SOURCE IDENTIFICATION OF CAPTURED VIDEO USING PHOTO RESPONSE NON-UNIFORMITY NOISE PATTERN AND SVM CLASSIFIERS

ROYA ESMAEILANI

A dissertation submitted in partial fulfillment of the requirements for the award of the degree of
Master of Science (Computer Science)

Faculty of Computing
Universiti Teknologi Malaysia

JUNE 2014
This dissertation is dedicated to my beloved parents and sister for their endless support and encouragement.
ACKNOWLEDGEMENT

First and foremost, I would like to express heartfelt gratitude to my supervisor Prof. Dr. Ghazali Bin Sulong for his constant support during my study at UTM. He inspired me greatly to work on this project. His willingness to motivate me contributed tremendously to our project. I have learned a lot from him and I am fortunate to have him as my mentor and supervisor. Also, I would like to thank my sister for her help facilities and for her nice guidance during my project preparation. Furthermore, I would like to thank my parents and all my friends that have given their moral support during the ups and down of my project work life cycle.

Last but not least, I would like to thank the authority of Universiti Teknologi Malaysia (UTM) for providing me with a good environment and facilities such as Computer laboratory to complete this project with software which I need during process.
ABSTRACT

Recent works have shown that passive capturing source detection methods based on Photo-Response-Non-Uniformity (PRNU) extraction are the most reliable ones in comparison with techniques that based on lens properties or compression artifacts. Some important issues in this field include: employing an effective method for extracting PRNU, calculating the similarity and categorizing videos according to source of camera. In this study, a comprehensive algorithm is proposed to compare and evaluate the performance of different source detection methods in terms of filters used and partitioning process applied for PRNU extraction coupled with SVM classifier. Moreover, in consideration of observations, a new method is proposed for sampling selection using SVM classifier. Furthermore, the capabilities of employing and combining the results of different color parts of videos are used instead of changing them to grayscale. The proposed algorithm is based on three essential steps: Firstly, fingerprint of each camera, which is regarded as reference PRNU, is calculated by extracting PRNU of blue-sky videos. Secondly, the PRNU similarities of sample videos with reference PRNU are measured by calculating cross correlation and Peak to Correlation Energy (PCE) metrics. Finally, the sample videos are classified based on calculated PCE with SVM classifier. Experimental results revealed that Zero-mean and Wiener filters have small influences on PRNU, thus they can be ignored. Experimental results also revealed that eliminating the partitioning step considerably increases the performance of detection success rate by 15%. Among SVM classifiers, “RBF” and “MLP” types have the best identification rate of 75%.
ABSTRAK

