Transportation-theoretic image counterforensics to First Significant Digit histogram forensics

Cecilia Pasquini
cecilia.pasquini@unitn.it

Giulia Boato
boato@disi.unitn.it

University of Trento,
DISI, Multimedia Signal Processing and Understanding Lab

Pedro Comesaña-Alfaro
comesana@gts.uvigo.es

Fernando Pérez-González
perez@gts.uvigo.es

University of Vigo,
Signal Processing in Communications Group
**Motivation**

**Multimedia forensic scenario**

1. Manipulation
2. Forensic analysis

**Adversarial environment**

1. Manipulation
2. Counter-forensics action
3. Forensic analysis

---

*MMSP 2014, GTTI Thematic Meeting, Forni di Sopra (UD), 16-18/02/2014*
Adversarial signal processing

- Study of general theoretic solutions to problems involving an attacker and a defender
  [Barni and Pérez-González, 2013]

- For the case of multimedia forensics, tradeoff between the fidelity of the statistics recovered and the distortion introduced in the object
  [Kirchner and Boehme, 2013]

- Analysis of the optimal strategy for one or both players under specific hypothesis
  [Barni and Tondi, 2013] [Barni et al., 2012]
  [Comesaña-Alfaro and Pérez-González, 2013]
First Significant Digits in JPEG image forensics

Detectors discriminating single, double or multiple compressed images using FSD probabilities as features
[Li et al., 2008]
[Milani et al., 2012]

Attacks modifying the distribution of FSDs in order to mislead the forensic analysis
[Milani et al., 2013]
[Pasquini and Boato, 2013]

GENERAL PROBLEM: Is there an optimal strategy for a generic detector?

Number of compressions

0
1
2
n
Outline

- Transportation-theoretic formulation of the problem as a two-step optimization process

- Design of a novel technique providing an approximated solution of the optimization problem for a class of distortion measures (including the MSE)

- Analysis of the distortion introduced on a dataset of images and comparison with state-of-the-art method
Basic definitions

**Pixel** \( x \in \mathbb{R}^N \) \quad \rightarrow \quad **DCT** \quad y = DCT(x) \quad \rightarrow \quad **FSD** \quad d = FSD(y) \quad \rightarrow \quad **FSD histogram** \quad h = H(d)

\[ y \in \mathbb{R}^N \quad d \in \{0, \ldots, 9\}^N \quad h \in \{0, \ldots, N\}^{10} \]

**Distance** \( g^y(y,y') \) between the DCT vectors \( y \) and \( y' \)

Fixed \( \bar{y} \in \mathbb{R}^N \), we define \( \bar{d} = FSD(y) \) and

\[ g^d(\bar{d}, d) := \min_{\{y|FSD(y)=\bar{d}\}} g^y(\bar{y}, y) \quad \text{distance between the FSD vectors } d \text{ and } \bar{d} \]

**Problem**: given a starting DCT vector \( \bar{y} \) and a target histogram \( h^* \), modify \( \bar{y} \) minimizing the distortion:

\[ y^* = \arg\min_{\{y|H(FSD(y))=h^*\}} g^y(\bar{y}, y) \]
The problem can be formulated as a two-step optimization process:

\[
\begin{align*}
\mathbf{d}^\# &= \arg \min_{\mathbf{d}} g^d(\tilde{\mathbf{d}}, \mathbf{d}), \\
\mathbf{y}^* &= \arg \min_{\mathbf{y}} g^y(\tilde{\mathbf{y}}, \mathbf{y})
\end{align*}
\]

Computationally unfeasible!
Approximated optimization technique I

**Assumption:** \[ g^y(\bar{y}, y) = \sum_{j=1}^{N} g(y_j, \bar{y}_j) \] where \( g(\cdot, \cdot) \) is a symmetric convex function depending on the difference between its input arguments (it holds for any metrics induced by a \( L^p \) norm, like the MSE).

Given the subset \( S \subset \mathbb{R} \) to which we can move the initial values, we define

\[
 f_S(y, d) := \arg \min_{\{y' \in S | FSD(y') = d\}} |y - y'|, \quad \text{Dist}_S(y, d) := |y - f_S(y, d)|
\]

so that each \( g(\bar{y}_j, y_j) \) is minimized. \( \text{EX: } S \text{ integer numbers, } f_S(34, 3) = 34, f_S(34, 2) = 29 \)

**NOVEL PROCEDURE:**
- a new DCT vector \( z \) is determined starting from a given input vector \( \bar{y} \) and a target histogram \( h^* \)
- for every component of \( \bar{y} \) in descending order of magnitude, a new FSD is chosen by means of the map \( f_S(\cdot, \cdot) \)

MMSP 2014, GTTI Thematic Meeting, Forni di Sopra (UD), 16-18/02/2014
**Approximated optimization technique II**

\[ D_0 \] set of \( N \) digits with histogram \( h^* \)

\( \bar{y} \) is sorted in descending order

**ITERATIVE PROCEDURE**

For \( j = 0, \ldots, N - 1 \)

\[
\begin{bmatrix}
  \cdots \\
  y_j \\
  \cdots \\
\end{bmatrix}
\]

\( D_j \)

