TECHNISCHE UNIVERSITÄT DRESDEN

Digitale Bildforensik – Spuren in Digitalbildern

Matthias Kirchner

23C3, Berlin, 27. Dezember 2006

Die Fotografie ist unwiderlegbar. Sie ist gar nicht zu schlagen. [...] Der Zeichner kann sich etwas ausdenken. Der Fotograf nicht.

Kurt Tucholsky (1925)

Historische ...

... und aktuelle Bildmanipulationen

Digitale Bilder – (K)ein Abbild der Wirklichkeit?

- In Zeiten von Digitalkameras und ausgereifter Bearbeitungs-software kann heute nahezu jeder digitale oder digitalisierte Bilder und deren Aussage manipulieren
- Nachbearbeitung alltäglich
- "Intelligente" Kameras
- Speziell ausgerichtete Software

  (z.B. Touristremover, "Face Beautification", ...)

⇒ Bildmanipulationen allgemein akzeptiert?

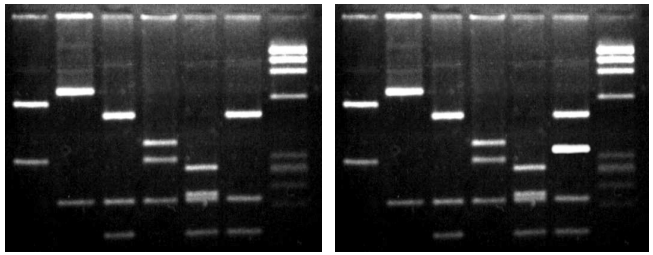
http://www.funpic.hu/funblog/allatok/allatok.html

Beispiel: "Face Beautification"

http://www.spiegel.de/international/spiegel/0,1518,439709,00.html

Schutz der Bildauthentizität

- **In einer Mediengesellschaft unverzichtbar**
- Aktive Verfahren (z.B. Einbetten eines Wasserzeichens) sind in vielen Fällen praktisch nicht einsetzbar:
  - Journalismus
  - Wissenschaftliche Veröffentlichungen
  - Blogs, ...



Digitale Bildforensik

Original unbekannt

Digitales Bild

Statistische Analyse

Modell:
- Bildinhalt oder
- Digitalisierungstechnik

Identifikation der Bildquelle

Detektion von Bildmanipulationen

Identifikation der Bildquelle

- Computergeneriert (CG) vs. Natürliche Fotografie

http://www.fakeorfoto.com

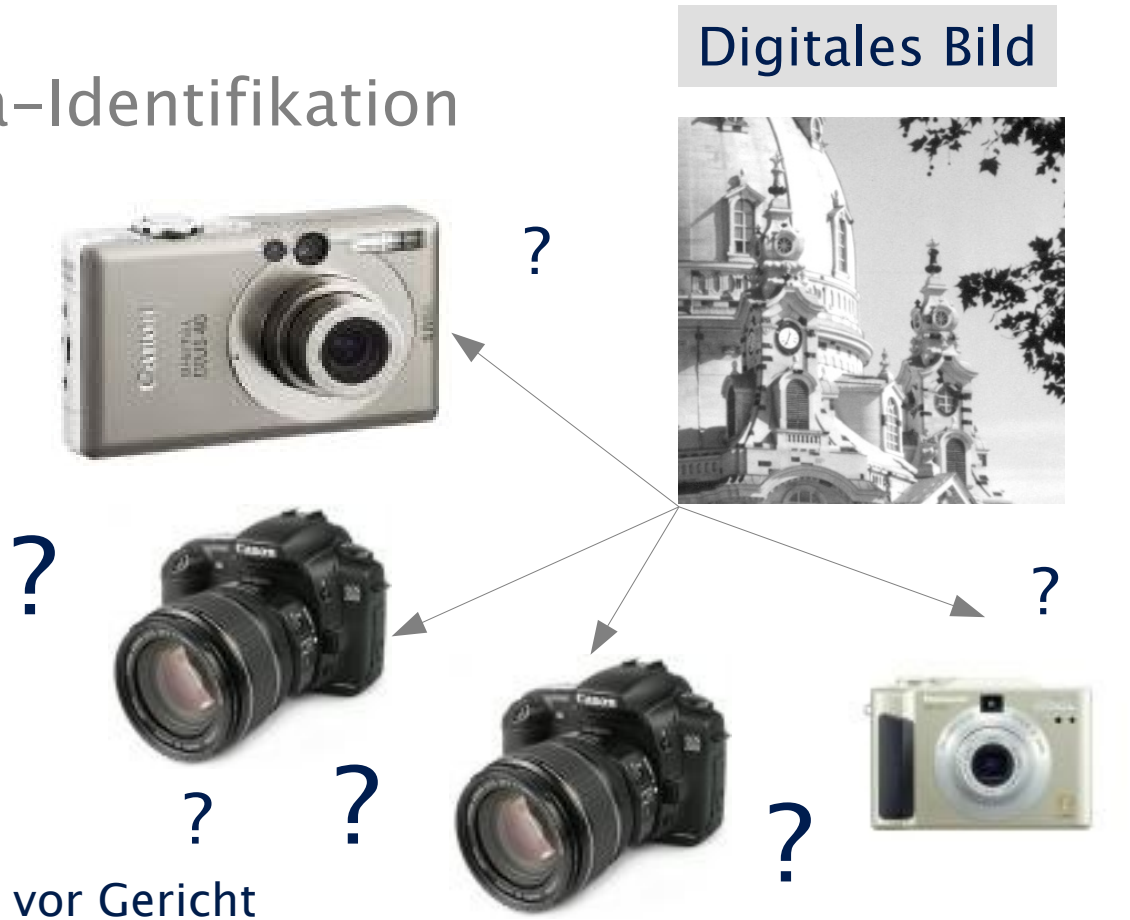
- Digitalkamera-Identifikation

Detektion von Manipulationen

- Meist auf spezielle Manipulationen ausgerichtete Verfahren, z.B.:
  - Fotomontagen
  - Copy & Paste
  - Affine Transformationen (Skalierung, Verzerrung, Rotation)

- Hinweise auf Manipulationen, z.B.:
  - Lokales Rauschen
  - Doppelte JPEG-Kompression