Pengkajian terbaharu telah menunjukkan bahawa kaedah pengesanan sumber rakaman pasif berdasarkan Photo-Response-Non-Uniformity (PRNU) adalah sangat boleh dipercayai berbanding dengan teknik yang berdasarkan sifat-sifat kanta atau artifak mampatan. Beberapa isu penting dalam bidang ini termasuklah penggunaan kaedah yang lebih berkesan untuk pengekstrakan PRNU, pengiraan kesamaan dan pengkategorian video menurut sumber kamera. Dalam kajian ini algoritma komprehensif dicadangkan bagi membandingkan dan menilai prestasi kaedah pengesanan sumber yang berbeza dari segi penapis yang digunakan dan proses pembahagian yang diterapkan untuk pengekstrakan PRNU. Selanjutnya, dalam pertimbangan pemerhatian kaedah baharu dicadangkan untuk pemilihan pensampelan menggunakan pengkelas SVM. Selain itu keupayaan menggunakan dan menggabungkan hasil video bagi bahagian-bahagian warna yang berbeza digunakan bukannya mengubah hasil video tersebut kepada skala kelabu. Algoritma ini berdasarkan tiga langkah penting: Pertama, cap jari setiap kamera dikira dengan pengekstrakan PRNU video langit biru. Kedua, kesamaan PRNU video sampel dengan cap jari kamera diukur dengan metrik korelasi tenaga PCE. Akhir sekali, video sampel dikelaskan berdasarkan PCE yang dikira dengan pengkelas SVM. Keputusan eksperimen menunjukkan bahawa penggunaan zero-mean dan penapis Wiener mempunyai pengaruh yang kecil kepada ketepatan PRNU dan boleh diabaikan. Sementara itu, langkah menghapuskan pembahagian secara ketara meningkatkan prestasi kadar kejayaan pengesanan sebanyak 15%. Tambahan lagi, antara jenis pengkelas SVM, "RBF" dan "MLP" memiliki kadar kejayaan yang terbaik, iaitu sebanyak 75%
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>DECLARATION</td>
<td>ii</td>
<td></td>
</tr>
<tr>
<td>DEDICATION</td>
<td>iii</td>
<td></td>
</tr>
<tr>
<td>ACKNOWLEDGMENT</td>
<td>iv</td>
<td></td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>v</td>
<td></td>
</tr>
<tr>
<td>ABSTRAK</td>
<td>vi</td>
<td></td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td>vii</td>
<td></td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xi</td>
<td></td>
</tr>
<tr>
<td><strong>1</strong> INTRODUCTION</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.1 Introduction</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>1.2 Problem Background</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>1.3 Problem Statement</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>1.4 Research Questions</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>1.5 Aim of Study</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>1.6 Objectives</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>1.7 Scope of Study</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>1.8 Significance of study</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>1.9 Thesis Organization</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td><strong>2</strong> LITERATURE REVIEW</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.1 Introduction</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>2.2 Multimedia Forgeries</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>2.2.1 Tampering in Spatial Domain</td>
<td>11</td>
<td></td>
</tr>
</tbody>
</table>
2.2.2 Tampering in Temporal Domain 12
2.2.3 Tampering in Spatio-Temporal Domain 13
2.2.4 Tampering in Source of Digital Files 13
2.3 Forgery Detection Methods without Pre-processing Activities 14

2.3.1 Source Identification Based Forgery Detection Methods 14

2.3.1.1 Noise Characteristics of Acquisition Device 16
2.3.1.2 Lens Characteristics of Acquisition Device 20

2.3.2 Coding Artifacts Based Video Forgery Identification Methods 21

2.3.3 Video Forgery Identification Methods Based on Detecting Inconsistency in Context 23

2.3.4 Detection of Copy-Move Forgeries 24

2.4 Support Vector Machines (SVM) 25

2.4.1 Linear 27
2.4.2 Polynomial 27
2.4.3 Gaussian Radial Basis Function 27
2.4.4 Multi-Layer Perceptron 28
2.4.5 Quadratic Kernel 28
2.4.6 Optimization Problem in SVM Classifiers 29

2.4.6.1 Interior Point Algorithms 30
2.4.6.2 Chunking and Sequential Minimal Optimization (SMO) 30

2.5 Related studies 31

2.6 Summary 35

3 RESEARCH METHODOLOGY 36

3.1 Introduction 36
3.2 Research Environment 36
3.3 Research Framework 37
3.4 Extract PRNU Noise for Each Video 38
3.5 Extract Reference PRNU Noise for Each Camera 41
3.6 Apply PRNU Extraction Enhancement Filters 42
3.7 Calculate Peak to Correlation Energy (PCE) 44
3.8 Classify Videos Based on Their Sources using SVM Classifier 46
  3.8.1 Training Phase 46
  3.8.2 Source Identification and Classification Phase 47
3.9 Summary 49

4 EXPERIMENTAL RESULT AND DISCUSSION
4.1 Introduction 50
4.2 Dataset 51
4.3 Experimental Results 51
  4.3.1 Results of Extracting PRNU Factor 52
  4.3.2 Results of Comparing Fingerprints 56
  4.3.3 Schematic Comparison of Video Classification by SVM 61
  4.3.4 Results of Classification Videos 65
4.4 Summary 75

