Source Camera Identification
Forensics Based on Wavelet Features

Bo Wang, Yiping Guo, Xiangwei Kong, Fanjie Meng

Dalian University of Technology, China
Outline

- Introduction
- Image features based identification
- Kharrazi’s method
- Our method
- Experimental results and conclusions
Introduction

Source Camera Identification:
Identifying the source camera of a digital photograph

Used for:
Establishing the origin of legal photographic evidence
Active and Passive Identification

- **Active Identification**
  - Embed watermarks
  No watermarks in most of digital photographics

- **Passive Identification**
  - Do not need embed any information
  - Only using image data
  Our method is a passive identification
Using EXIF for Identification

- which one is its original EXIF? The left one.
- The EXIF of the right one is replaced by another image.
Image Features Based Identification

- **Imaging pipeline in digital cameras**

  ![Diagram of imaging pipeline]

  - Scene Radiance
  - Optical system (Lens Distortion, Optical lowpass)
  - Sensor noise
  - Color Filter Array + Sensor
  - Demoscaicing
  - White balancing, Image sharpening, Contrast enhancement, Gamma Correction
  - Digital image

Differences in the processing details of each stage among various models of digital cameras

Differences of **image features** in the output images from cameras of different models
Kharrazi’s Method

- Polytechnic University, Brooklyn, NY, USA: Mehdi Kharrazi, Husrev T. Sencar, Nasir Memon

- Using Pattern Recognition

- Image Features: color features, IQM features, mean of wavelet coefficients
Can we do better?

- Shortage of Kharrazi's method
- Identification accuracy is not reliable
- Why?
- Image Features used are not effective
- What we do?
- Extract more effective features
Our method

- Features Extraction
- Features Selection
- Classification
Wavelet Features

• Higher-order wavelet statistics

  ➢ Statistics of linear prediction of wavelet coefficients
    - A kind of filter operation in wavelet domain
    - Less dependence on image content

• Wavelet Coefficient Co-occurrence statistics

  ➢ Distances of co-occurrence matrices in the same orientation between different scales
Higher-order Wavelet Features

Digital Image $\xrightarrow{\text{four-scale wavelet decomposition}}$ $HL_{r,g,b}^{1,2,3}$ $HH_{r,g,b}^{1,2,3}$ $LH_{r,g,b}^{1,2,3}$ $\xrightarrow{\text{mean variance skewness kurtosis}}$ statistical features

\[ |V_i^g(x, y)| = \omega_1 |V_i^g(x - 1, y)| + \omega_2 |V_i^g(x + 1, y)| + \omega_3 |V_i^g(x, y - 1)| + \omega_4 |V_i^g(x, y + 1)| + \omega_5 |V_{i+1}^g(x, y/2)| + \omega_6 |V_i^g(x, y)| + \omega_7 |D_{i+1}^g(x/2, y/2)| + \omega_8 |V_{i}^{\tau}(x, y)| + \omega_9 |V_i^b(x, y)| \]

\[ \bar{\nu} = Q\hat{\omega} \quad E(\hat{\omega}) = [\bar{\nu} - Q\hat{\omega}]^2 \quad \frac{dE(\hat{\omega})}{d\hat{\omega}} = 2Q^T(\bar{\nu} - Q\hat{\omega}) \]

\[ \hat{\omega} = (Q^TQ)^{-1}Q^T\bar{\nu} \quad \bar{p} = \log(\bar{\nu}) - \log(|Q\hat{\omega}|) \]
Wavelet Coefficient Co-occurrence Statistics

\[
\begin{align*}
    DC(V_i^c) &= CV_i^c - CV_{i+1}^c \\
    DC(H_i^c) &= CH_i^c - CH_{i+1}^c \\
    DC(D_i^c) &= CD_i^c - CD_{i+1}^c \\
    CV_i^c, CH_i^c, CD_i^c : & \text{ vertical, horizontal, and diagonal subbands' co-occurrence matrices} \\
    i &= 1, 2, 3, 4 \quad c = r, g, b \\
\end{align*}
\]