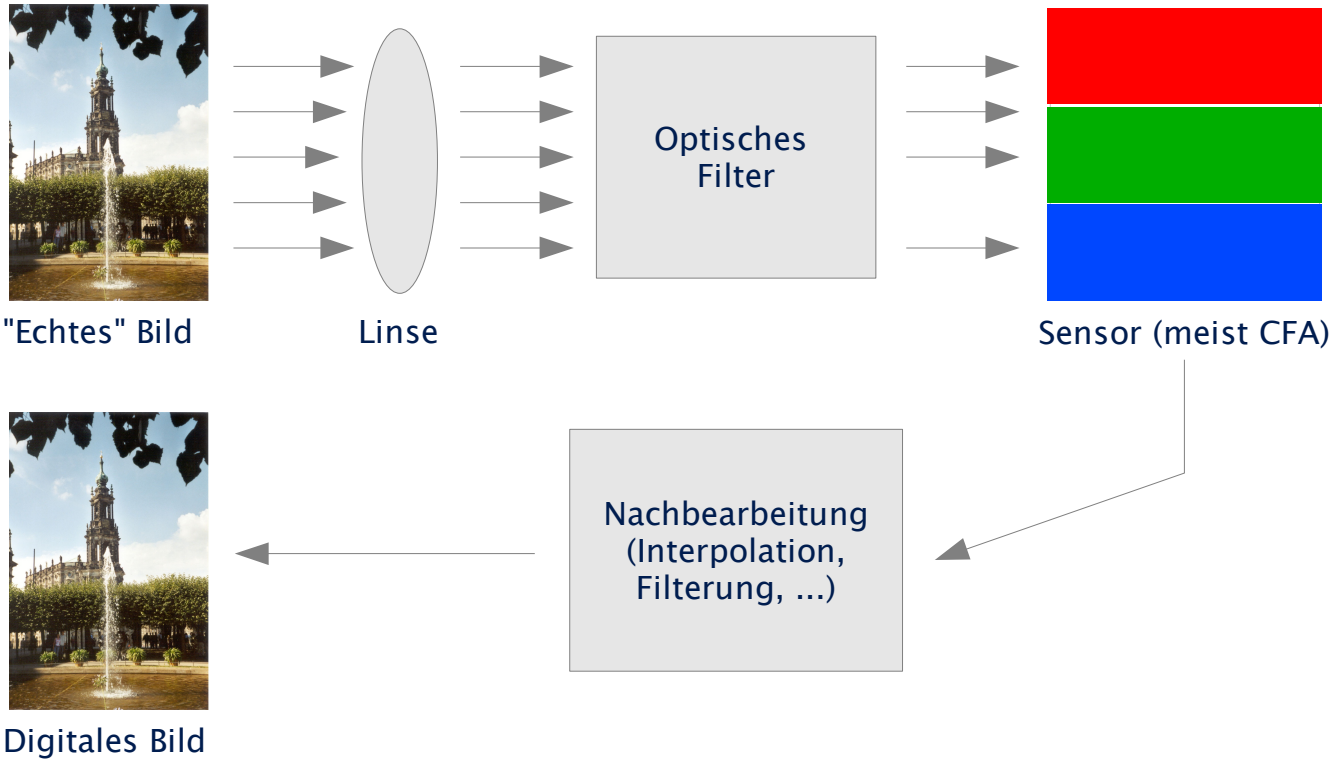
Digitalkamera-Identifikation



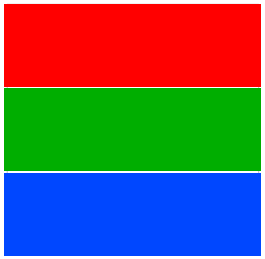
- Anwendung:

  z.B. Beweisbilder vor Gericht

Signalbearbeitung in einer Digitalkamera



Rauschen



Sensor

Rauschen

- Schrotrauschen
- Dunkelstrom
- Photo Response Non-Uniformity

=

Pattern Noise

Fixed Pattern Noise

- Durch Dunkelströme
- Additiver Charakter
- Korrektur durch X=X-D

PRNU

- Pixel-zu-Pixel Ungleichheiten
- Hochfrequenter Charakter
- Nutzung zur Kamera-Identifikation

Kamera-Identifikation mit Rauschmustern
(Lukáš et al., 2005)

Annahme: PRNU ist normalverteiltes Rauschen

⇒ Digitalkamera-Identifikation = Detektion eines "Wasserzeichens"

↪ Extraktion mit geeignetem Rauschfilter *F*, Korrelationsmaß

$$\rho_C(Y) = corr(Y - F(Y), P_C) = \frac{(Y - F(Y) - E[Y - F(Y)])(P_C - E[P_C])}{\|Y - F(Y) - E[Y - F(Y)]\| \|P_C - E[P_C]\|}$$

F … Rauschfilter
Y … zu untersuchendes Bild
$Y - F(Y)$ … Approximation des gesuchten Pattern Noise
P_C … Referenz-Rauschmuster einer Kamera C

Anwendungsszenarien

- Wähle aus einer Menge von Kameras die, mit welcher am wahrscheinlichsten ein bestimmtes Bild aufgenommen wurde

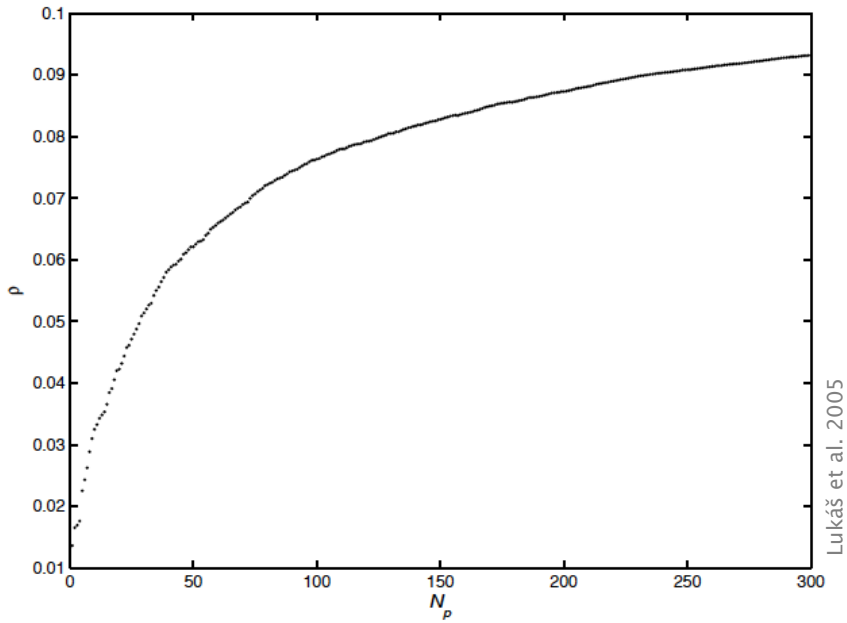
  ⇒ Kamera, deren Referenz-Rauschmuster die höchste Korrelation mit dem extrahierten Rauschmuster aufweist