5 CONCLUSION
5.1 Introduction 76
5.2 Research Contribution 77
5.3 Future Work 78

REFERENCES 80
### LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Illumination-Independent Noise Types</td>
<td>19</td>
</tr>
<tr>
<td>2.2</td>
<td>Illumination-dependent Noise Types</td>
<td>19</td>
</tr>
<tr>
<td>2.3</td>
<td>Digital Processing Noise Types</td>
<td>20</td>
</tr>
<tr>
<td>2.4</td>
<td>Summary of Related Studies</td>
<td>32</td>
</tr>
<tr>
<td>4.1</td>
<td>Best Minimum success rates of classification by selecting a maximum of PCE as sample type and 3 out of 3 from RGB results</td>
<td>74</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Processing pipeline diagram inside digital cameras</td>
<td>4</td>
</tr>
<tr>
<td>2.1</td>
<td>Multimedia authentication techniques</td>
<td>10</td>
</tr>
<tr>
<td>2.2</td>
<td>Complete classification of pattern noise</td>
<td>17</td>
</tr>
<tr>
<td>2.3</td>
<td>Model of existing noises in captured images and frames</td>
<td>18</td>
</tr>
<tr>
<td>2.4</td>
<td>Distortion of a rectangular grid. Left: Undistorted grid. Middle: Grid with barrel distortion. Right: Grid with pincushion distortion</td>
<td>21</td>
</tr>
<tr>
<td>3.1</td>
<td>Capturing Source Detection Operational framework</td>
<td>37</td>
</tr>
<tr>
<td>3.2</td>
<td>PRNU extraction process</td>
<td>43</td>
</tr>
<tr>
<td>3.3</td>
<td>The existence of peak in correlation matrix (match)</td>
<td>45</td>
</tr>
<tr>
<td>3.4</td>
<td>There is no peak in correlation matrix (mismatch)</td>
<td>45</td>
</tr>
<tr>
<td>3.5</td>
<td>Detailed Algorithm Steps Chart</td>
<td>48</td>
</tr>
<tr>
<td>4.1</td>
<td>Red colour channel PRNU factor of a blue-sky video captured by camera number 4 with 512*512 block size</td>
<td>53</td>
</tr>
<tr>
<td>4.2</td>
<td>Green colour channel PRNU factor of a blue-sky video captured by camera number 1 with 512*512 block size</td>
<td>53</td>
</tr>
<tr>
<td>4.3</td>
<td>Fingerprints after applying Zero-mean filter and WienerInDFT filters and without using RGB-to-greyscale filter of a blue-sky video captured by camera number 4 with block size 512*512</td>
<td>54</td>
</tr>
<tr>
<td>4.4</td>
<td>Fingerprint after applying Zero-mean filter, WienerInDFT filter and RGB-to-greyscale filter of a blue-sky video captured by camera number 4 with block size 512*512</td>
<td>54</td>
</tr>
</tbody>
</table>
| 4.5    | Fingerprints after applying Zero-mean filter and WienerInDFT filters and without using RGB-to-
grey scale filter of a blue-sky video captured by camera number 1 with block size 512*512

4.6 Fingerprint after applying Zero-mean filter and WienerInDFT filters and using RGB-to-grey scale filter of a blue-sky video captured by camera number 1 with block size 512*512