Minimization of \( Dist_S(y_j, \cdot) \)

\[
\begin{bmatrix}
  \cdots \\
  z_j \\
  \cdots \\
\end{bmatrix}
\]

Definition of \( z_j := f_S(y_j, p) \)

Extraction from the set \( D_j \) of the digit \( p \) obtained

After a reordering operation, \( z \) is obtained as approximated solution instead of the exact one \( y^* \)

---

**EX:**

\[ y_j = 48 \]

\[ D_j = \{1, 2, 6, 9, 9\} \]

\[ p = 6 \]

\[ z_j = 60 \]
Experimental setting

- Three different binary hypothesis testing problems:
  
  \[ H_0: 0 \text{ compressions} \quad \text{vs.} \quad H_1: 1 \text{ compression} \]
  
  \[ H_0: 0 \text{ compressions} \quad \text{vs.} \quad H_1: 2 \text{ compressions} \]
  
  \[ H_0: 1 \text{ compression} \quad \text{vs.} \quad H_1: 2 \text{ compressions} \]

  where different quality factors have been considered.

- UCID dataset: 600 images for creating the reference target histograms (by means of an averaging operation), 738 images for the testing phase;

- Comparison with a state-of-the-art method based on waterfilling techniques in terms of mean MSE (converted to PSNR) over the testing set [Milani et al., 2013]
## Experimental results

### $1 \rightarrow 0$

<table>
<thead>
<tr>
<th>QF1</th>
<th>50</th>
<th>60</th>
<th>70</th>
<th>80</th>
<th>90</th>
</tr>
</thead>
<tbody>
<tr>
<td>TT</td>
<td>41.03</td>
<td>41.56</td>
<td>42.11</td>
<td>43.38</td>
<td>46.22</td>
</tr>
<tr>
<td>WF</td>
<td>34.47</td>
<td>34.22</td>
<td>34.58</td>
<td>34.91</td>
<td>43.77</td>
</tr>
</tbody>
</table>

### $2 \rightarrow 0$

<table>
<thead>
<tr>
<th>QF2</th>
<th>50</th>
<th>60</th>
<th>70</th>
<th>80</th>
<th>90</th>
</tr>
</thead>
<tbody>
<tr>
<td>TT</td>
<td>38.21</td>
<td>38.58</td>
<td>41.33</td>
<td>44.17</td>
<td>42.83</td>
</tr>
<tr>
<td>WF</td>
<td>33.14</td>
<td>33.14</td>
<td>34.29</td>
<td>35.86</td>
<td>35.12</td>
</tr>
</tbody>
</table>

### $2 \rightarrow 1$

<table>
<thead>
<tr>
<th>QFt</th>
<th>QF2</th>
<th>50</th>
<th>60</th>
<th>70</th>
<th>80</th>
<th>90</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>TT</td>
<td>42.71</td>
<td>41.37</td>
<td>38.49</td>
<td>39.20</td>
<td>38.61</td>
</tr>
<tr>
<td></td>
<td>WF</td>
<td>33.65</td>
<td>37.32</td>
<td>36.33</td>
<td>35.03</td>
<td>36.05</td>
</tr>
<tr>
<td>60</td>
<td>TT</td>
<td>42.08</td>
<td>42.95</td>
<td>41.33</td>
<td>39.08</td>
<td>40.70</td>
</tr>
<tr>
<td></td>
<td>WF</td>
<td>33.65</td>
<td>38.20</td>
<td>36.49</td>
<td>36.11</td>
<td>36.55</td>
</tr>
<tr>
<td>70</td>
<td>TT</td>
<td>39.23</td>
<td>42.25</td>
<td>42.98</td>
<td>42.16</td>
<td>43.57</td>
</tr>
<tr>
<td></td>
<td>WF</td>
<td>32.34</td>
<td>32.77</td>
<td>37.66</td>
<td>37.36</td>
<td>37.45</td>
</tr>
<tr>
<td>80</td>
<td>TT</td>
<td>36.99</td>
<td>38.93</td>
<td>41.93</td>
<td>44.04</td>
<td>43.82</td>
</tr>
<tr>
<td></td>
<td>WF</td>
<td>32.17</td>
<td>32.40</td>
<td>33.31</td>
<td>39.13</td>
<td>34.57</td>
</tr>
<tr>
<td>90</td>
<td>TT</td>
<td>36.17</td>
<td>36.73</td>
<td>39.22</td>
<td>44.16</td>
<td>40.85</td>
</tr>
<tr>
<td></td>
<td>WF</td>
<td>32.99</td>
<td>32.46</td>
<td>33.41</td>
<td>35.14</td>
<td>34.35</td>
</tr>
</tbody>
</table>
Visual comparison

Starting FSD histograms: 2 compressions (75, 90). Target FSD histograms: 1 compression (70).
Final discussions and future work

- Extension to other distortion measures based on visual perception, like WPSNR or SSIM
- Joint analysis with the problem of DCT histogram modification
- Experiments on a wider datasets composed by images of different source and size
- Dealing with multiple compression steps
Thank you!
**References**

<table>
<thead>
<tr>
<th>Reference</th>
<th>Title</th>
</tr>
</thead>
</table>