- Bewerte die Aussage, dass mit einer speziellen Kamera ein bestimmtes Bild aufgenommen wurde

  ⇒ Grenzwerte nötig, welche die Zuverlässigkeit der Entscheidung garantieren

Referenz-Rauschmuster

- Kann durch Mittelung des Rauschens aus mehreren Bildern gewonnen werden



Abhängig von der Anzahl der verwendeten Bilder (mindestens 50)

Untersuchte Kameras

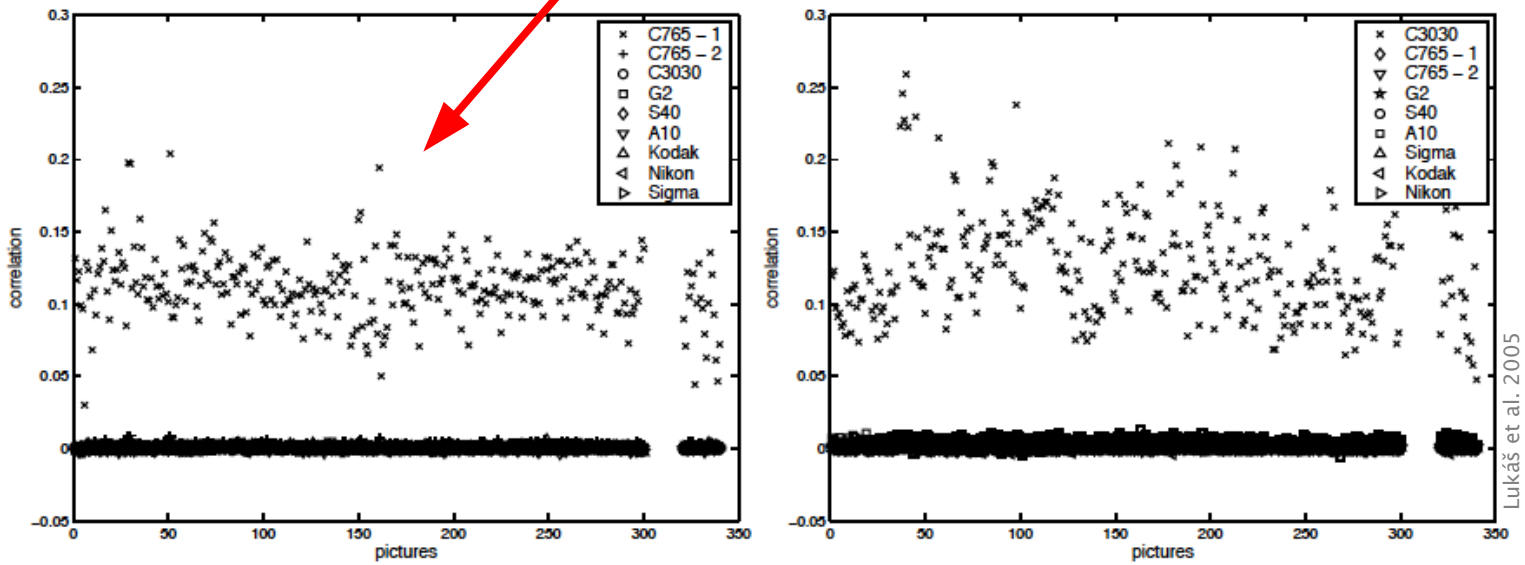
Zwei Kameras des gleichen Typs

Camera brand	Sensor	Maximal resolution	Image format
Canon PowerShot A10	1/2.7-inch CCD	1280×960	JPEG
Canon PowerShot G2	1/1.8-inch CCD	2272×1704	CRW, JPEG
Canon PowerShot S40	1/1.8-inch CCD	2272×1704	CRW, JPEG
Kodak DC290		1792×1200	TIFF, JPEG
Olympus Camedia C765 UZ - 1	1/2.5-inch CCD	2288×1712	TIFF, JPEG
Olympus Camedia C765 UZ - 2	1/2.5-inch CCD	2288×1712	TIFF, JPEG
Nikon D100	23.7×15.5 mm Nikon DX CCD	3008×2000	NEF-RAW, TIFF, JPEG
Sigma SD9	20.7×13.8 mm CMOS-Foveon X3	2268×1512	X3F-RAW
Olympus Camedia C3030	1/1.8-inch CCD	2048×1536	TIFF, JPEG

Lukáš et al. 2005

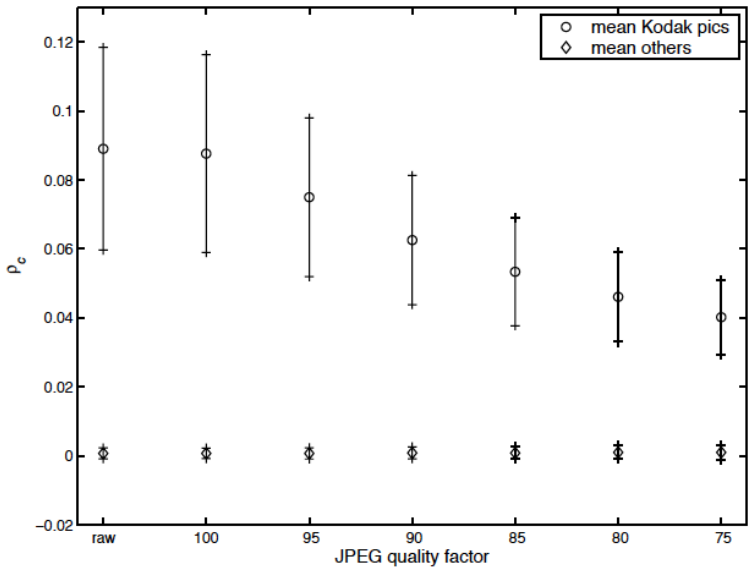
Ergebnisse (unkomprimierte Bilder)