4.7 Cross correlation of two extracted fingerprints from two equal blue-sky videos captured by camera number 4

4.8 Cross correlation of two extracted fingerprints from two parts of a blue video captured by camera number 4

4.9 Cross-correlation of two extracted fingerprints from different videos captured by the same camera

4.10 Cross correlation of two extracted fingerprints from two videos captured by different cameras

4.11 Average of PCE values between a conventional videos and blue videos of a camera

4.12 Video classification for Camera No. 1 by using Linear kernel, QP type SVM

4.13 Video classification for Camera No. 2 by using Quadratic kernel, QP type SVM

4.14 Video classification for Camera No. 3 by using Quadratic kernel, SMO type SVM

4.15 Video classification for Camera No. 4 by using Polynomial kernel, SMO type SVM

4.16 Video classification for Camera No. 4 by using MLP kernel, LS type SVM

4.17 Video classification for Camera No. 3 by using RBF kernel, QP type SVM

4.18 Success rates of video files classification by selecting properties as follow: not applying Zero-mean, WienerInDFT, RGB-to-grey scale and partitioning step, using RBF kernel QP type SVM

4.19 Success rates of video files classification by selecting properties as follow: not applying Zero-mean, WienerInDFT, RGB-to-grey scale and
partitioning step, using RBF kernel SMO type SVM

4.20 Success rates of video files classification by selecting properties as follow: not applying Zero-mean, WienerINDFT, RGB-to-greyscale and partitioning step, using RBF kernel LS type SVM

4.21 Success rates of video files classification by selecting properties as follow: not applying Zero-mean, WienerINDFT, RGB-to-greyscale and partitioning step, using MLP kernel SMO type SVM

4.22 Success rates of video files classification by selecting properties as follow: not applying Zero-mean, WienerINDFT, RGB-to-greyscale and partitioning step, using MLP kernel LS type SVM

4.23 Success rates of video files classification by selecting properties as follow: not applying Zero-mean, WienerINDFT, RGB-to-greyscale and partitioning step, using RBF kernel QP type SVM

4.24 Success rates of video files classification by selecting properties as follow: applying Zero-mean, but not applying WienerINDFT, RGB-to-greyscale and partitioning step, using RBF kernel SMO type SVM

4.25 Success rates of video files classification by selecting properties as follow: applying Zero-mean but not applying WienerINDFT, RGB-to-greyscale and partitioning step, using RBF kernel LS type SVM

4.26 Success rates of video files classification by selecting properties as follow: applying Zero-mean but not applying WienerINDFT, RGB-to-greyscale and partitioning step, using MLP kernel SMO type SVM

4.27 Success rates of video files classification by selecting properties as follow: applying Zero-mean but not applying WienerINDFT, RGB-to-greyscale and partitioning step, using MLP kernel LS type SVM

4.28 Success rates of video files classification by selecting properties as follow: applying Zero-mean and WienerINDFT but not applying RGB-to-greyscale and partitioning step, using RBF kernel
QP type SVM

4.29 Success rates of video files classification by selecting properties as follow: applying Zero-mean and WienerInDFT but not applying RGB-to-greyscale and partitioning step, using RBF kernel SMO type SVM

4.30 Success rates of video files classification by selecting properties as follow: applying Zero-mean and WienerINDFT but not applying RGB-to-greyscale and partitioning step, using RBF kernel LS type SVM

4.31 Success rates of video files classification by selecting properties as follow: applying Zero-mean and WienerInDFT but not applying RGB-to-greyscale and partitioning step, using MLP kernel SMO type SVM

4.32 Success rates of video files classification by selecting properties as follow: applying Zero-mean and WienerINDFT but not applying RGB-to-greyscale and partitioning step, using MLP kernel LS type SVM
CHAPTER 1

INTRODUCTION

1.1 Introduction

Nowadays, digital file producing devices like several kinds of cell phones, camcorders, cameras and scanners and the possibility of sharing information through internet causes usage of digital videos to increase in different aspects of human lives such as surveillance cameras, home video and so on. The improvements in digital technology with providing easy access to digital video, besides a lot of advantages can create some problems which are the results of illegal activities and may lead to creating diverse origins or some alterations to contents of video files.

Video authentication can be defined as a process which proves the given video content is exactly same as when it was captured by camera and due to its usage can be achieved by searching and detecting different types of forensics which maybe done about video through several ways (Upadhyay and Singh, 2012).

