Image Fraud

Ariel Shin

March 23, 2018

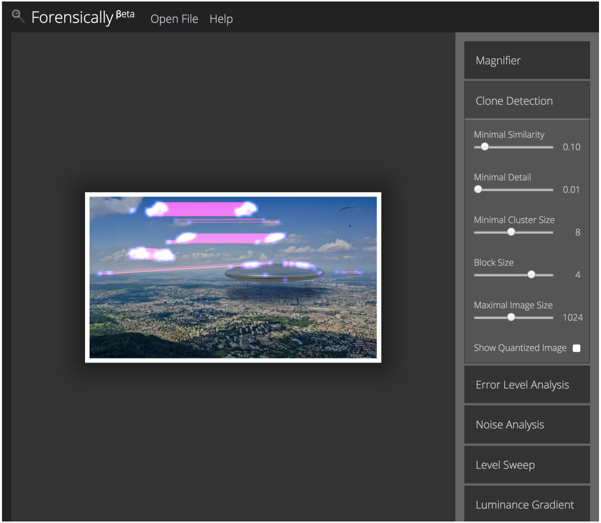
Copy-Paste Fraud

original

tampered

Existing Tools: Forensically

Based on an algorithm from "Detection of Copy-Move Forgery in Digital Images" by Fridrich et al.



Clone Detection Parameters

- Minimal Similarity
- Minimal Detail
- Minimal Cluster Size
- Block Size

Forensically's algorithm

Create a Table
Move a window over the image, for each position of the window
    Use all of the pixels in the window as a key
    If the key is already in the table
        We found a clone! Mark it.
    Else
        Add the key to the table

"Detection of Copy-Move Forgery in Digital Images" by Fridrich et al

Provides Four Methods

- Exhaustive search
- Autocorrelation
- Exact Match
- Robust Match

Exhaustive Search

The image and its circularly shifted version are overlaid looking for closely matching image segments.

Autocorrelation

- The original and copied segments will introduce peaks in the autocorrelation for the shifts that correspond to the copied-moved segments

Exact Match

1. User specifies minimal size of the segment. Suppose the segment is BxB pixels.
2. Copy pixel values from each block to a vector of size B^2
3. Search for two vectors that are identical

Robust Match

- Similar to exact match except we do not order and match the pixel representation of the blocks but their robust representation that consists of quantized DCT coefficients.

Principle Component Analysis

- Non-parametric method of extracting relevant information from confusing data sets
- If there are n dimensions in original data
  - Calculate n eigenvectors and eigen values
  - Choose the first p eigenvectors, based on their eignvalues
- Final dataset has only p dimensions

Popescu and Farid's PCA Anglorithm

1. Let N be the total number of pixels in a grayscale or color image
2. Initialize the parameters:
   - b: number of pixels per block ($\sqrt{b} \times \sqrt{b}$ pixels in dimension) – there are $N_b = (\sqrt{N} - \sqrt{b} + 1)^2$ such blocks
   - ϵ: fraction of the ignored variance along the principal axes
   - Q: number of quantization bins
   - N_n: number of neighboring rows to search in the lexicographically sorted matrix
   - N_f: minimum frequency threshold
   - N_d: minimum offset threshold
3. Using PCA, compute the new N_t-dimensional representation, $\vec{a}_i$, $i = 1, \ldots, N_b$, of each b pixel image block (for color images: (1) analyze each color channel separately; or (2) build a single color block of size $3b$ pixels). The value of N_t is chosen to satisfy: $1 - \epsilon = \frac{\sum_{i=1}^{N_t} \lambda_i}{\sum_{i=1}^{b} \lambda_i}$, where λ_i are the eigenvalues as computed by the PCA.
4. Build a $N_b \times b$ matrix whose rows are given by the component-wise quantized coordinates: $\lfloor \vec{a}_i / Q \rfloor$.
5. Sort the rows of the above matrix in lexicographic order to yield a matrix S. Let $\vec{s}_i$ denote the rows of S, and let (x_i, y_i) denote the position of the block's image coordinates (top-left corner) that corresponds to $\vec{s}_i$.
6. For every pair of rows $\vec{s}_i$ and $\vec{s}_j$ from S such that $|i - j| < N_n$, place the pair of coordinates (x_i, y_i) and (x_j, y_j) onto a list.
7. For all elements in this list, compute their offsets, defined as:

$$\begin{array}{rl} (x_i - x_j, y_i - y_j) & \text{if} \quad x_i - x_j > 0 \\ (x_j - x_i, y_i - y_j) & \text{if} \quad x_i - x_j < 0 \\ (0, |y_i - y_j|) & \text{if} \quad x_i = x_j \end{array}$$

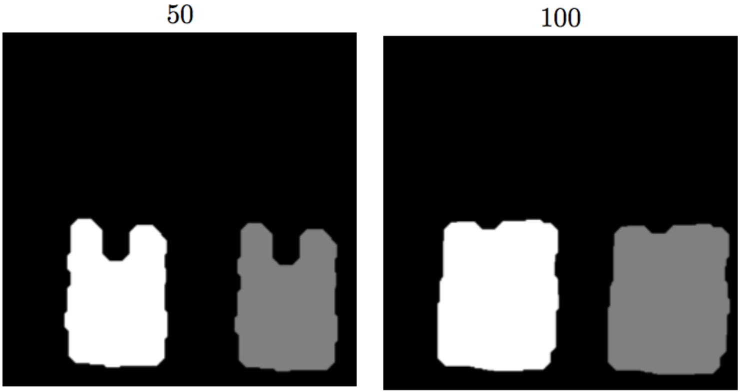
8. Discard all pairs of coordinates with an offset frequency less than N_f.
9. Discard all pairs whose offset magnitude, $\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$, is less than N_d.
10. From the remaining pairs of blocks build a duplication map by constructing a zero image of the same size as the original, and coloring all pixels in a duplicated region with a unique grayscale intensity value.

PCA Results

original

tampered

PCA Results



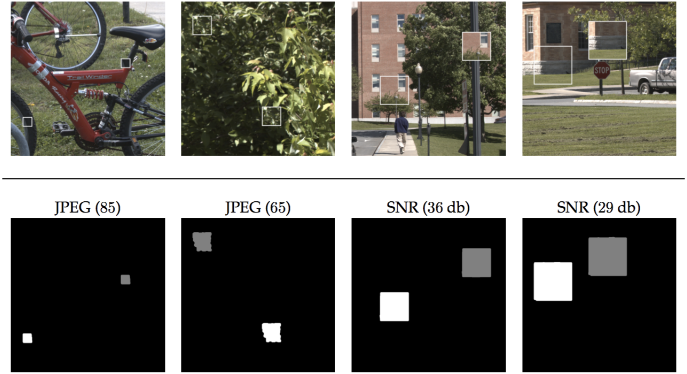
PCA Results

original

tampered

PCA Results

PCA Results



Works Cited

- Fridrich, A. Jessica, B. David Soukal, and A. Jan Luk. "Detection of copy-move forgery in digital images." in Proceedings of Digital Forensic Research Workshop. 2003.
- Popescu, A. C., and H. Farid. "Exposing digital forgeries by detecting duplicated image regions. Department Computer Science, Dartmouth College, Technology Report TR2004-515." (2004).
- Wagner, Jonas. Forensically, Photo Forensics for the Web. 29a.Ch by Jonas Wagner, 16 Aug. 2015, 29a.ch/2015/08/16/forensically-photo-forensics-for-the-web.