Kameras gleichen Typs unterscheidbar

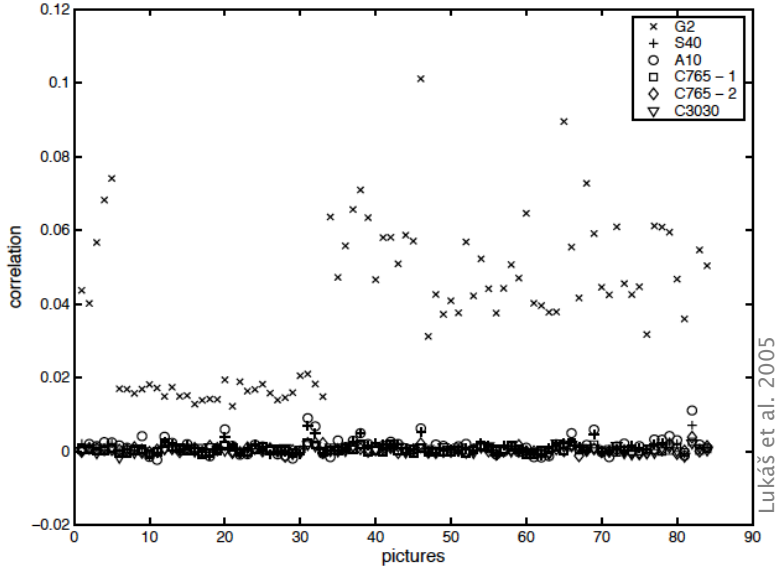


(Bei nicht passenden Referenz-Mustern wird das größere Muster beschnitten)

Ergebnisse (JPEG-Kompression, Re-Sampling)



JPEG-Kompression verkleinert Mittelwert u. Varianz des Korrelationsmaßes



Identifikation auch nach Bildvergrößerung und JPEG-Kompression möglich

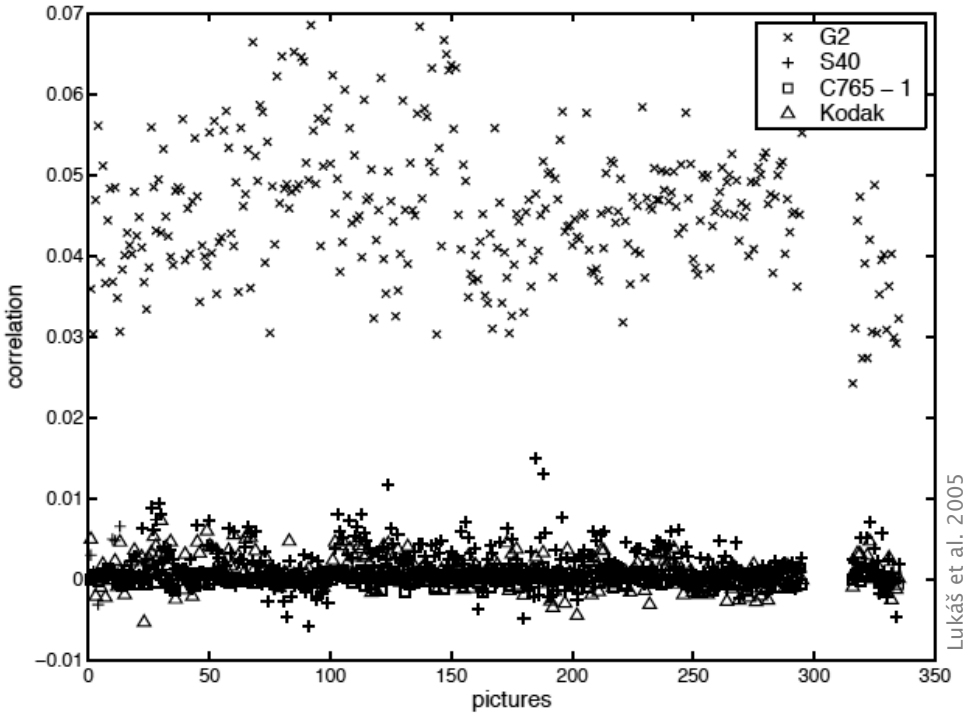
Robustheit (1)

Mögliche Angriffe

- Entfernen des Kamera-typischen Rauschmusters
  - De-Synchronisation (klassisches Wasserzeichen-Problem)
  - Wenn Referenzmuster oder Rauschfilter bekannt

- Austauschen des Kamera-typischen Rauschmusters
  - Wenn betreffende Referenzmuster bekannt

Robustheit (2)



Lukáš et al. 2005

Richtiges Referenzmuster erzeugt auch nach Entfernung des Rauschens aus dem Bild (Rauschfilter) die höchste Korrelation

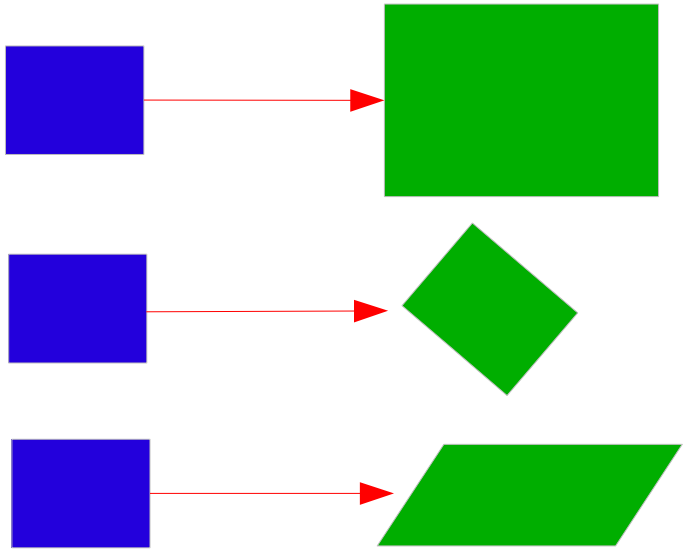
Weitere Verfahren

- Identifikation als Klassifikationsaufgabe (Kharrazi et al., 2004)
- Bestimung des verwendeten Algorithmus bei der CFA-Interpolation
  (Bayram et al., 2005; Popescu & Farid, 2005; Swaminathan et al., 2006)
- Analyse der JPEG-Quantisierungsmatrix (Farid, 2006)

Detektion von Re-Sampling
(Popescu & Farid, 2005)

Re-Sampling: im Allgemeinen nach jeder geometrischen Transformation des Bildes (oder von Bildteilen), z.B.:

- Skalierung
- Rotation
- Scherung



Beschreibbar als Koordinatentrans-formation
⇒ Umtastung auf ein neues Bildgitter

Re-Sampling eines 1D-Signals (1)

x[k] mit **m** Samples —Faktor p/q→ y[k'] mit **n** Samples

1) Up-Sampling

$$x_u[t'] = \begin{cases} x[k] & t' = p \cdot k \quad (k = 1,2,\ldots,m) \\ 0 & \text{sonst} \end{cases}$$