- Energy: \( \sum \sum DC[i, j]^2 \)
- Entropy: \( -\sum \sum DC[i, j] \log_2 DC[i, j] \)
- Contrast: \( \sum \sum (i-j)^2 DC[i, j] \)
- Homogeneity: \( \sum \sum \frac{DC(i, j)}{1 + |i-j|} \)
- Correlation: \( \frac{\sum \sum (i-\mu_i)(j-\mu_j)DC[i, j]}{\sigma_i \sigma_j} \)
Feature Selection and Classification

• Sequential Forward Feature Selection (SFFS)

• Support Vector Machine (SVM)
  – C-support vector classification with non-linear RBF kernel
## Experiment

- Experiment samples and parameters

<table>
<thead>
<tr>
<th>Cameras</th>
<th>Camera Parameters</th>
<th>Sample image parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sensor</td>
<td>Max resolution</td>
</tr>
<tr>
<td>Kodak DC290</td>
<td>Unspecified CCD</td>
<td>2240*1500</td>
</tr>
<tr>
<td>Nikon E5700</td>
<td>2/3 inch CCD</td>
<td>2560*1920</td>
</tr>
<tr>
<td>Sony DSC-F828</td>
<td>2/3 inch CCD</td>
<td>3264*2448</td>
</tr>
<tr>
<td>Canon PowerShot Pro1</td>
<td>2/3 inch CCD</td>
<td>3264*2448</td>
</tr>
<tr>
<td>Canon PowerShot G2</td>
<td>1/1.8 inch CCD</td>
<td>2272*1704</td>
</tr>
<tr>
<td>Canon PowerShot G3</td>
<td>1/1.8 inch CCD</td>
<td>2272*1704</td>
</tr>
</tbody>
</table>
Experiment result of our method

- Confusion matrix

<table>
<thead>
<tr>
<th>Camera</th>
<th>Kodak</th>
<th>Nikon</th>
<th>Sony</th>
<th>CanonPro1</th>
<th>CanonG2</th>
<th>CanonG3</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kodak DC290</td>
<td>150</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100%</td>
</tr>
<tr>
<td>Nikon 5700</td>
<td>0</td>
<td>148</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>98.7%</td>
</tr>
<tr>
<td>Sony DSC-F828</td>
<td>0</td>
<td>2</td>
<td>148</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>98.7%</td>
</tr>
<tr>
<td>Canon PowerShot Pro1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>145</td>
<td>4</td>
<td>0</td>
<td>96.7%</td>
</tr>
<tr>
<td>Canon PowerShot G2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>143</td>
<td>4</td>
<td>95.3%</td>
</tr>
<tr>
<td>Canon PowerShot G3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>148</td>
<td>98.7%</td>
</tr>
</tbody>
</table>
Comparison with Kharrazi’s method

<table>
<thead>
<tr>
<th>Camera</th>
<th>Kodak</th>
<th>Nikon</th>
<th>Sony</th>
<th>CanonPro1</th>
<th>CanonG2</th>
<th>CanonG3</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kharrazi’s method</td>
<td>94.7%</td>
<td>91.3%</td>
<td>96.3%</td>
<td>85.3%</td>
<td>84.7%</td>
<td>93.3%</td>
<td>90.9%</td>
</tr>
<tr>
<td>Our method</td>
<td>100%</td>
<td>98.7%</td>
<td>98.7%</td>
<td>96.7%↑1.4%</td>
<td>95.3%↑10.6%</td>
<td>98.7%↑5.4%</td>
<td>98.2%</td>
</tr>
</tbody>
</table>

IIH-MSP-2009
Dalian University of Technology
Conclusions

1. Introduce feature based source camera identification

2. Discuss a classic feature based identification method

3. Give a new source camera identification method based on wavelet features
Thank you!