A video clip can be doctored by several forensics such as altering, combining, or creating new video contents. All these forgery activities that are done usually with criminal goals lead to producing a fake video which conceals or changes the determinative facts, events and important details in the recorded scene; So that, authenticity of video data can be considered as an important factor in cases such as forensic investigation, law enforcement, video surveillance and content ownership.
For example, in surveillance videos which usually are used as legal evidences in many courts, the accuracy of content is essential. These surveillance videos can be easily forged using video editing tools such as Premier, Vegas, etc. Moreover, the forger might replace the whole video taken by the surveillance camera with another video. In the cases that forged surveillance videos are used as court evidence, it may potentially cause the wrong person to be convicted mistakenly. Generally, the most important and sensitive cases which clarify the necessity of video authentication are related to scenarios in which video clips used as evidence in court law and even a little modification can change the sentence (Pradeep K et al, 2009; Upadhyay and Singh, 2012; Pradeep K et al, 2004).

Digital forensics often leaves some traces in resulting signal which called “fingerprints”. Fingerprint are usually hidden in digital video and it is needed to complex signal processing for analyzing and detecting them. Uncovering and inspecting these fingerprints provide a kind of reverse engineering to find and understand the processing steps which is performed on video from its first generation to its actual form and can be employed for finding tampering areas (Bestagini et al, 2012; Tsai and Shih, 2009).

1.2. Problem Background

Nowadays, fast improvements in digital technologies and widely use of digital video recording systems together with sophisticated video editing software, high quality processing tools and algorithms, and accessibility of low-cost and easy-operable digital multimedia devices, increase the trend for tampering videos and make the authenticating multimedia content as a challenging issue.

There are two general approaches for authenticating video clips. In the first approach of forgery detection, a watermark or digital signature is embedded into the video. These determinant files can be saved into video content, header file or even as independent files. By using embedded watermarks in video content, whenever a
video is forged, this watermark also be modified and can be processed by authentication systems as a clue of forgery. In the cases of using digital signatures, tampering in videos can be determined by extracting the digital signature and matching the data of video content with information obtained from the digital signature. In these methods, the requirement of a preprocessing modules which should be embedded in the device and it can influence on video quality, can be considered as an important drawback (Upadhyay and Singh, 2012).

In the second approach, in comparison to first one, there is no need to any pre-processing activities. These methods rely on intrinsic features like pixel value and statistical features and characteristics of video files. The main steps of these approaches is extracting different types of fingerprints and then applying pattern recognition techniques in order to detect forgery. Some fingerprints which can be used in these techniques include: noise patterns, lens distortion, double compression artifacts, inconsistence-related artifacts and so on (Kancherla and Mukkamala, 2012).

One of the methods which is considered in recent years in order to detect video forgery, are techniques that are based on identifying of acquisition device and able to detect whether two video clips originate from the same source. Based on Alex C. Kot and Hong Cao, because of statistical source features fragile nature toward tampering and manipulations, it can be used to address tamper issue.

Researches on camera model identification focus on certain stages of processing pipeline inside digital cameras because each part of this process could be implemented differently in various camera types. An overview of this processing pipeline is illustrated in Figure 1.1(Xu and Shi, 2012).
Processing pipeline stages that are shown in Figure 1.1 can be divided into two main categories: Software stages and hardware parts. Camera identification methods rely on software parts consists of compression procedure parameters and examining the digital file’s header which contains the information about camera type, date and time and exposure. But these methods are not so reliable and not widely used because each camera could be adjusted in different settings and the resulted image could be compressed even more by additional software on computers before usage.

Because of that, it’s better to switch on hardware parts and identify the source of digital files based on the imperfections of camera hardware. These hardware parts consist of lens and CCD or CMOS sensors. Each of these hardware devices leave some fingerprints in resulted images or videos which can be used as intrinsic features of each camera type (Van Lanh et al., 2007; Irie et al., 2008).