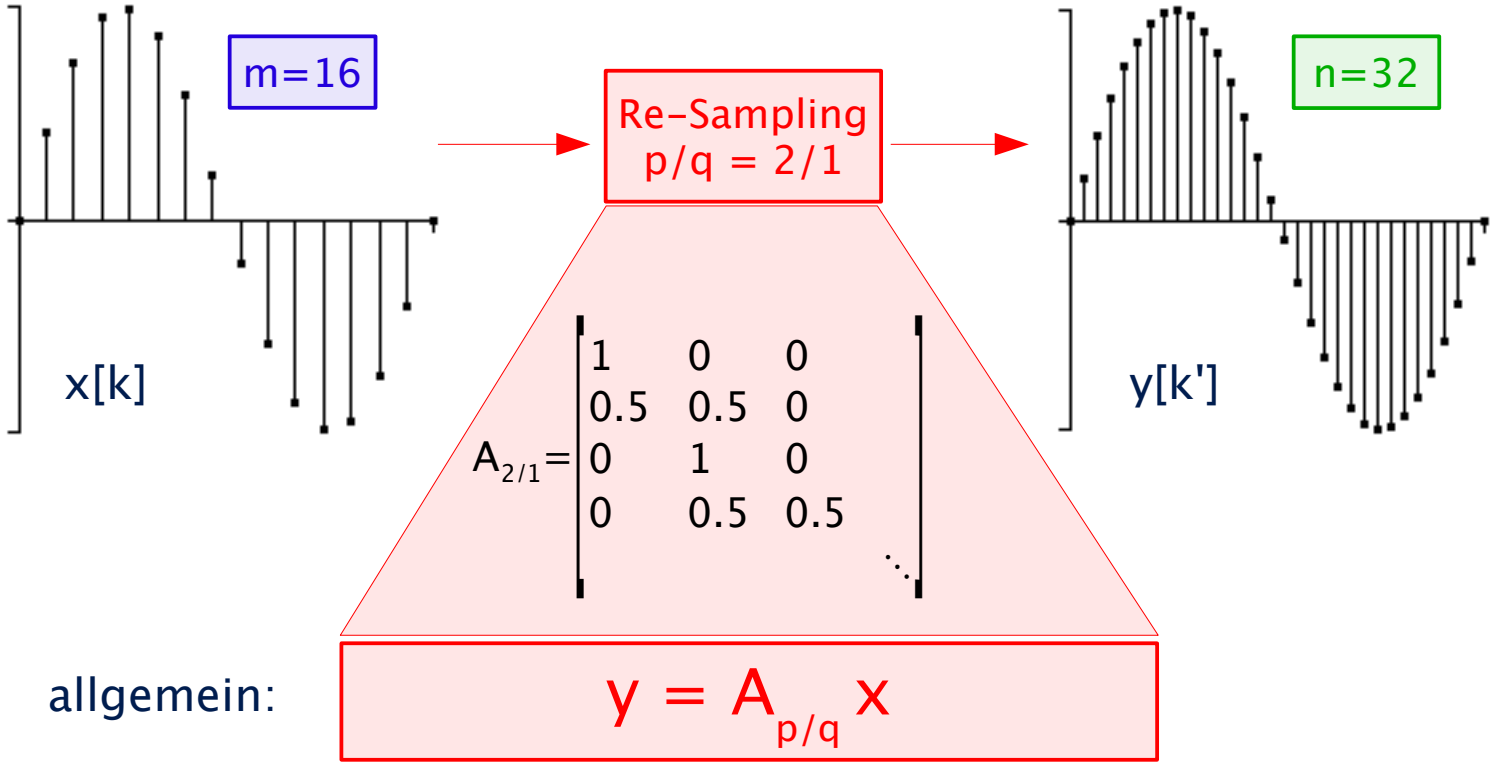
2) Interpolation

$$x_i[t'] = x_u[t'] * h[t']$$

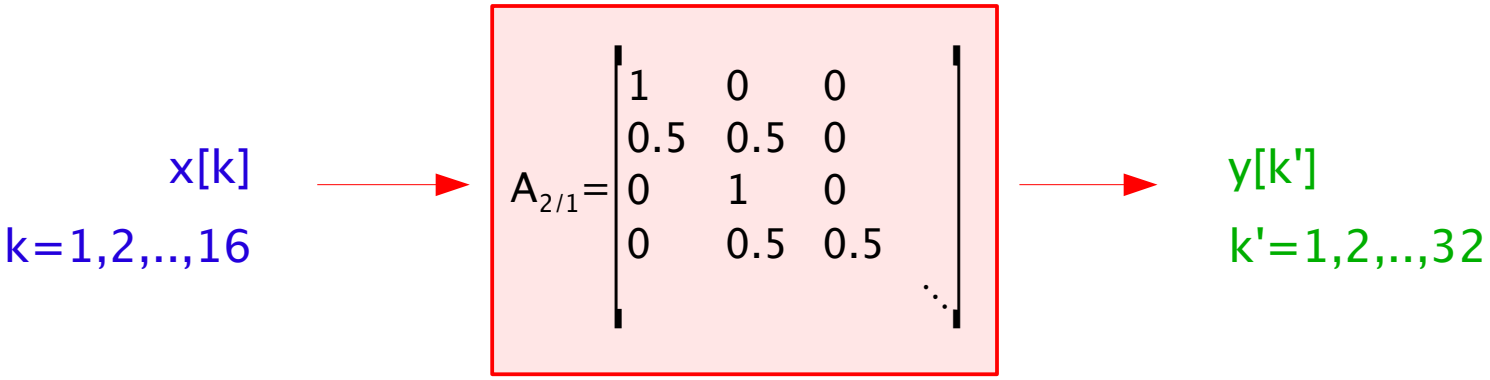
3) Down-Sampling

$$x_d[k'] \equiv y[k'] = x_i[q \cdot t'] \quad (k' = 1,2,\ldots,n)$$

Re-Sampling eines 1D-Signals (2)



Re-Sampling eines 1D-Signals (3)



$$y[2i-1] = x[i]$$

$$y[2i] = 0.5x[i] + 0.5x[i+1]$$

$$y[2i] = 0.5y[2i-1] + 0.5y[2i+1]$$

⇒ Korrelation zwischen benachbarten Samples

Re-Sampling eines 1D-Signals (4)

allgemein:

Abtastwerte korrelieren

Korrelierte Samples bekannt

Zeilen der Re-Sampling-Matrix korrelieren

$$y[i]=\sum_{k=-N}^{k=N}\alpha_k\cdot y[i+k] \quad (\alpha_0=0) \iff a_i=\sum_{k=-N}^{k=N}\alpha_k\cdot a_{i+k} \quad (\alpha_0=0)$$

