FastICA, JADE and kernel-independent com-ponent analysis. Among which FastICA is used in this paper

Fig. 6. Sample images of pi cam

Fig. 7. Sample images of web cam

because this algorithm is rapid and steady convergence speed advantage and it has higher separation accuracy. The ICA algorithm is tested using ten different digital cameras which are listed below. A database is formed using these cameras. From each camera 25 images have taken and total the dataset contains 250 images.

Fig. 8. Reference digital cameras

TABLE I

LIST OF DIGITAL CAMERAS

1	Sony HDR XR150
2	Sony DSC W320
3	Panasonic lumix DMC TZ70
4	Kodak M580
5	Nikon coolpix p610
6	Nikon coolpix L840
7	Canon EOS 1200D
8	Nikon coolpix S7000
9	Fujifilm X70
10	Sony cyber-shot DSC-RX1R II

There are different types of classifier functions available such as SVM, Naive Bayes classifier etc. These classifiers are small functions which will read some parameters of camera and will classify accordingly. Here it is just classifying the camera according to the standard deviation of the noises present in the camera. CV2 modules are available for cal-culating standard deviation. The coding is done in such a way that the source camera is the one having lowest difference in standard deviations of SPN of unknown image and that of the reference camera.

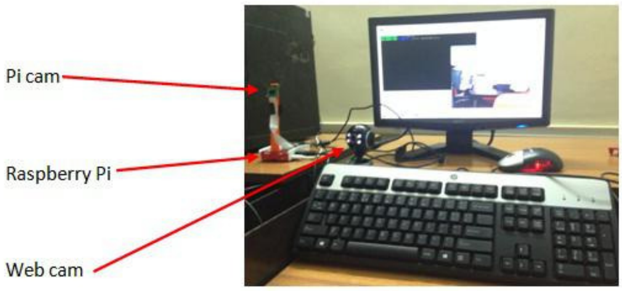

Fig. 9. Experimental setup

The sensor pattern noise parameters and the noise residual signal parameters of the test image have calculated already. These parameter can be used to train the cameras. The training process is a kind of sorting. Training will become more diffi-cult when the number of cameras increase. The training vector is formed, which includes the mean. variance and standard deviations of noise patterns of camera. Any parameters from the training vector can be used for matching. Correct detection of camera showed by taking only the standard deviation. Finally, the decision whether the unknown image is acquired by the reference camera can be made. Total 250 images were used for training the cameras.

IV. CONCLUSION

Digital imaging devices play an important role in todays world. The increased use of these devices like digital camera and smart phones help to create and manipulate images. By the use of image processing tools or software the images can manipulate in any manner. As a consequence such images can no longer authenticate. Source camera identification is one of the methods of image authentication. This paper presents a technique for identifying digital camera based on a new algorithm known as ICA (Independent Component Analysis) using SPN (Sensor Pattern Noise). FastICA is selected for implementing the system. This algorithm is tested with 10 different digital camera models. This system can be imple-mented for criminology investigation purposes. The system has implemented using Raspberry Pi. The results showed that the proposed technique could perform well and have better accuracy.

Received the manuscript on march 2016, reviewed on December 2016,April 2017,Dec 2017 . Finalised on April 2018

IJIIE/2018041002

REFERENCES

[1] Jan Luks, Jessica Fridrich and Miroslav Goljan, ”Digital Camera Identification From Sensor Pattern Noise”, IEEE transactions on information forensics and security, vol. 1, no. 2, june 2006
[2] Chang-Tsun Li, ”Source Camera Identification Using Enhanced Sensor Pattern Noise”, IEEE transactions on information forensics and security, vol. 5, no. 2, june 2010
[3] Ahmet Emir Dirik, Husrev Taha Sencar, and Nasir Memon, ”Digital Single Lens Reflex Camera Identification From Traces of Sensor Dust”, IEEE transactions on information forensics and security, vol. 3, no. 3, september 2008
[4] Miroslav Goljan, Jessica Fridrich, and Mo Chen, ”Defending Against Fingerprint-Copy Attack in Sensor-Based Camera Identification”, IEEE transactions on information forensics and security, vol. 6, no. 1, march 2011
[5] Thanh Hai Thai, Rmi Cogranne and Florent Retraint, ”Camera Model Identification Based on the Heteroscedastic Noise Model”, IEEE transactions on image processing, vol. 23, no. 1, january 2014
[6] Xiangui Kang, Yinxiang Li, Zhenhua Qu, and Jiwu Huang, ”Enhancing Source Camera Identification Performance With a Camera Reference Phase Sensor Pattern Noise”, IEEE transactions on information forensics and security, vol. 7, no. 2, april 2012
[7] Yoichi Tomioka, Yuya Ito, and Hitoshi Kitazawa, ”Robust Digital Camera Identification Based on Pairwise Magnitude Relations of Clustered Sensor Pattern Noise”, IEEE transactions on information forensics and security, vol. 8, no. 12, december 2013
[8] Amit Pande, Shaxun Chen, Prasant Mohapatra, and Joseph Zambreno, ”Hardware Architecture for Video Authentication Using Sensor Pattern Noise”, IEEE transactions on circuits and systems for video technology, vol. 24, no. 1, january 2014
[9] Yonggang Huang, Member, Jun Zhang, and Heyan Huang, ”Camera Model Identification With Unknown Models”, IEEE transactions on information forensics and security, vol. 10, no. 12, december 2015
[10] Ce Liu William T. Freeman and Richard Szeliski Sing Bing Kang, ”Noise Estimation from a Single Image”, IEEE Conference on Computer Vision and Pattern Recognition, 2006
[11] Mo Chen, Jessica Fridrich , Miroslav Goljan, and Jan Luks Determining Image Origin and Integrity Using Sensor Noise, IEEE Transactions on information forensics and security, VOL. 3, NO. 1, MARCH 2008
[12] Ashref Lawgaly, Fouad Khelifi, Ahmed Bouridane Image Sharpening for Efficient Source Camera Identification Based on Sensor Pattern Noise Estimation, Fourth International Conference on Emerging Security Technologies, 2013
[13] Yue Xufeng Lin, Chang-Tsun Li, Enhancing Sensor Pattern Noise via Filtering Distortion Removal, IEEE SIGNAL PROCESSING LETTERS, VOL. 23, NO. 3, MARCH 2016
[14] Xufeng Lin and Chang-Tsun Li Preprocessing Reference Sensor Pattern Noise via Spectrum Equalization, IEEE Transactions on information forensics and security, VOL. 11, NO. 1, JANUARY 2016
[15] Savitha R V, S.R.Breesha, X.Felix Joseph Pre Processing the Abdominal ECG signal using combination of FIR filter and Principal Component Analysis, International Conference on Circuit, Power and Computing Technologies, 2015