Lens distortion parameters, can capture the geometric fingerprints which are left by the camera (lens system) on the digital files. Among different types of lens aberrations, lens radial distortion seems more suitable for camera source identification; however the lens distortion parameter is affected by manual zooming and this causes the performance of this method to limit (Choi et al., 2006).
As it is mentioned above, another set of video source identification methods are based on extracting and measuring noise characteristics which are resulted of camera sensors. Generally, noise is a random, unwanted and fluctuation of pixel values in digital file such as images and videos that is produced by sensors. The noise patterns which are mostly used as one part of source identification process because of their deterministic properties, come from CCD sensors (Bayram et al., 2005).

Among all CCD’s noise types which can be detected in a captured video, PRNU can be used as an effective fingerprint for camera identification because of some its properties such as resistance against humidity, temperature, light refraction on dust particles, optical surfaces and optical zoom setting and also its source that is camera sensor (Chen et al., 2007).

1.3. Problem Statement

As it is mentioned in the last section, capturing source detection methods is considered as an effective approach in authenticating video files. Based on previous studies which are performed in this area, PRNU noise is determined as one of the best deterministic fingerprint for identifying the source (Bayram et al., 2005; Chen et al. 2007; Peng et al., 2013; Goljan et al. 2009). One of the challenging issues which raises in these studies, is implementing an appropriate method for evaluating the accuracy of extracted PRNUs. In this study, the efficiency of PRNU-based identification methods is investigated by proposing the comprehensive evaluating algorithm. Moreover, the best and most effective methodology in order to identify the videos’ capturing source by means of PRNU is introduced. The presented algorithm in comparison with the last study in which videos should belong enough to obtain more reliable results, is independent of videos’ length (Chen and et al., 2007). Finally, the proposed method uses artificial intelligence classifiers for classifying a
large number of videos in consideration to extracted PRNUs. The success rates of these classifiers can be considered for evaluating the accuracy of extracted PRNU.

1.4. Research Questions

1. How to detect PRNU noise from captured video by digital cameras?
2. How to extract PRNU noise? Is it efficient to use filters? Is it efficient to use partitioning process?
3. How to authenticate the source of the video by using PRNU noise?
4. What is the best classifier for classifying videos based on extracted PRNU?

1.5. Aim of Study

The aim of this study is to propose and implement a comprehensive evaluating algorithm for video forgery detection method based on source of camera used and classify videos according to similarity metric (Peak to Correlation Energy).

1.6. Objectives

1. To detect PRNU noise from captured video by digital cameras.
2. To extract PRNU noise by employing different filters.
3. To classify videos based on captured PRNU noise and similarity metrics using the SVM classifier.
1.7. **Scope of Study**

1. Prior knowledge of the camera used is known.
2. SVM (Support Vector Machine) tools are used for classifying the video based on similarity metric (PCE).
3. Several numbers of videos which are taken by four types of camera and included “SONY DSC-W390”, ”CANON POWERSHOT GX1”, “CANON G12” and “SONY DSC-HX5V, are used for providing training and testing dataset.
4. Matlab2011a software is used as the processing tool for capturing the video noise pattern and implementing the evaluating phase.

1.8. **Significance of Study**

In recent years, increasing trend of using digital videos and sharing video clips easily throughout internet on one hand, and accessible enhanced tools and software tools which are produced in order to tamper videos on the other hand create an essential need to authenticate the video files. Authenticating of video files can be considered as an important issue since they can be used as critical evidences to prove criminal activities which are done by individuals and can change the sentence in court of law.

1.9 **Thesis Organization**

The research is comprised of five chapters. Chapter one presents an introduction to the research which includes the problem background, problem statements, research questions, aim of the study, the main objectives, scope and significance of the study. Chapter two reviews the literature about video
authentication and investigates the methods used in video forgery detection. Chapter three describes the methodology for the research. Chapter four looks into the design of algorithms and displays the results. Lastly, the conclusion is presented in chapter five.