Interpolations-schema bekannt

In der Regel aber weder noch bekannt

⇒ Expectation/Maximization (EM) Algorithmus zur Detektion von Re-Sampling

Expectation/Maximization

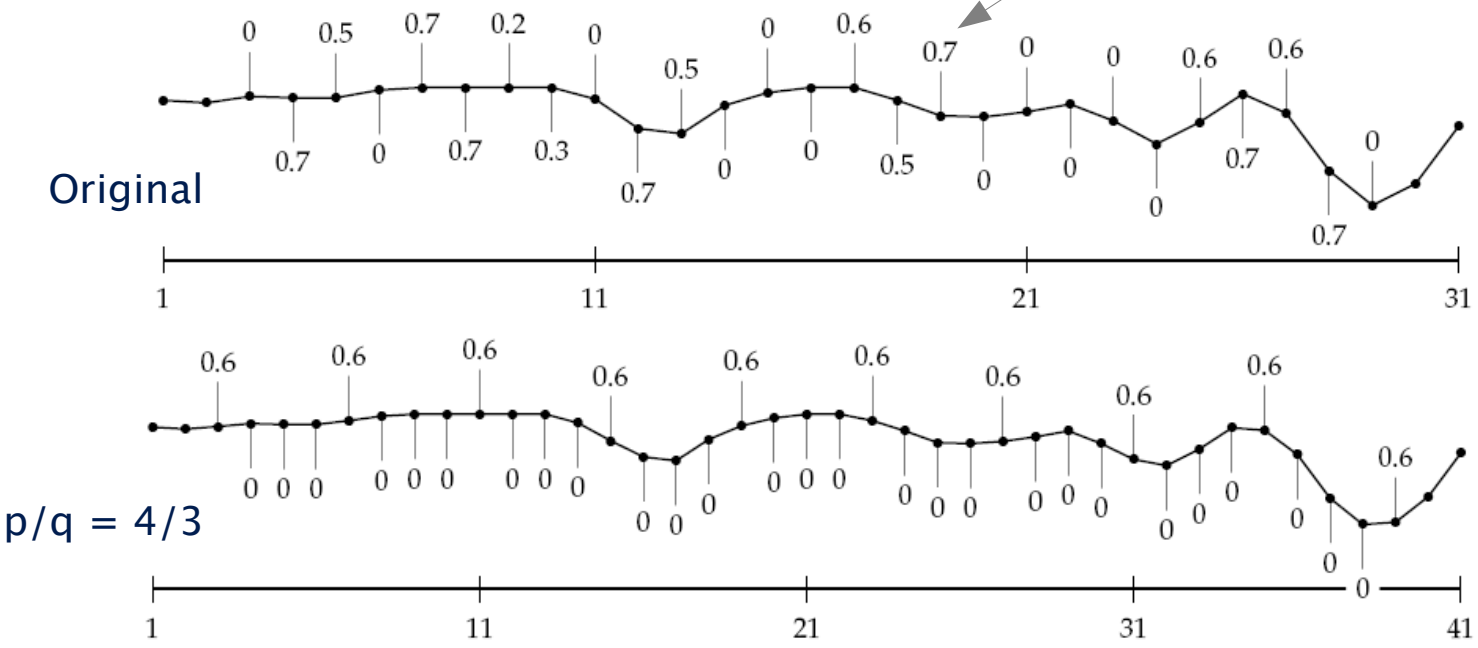
- Einteilung der Samples in zwei Klassen
  - $y[i] \in M_1$: y[i] korreliert mit seinen Nachbarn
  - $y[i] \in M_2$: y[i] korreliert nicht mit seinen Nachbarn

Iteration von E-Schritt und M-Schritt

- Expectation (E): Berechne Wahrscheinlichkeiten dafür, dass ein Sample zu M_1 gehört
- Maximization (M): Schätzen der Korrelationen zwischen den Samples (d.h. Schätzen der skalaren Gewichte α)

Detektion in 1D-Signalen

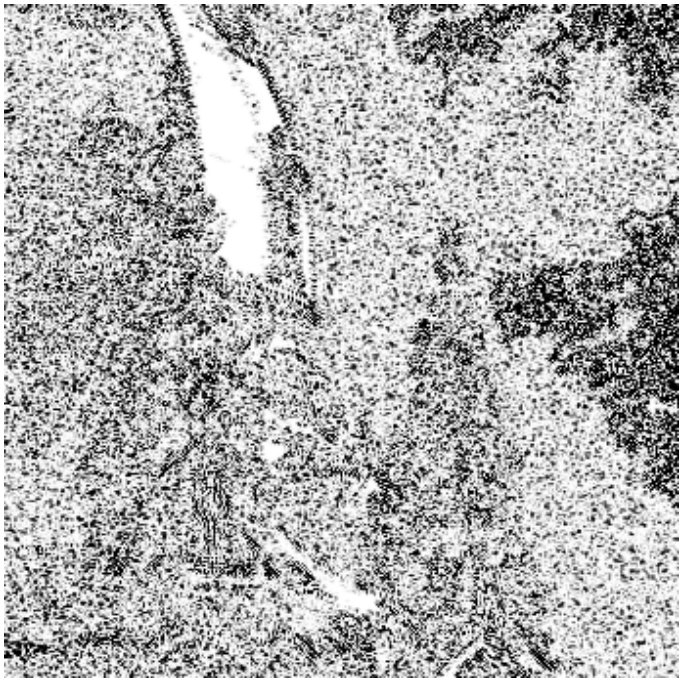
Wahrscheinlichkeit dafür, dass ein Sample mit seinen Nachbarn korreliert



Korrelationswahrscheinlichkeiten im umgetasteten Signal periodisch

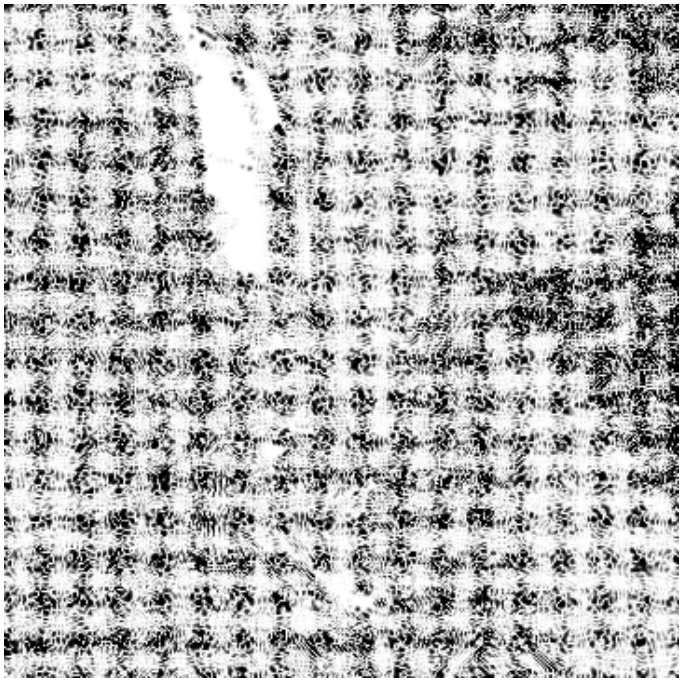
Detektion in Bildern (1)

