Fast and Reliable Resampling Detection by Spectral Analysis of Fixed Linear Predictor Residue

10th ACM Workshop on Multimedia and Security

Matthias Kirchner
matthias.kirchner@inf.tu-dresden.de

Image Forensics, Resampling Detection

Image Forensics
use image statistics for identification of source device or detection of manipulations

Resampling Detection

- image manipulations often rely on geometric transformations (scaling, rotation, ... of images or parts thereof)
- resampling to a new image grid; involves an interpolation step
- interpolation introduces specific artefacts in the processed image which can be detected by the analysis of predictor residue
Interpolation Artefacts

- example: doubling of image size (linear interpolation)

\[ \tilde{I}_{5,5} = \frac{1}{2}(\tilde{I}_{5,5} + \tilde{I}_{7,5}) \]
\[ \tilde{I}_{5,6} = \frac{1}{2}(\tilde{I}_{5,5} + \tilde{I}_{5,7}) \]
\[ \tilde{I}_{6,5} = \frac{1}{4}(\tilde{I}_{5,5} + \tilde{I}_{7,5} + \tilde{I}_{5,7} + \tilde{I}_{7,7}) \]

- interpolation causes **systematic dependencies** between groups of neighbouring pixels
Resampling Detection

**p-map**: measure each pixel's probability for being correlated to its neighbours

Resampling causes **periodic pattern** in the p-map and **characteristic peaks** in the p-map's DFT
Detection Scheme

[Popescu & Farid, 2005]
Detection Scheme

[Popescu & Farid, 2005]

$$I(\omega t') = \sum_{\forall t \in \mathbb{Z}} h(\omega t' - t) I(t)$$
Detection Scheme
[Popescu & Farid, 2005]

\[
e(\omega t') = I(\omega t') - \sum_{k=-K}^{K} \alpha_k I(\omega t' + \omega k)
\]

\( (\alpha_0 := 0) \)

p-map → periodic artifacts

\[\mathcal{F}(\cdot)\]

p-map’s DFT → characteristic peaks
Detection Scheme
[Popescu & Farid, 2005]

\[ e(\omega t') = I(\omega t') - \sum_{k=-K}^{K} \alpha_k I(\omega t' + \omega k) \]

which \( \bar{\alpha} \)?

\( \mathcal{F}(\cdot) \)

p-map’s DFT \( \rightarrow \) characteristic peaks

p-map \( \rightarrow \) periodic artifacts


Fast and Reliable Resampling Detection
EM estimate of $\alpha$

iterative weighted least squares (WLS) procedure

$\alpha^{[0]} \rightarrow$ calculate $\vec{e}$ and p-map for current estimate $\alpha^{[i]}$

p-map

$\rightarrow$ compute $\alpha^{[i+1]}$

by minimizing

$\sum p(\omega t') e(\omega t')^2$

$\|\alpha^{[i+1]} - \alpha^{[i]}\| < \epsilon$ ?

$\rightarrow$ no

$\rightarrow$ yes

$\alpha = \alpha^{[i+1]}$
Analysis of Predictor Residue

Prediction error is the origin of periodic artefacts in the p-map of a resampled signal.

\[ e(\omega t') = I(\omega t') - \sum_{k=-K}^{K} \alpha_k I(\omega t' + \omega k) \]

Alternatively, we can write

\[ e(\omega t') = \sum_{k=-K}^{K} \beta_k \sum_{\forall t \in \mathbb{Z}} h(\omega t' + \omega k - t) I(t) \]

\[ \beta_k = \begin{cases} 
1 & \text{for } k = 0 \\
-\alpha_k & \text{else}
\end{cases} \]
Analysis of Predictor Residue

prediction error is the origin of periodic artefacts in the p-map of a resampled signal

\[ e(\omega t') = I(\omega t') - \sum_{k=-K}^{K} \alpha_k I(\omega t' + \omega k) \]

alternatively we can write

\[ e(\omega t') = \sum_{k=-K}^{K} \beta_k \sum_{t \in \mathbb{Z}} h(\omega t' + \omega k - t) I(t) \]

\[ = \sum_{t \in \mathbb{Z}} \varphi_\omega(\omega t' - t) I(t) \]

\[ \beta_k = \begin{cases} 
1 & \text{for } k = 0 \\
-\alpha_k & \text{else} 
\end{cases} \]

\[ \varphi_\omega(x) = \sum_{k=-K}^{K} \beta_k h(x + \omega k) \]
Variance of Predictor Residue