Original



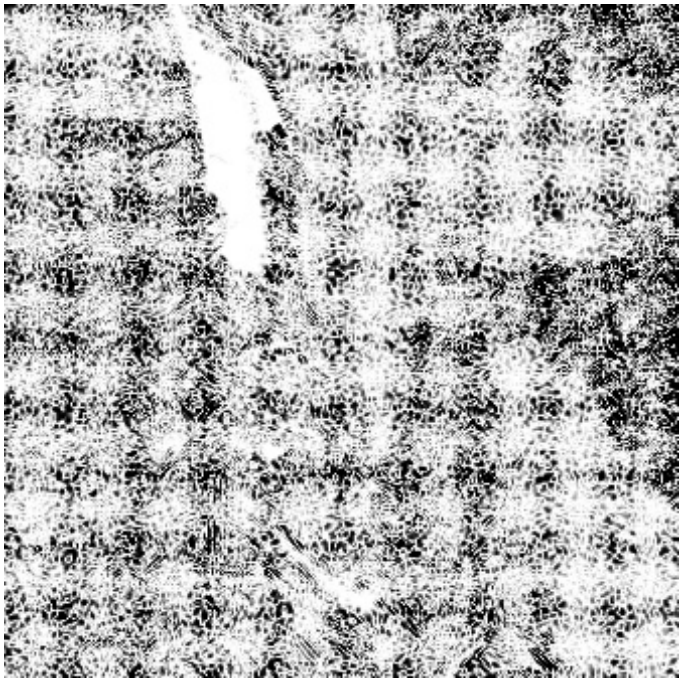
Detektion in Bildern (2)

Vergrößerung um 5%



Detektion in Bildern (3)

Verkleinerung um 2.5%



Detektion in Bildern (4)

Rotation um 5°

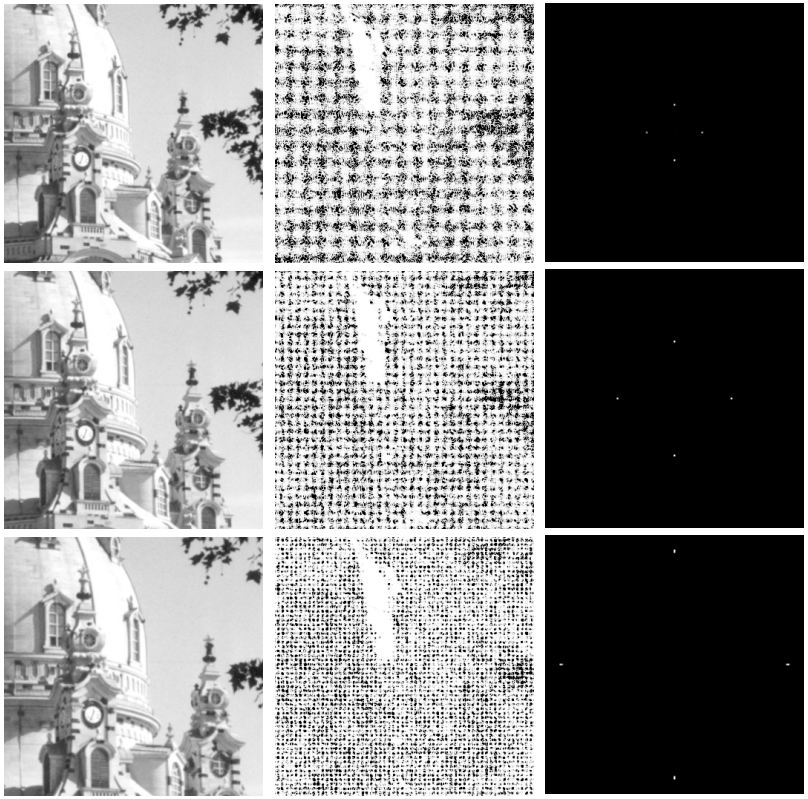
Detektion in Bildern (5)