Theorem

For wide sense stationary signals \( I(t), t \in \mathbb{Z}, \) with variance \( \text{Var}[I(t)] > 0 \) and arbitrary prediction weights \( \tilde{\alpha} \neq 0 \) it holds \( \forall x \in \mathbb{R} \) that \( \text{Var}[e(x)] = \text{Var}[e(x + 1)] \).
Variance of Predictor Residue

Theorem
For wide sense stationary signals $I(t), t \in \mathbb{Z}$, with variance $\text{Var}[I(t)] > 0$ and arbitrary prediction weights $\vec{\alpha} \neq 0$ it holds $\forall x \in \mathbb{R}$ that $\text{Var}[e(x)] = \text{Var}[e(x + 1)]$.

Why is $\text{Var}[e(x)]$ of interest?

- For a zero mean (generalized) Gaussian prediction error, variance governs the p-map’s appearance.
- Periodicity in variance can serve as a simple model to explain periodic pattern in the p-map of resampled signals.
Variance of Predictor Residue

Theorem
For wide sense stationary signals $I(t), t \in \mathbb{Z}$, with variance $\text{Var}[I(t)] > 0$ and arbitrary prediction weights $\vec{\alpha} \neq 0$ it holds $\forall x \in \mathbb{R}$ that $\text{Var}[e(x)] = \text{Var}[e(x + 1)]$.

Why is $\text{Var}[e(x)]$ of interest?

- For a zero mean (generalized) Gaussian prediction error, variance governs the p-map’s appearance.
- Periodicity in variance can serve as a simple model to explain periodic pattern in the p-map of resampled signals.

Extension to 2D
$\text{Var}[e(x_1, x_2)] = \text{Var}[e(x_1 + 1, x_2)] = \text{Var}[e(x_1, x_2 + 1)] = \text{Var}[e(x_1 + 1, x_2 + 1)]$
Periodic Variance

\[ \alpha = (0.5, 0, 0.5) \quad \omega = 0.3 \quad \omega = 0.5 \quad \omega = 0.9 \]
Periodic Variance

\[\tilde{\alpha} = (0.5, 0, 0.5)\]

\[\omega = 0.3 \quad \omega = 0.5 \quad \omega = 0.9\]

Resampling implies sampling of \(\text{Var}[e(x)]\) with sampling frequency \(f_s = \omega^{-1}\)
Periodic Variance

\[ \bar{\alpha} = (0.5, 0, 0.5) \quad \omega = 0.3 \quad \omega = 0.5 \quad \omega = 0.9 \]

resampling implies sampling of \( \text{Var}[e(x)] \) with sampling frequency \( f_s = \omega^{-1} \)
Implications

- complex and time-consuming **EM estimation** of scalar weights $\bar{\alpha}$ seems unnecessary
- in the most simple case, resampling detection involves only **linear filtering** with fixed coefficients
- using the simplified model, we can employ basic signal processing primitives to derive an **exact formulation** on how a specific transformation will influence the **position of characteristic peaks**
Fast Detection of Resampling

- drop computationally demanding EM estimation of the scalar weights
- compute prediction residue by applying a linear filter to the image under investigation
- calculate p-map as \( p = \lambda \exp \left( -\frac{|e|^\tau}{\sigma} \right) \)

<table>
<thead>
<tr>
<th>( \omega^{-1} )</th>
<th>( K = 1 )</th>
<th>( K = 2 )</th>
<th>( K = 3 )</th>
<th>fast</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.75</td>
<td>6.0 s</td>
<td>20.2 s</td>
<td>52.9 s</td>
<td>0.1 s</td>
</tr>
<tr>
<td>orig.</td>
<td>7.6 s</td>
<td>19.9 s</td>
<td>45.8 s</td>
<td>0.1 s</td>
</tr>
<tr>
<td>1.5</td>
<td>6.5 s</td>
<td>16.8 s</td>
<td>38.6 s</td>
<td>0.1 s</td>
</tr>
</tbody>
</table>
Which Filter Coefficients?

\[
\vec{\alpha} = \begin{bmatrix}
\alpha_1 & \alpha_4 & \alpha_7 \\
\alpha_2 & 0 & \alpha_8 \\
\alpha_3 & \alpha_6 & \alpha_9 
\end{bmatrix}
\]

\[
\vec{\alpha}^* = \begin{bmatrix}
\frac{-1}{4} & \frac{1}{2} & \frac{-1}{4} \\
\frac{1}{2} & 0 & \frac{1}{2} \\
\frac{-1}{4} & \frac{1}{2} & \frac{-1}{4}
\end{bmatrix}
\]
Detection Results

- in general very high detectability for a large range of resampling parameters

Detection rate

FAR < 3 %

original

scaling $S = \omega^{-1}$


Fast and Reliable Resampling Detection
Detection Results

- in general very high detectability for a large range of resampling parameters
- accelerated detector performs equally well except for strong downscaling
Detection Results

- in general very high detectability for a large range of resampling parameters
- accelerated detector performs equally well except for strong downscaling
Locating Characteristic Resampling Peaks (2D)

$V = \mathcal{F}(\text{Var}[e(x)])$ has distinct peaks

$f_{p_1} = (1, 0), f_{p_2} = (0, 1), f_{p_3} = (1, 1)$
Locating Characteristic Resampling Peaks (2D)

A = \[
\begin{bmatrix}
\cos \frac{\pi}{4} & \sin \frac{\pi}{4} \\
-\sin \frac{\pi}{4} & \cos \frac{\pi}{4}
\end{bmatrix}
\]

45° rotation:

V = \mathcal{F}(\text{Var}[e(\mathbf{x})]) \text{ has distinct peaks } \mathbf{f}_{p_1} = (1, 0), \mathbf{f}_{p_2} = (0, 1), \mathbf{f}_{p_3} = (1, 1)

affine transformation A:

\mathcal{F}(\text{Var}[e(A\mathbf{x})]) = V((A')^{-1}\mathbf{f})
Locating Characteristic Resampling Peaks (2D)

\[ A = \begin{bmatrix} \cos \frac{\pi}{4} & \sin \frac{\pi}{4} \\ -\sin \frac{\pi}{4} & \cos \frac{\pi}{4} \end{bmatrix} \]

45° rotation:

\[ V = \mathcal{F}(\text{Var}[e(x)]) \] has distinct peaks

\[ f_{p1} = (1, 0), f_{p2} = (0, 1), f_{p3} = (1, 1) \]

affine transformation \( A \):

\[ \mathcal{F}(\text{Var}[e(Ax)]) = V((A')^{-1}f) \]

sampling of \( \text{Var}[Ae(x)] \) implies a periodic extension of the original spectrum with the base band

\[ |f| \leq (0.5, 0.5) \]
Locating Characteristic Resampling Peaks (2D)

Examples

45° rotation

predicted peaks: (0.293, 0.293)

10% vertical shear

predicted peaks: (0, 0.1)
Conclusion and Future Work

• rather complex and computationally demanding state-of-the-art resampling detector can benefit from a deeper analysis of resampling artefacts
  – simple model to explain periodic pattern in the p-map
  – tremendous performance speed-up by dropping unnecessary procedures in detection procedure while keeping detection reliable
  – derivation of the affine transformation parameters (within a well-defined ambiguity)

• future work includes
  – quantitative evaluation of our model with respect to the correctness of predicted peak positions and
  – analysis of our recent attacks on resampling detection in terms of the presented model
Thanks for your attention

Questions?

Matthias Kirchner
matthias.kirchner@inf.tu-dresden.de

The author gratefully receives a doctorate scholarship from Deutsche Telekom Stiftung.

Ambiguous Peak Position (1D)

different inverse resampling rates $\omega$ result in the same peak position