Art des Re-Samplings erzeugt typisches Betragsspektrum

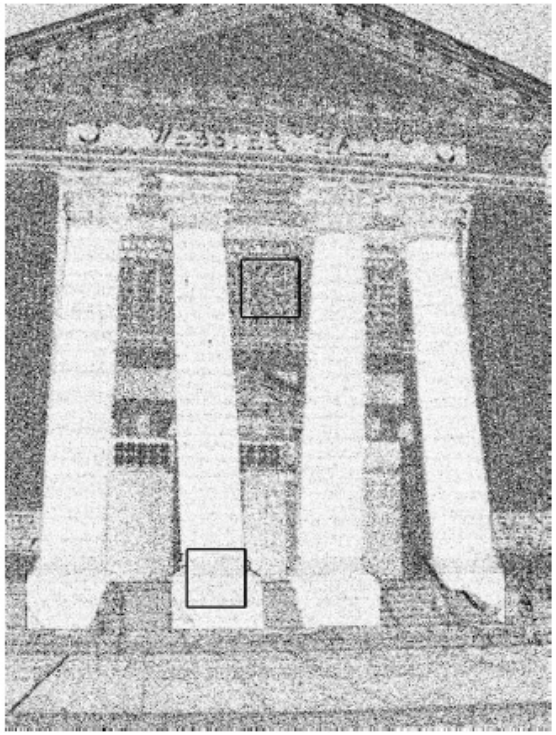
105%

110%

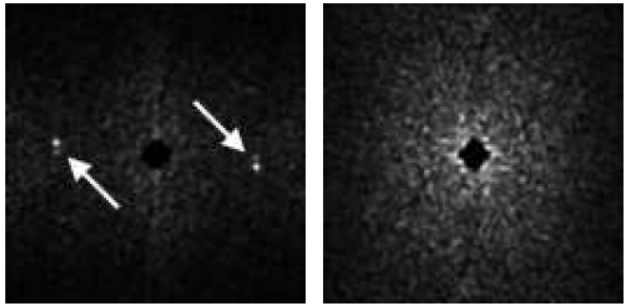
120%



Detektion lokaler Manipulationen



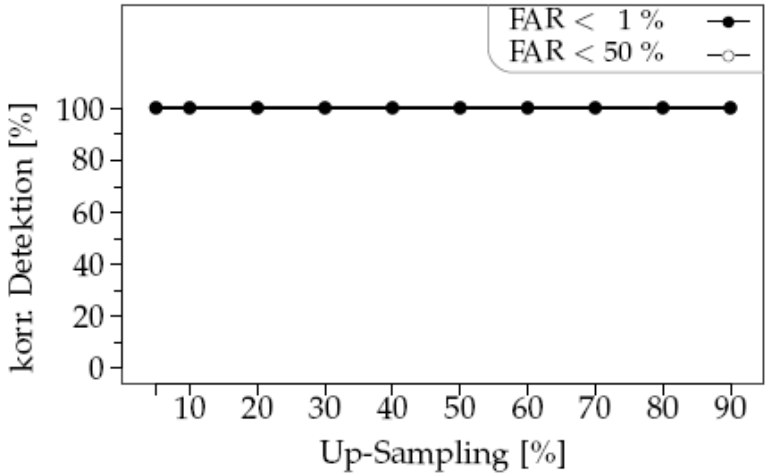
Popescu 2005

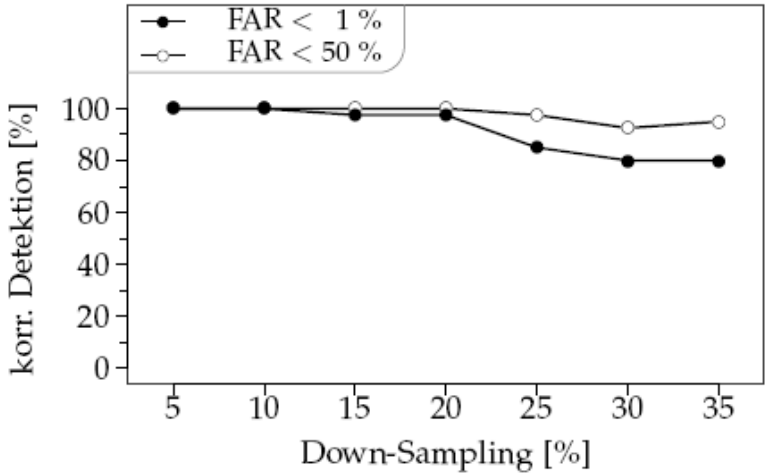


Ergebnisse

Grundlage: Vergrößerung / Verkleinerung von 40 Grauwertbildern

Vergleich mit synthetisch erzeugten periodischen Mustern



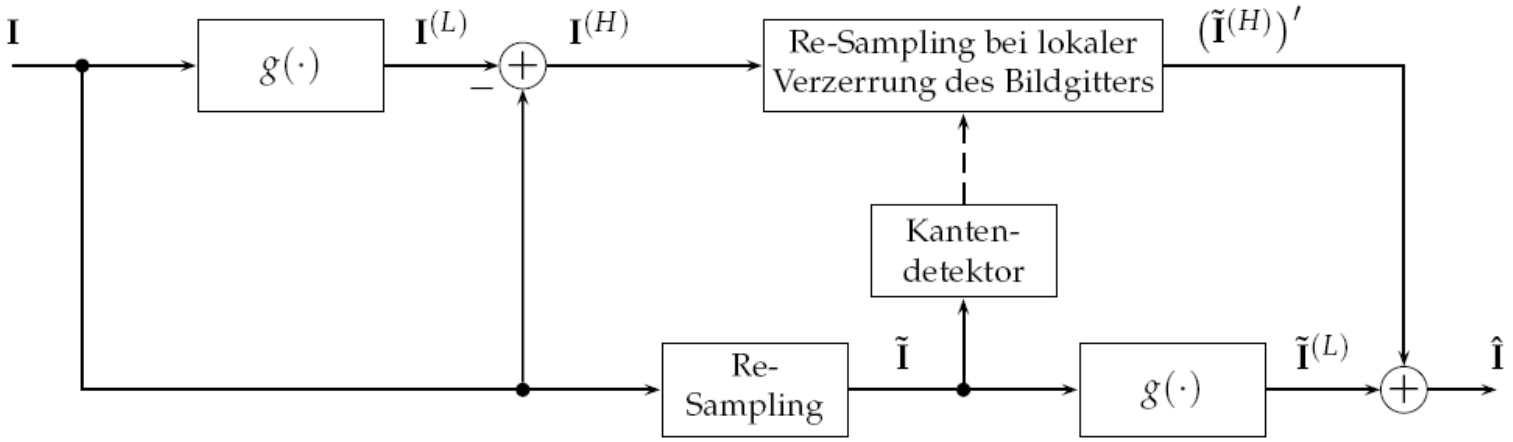


Robustheit

- Sichere Detektion auch nach einfachen Angriffen wie z.B.:
  - Additives Rauschen
  - Gamma-Korrektur
- Schwachstelle: Komprimierte Bilder (JPEG)
  - Block-Artefakte stören periodische Korrelationen
- Ausblick: Komplexere Angriffe

Ausblick (1)

- Angriff auf Basis nicht-linearer Glättung (z.B. Median-Filter) und Re-Sampling bei lokaler Verzerrung des Bildgitters



Ausblick (2)

Vergrößerung um 5%

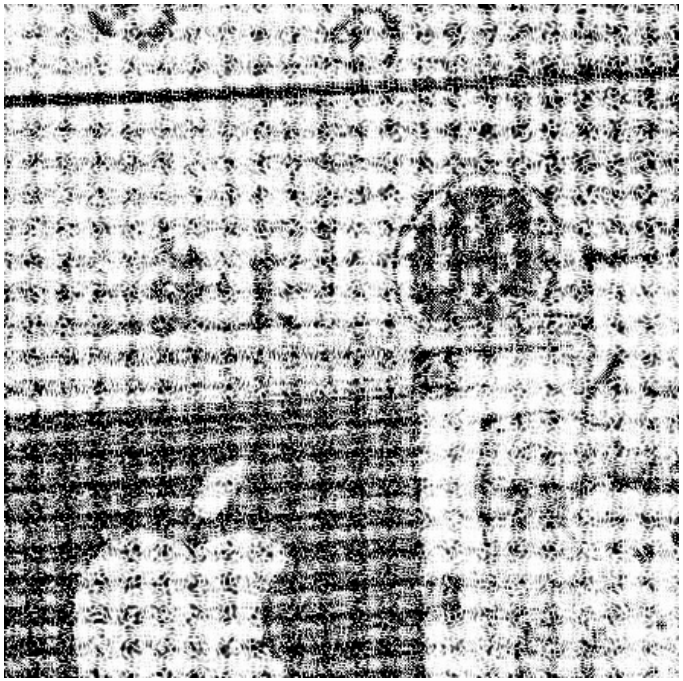
Ohne Angriff

Mit Angriff

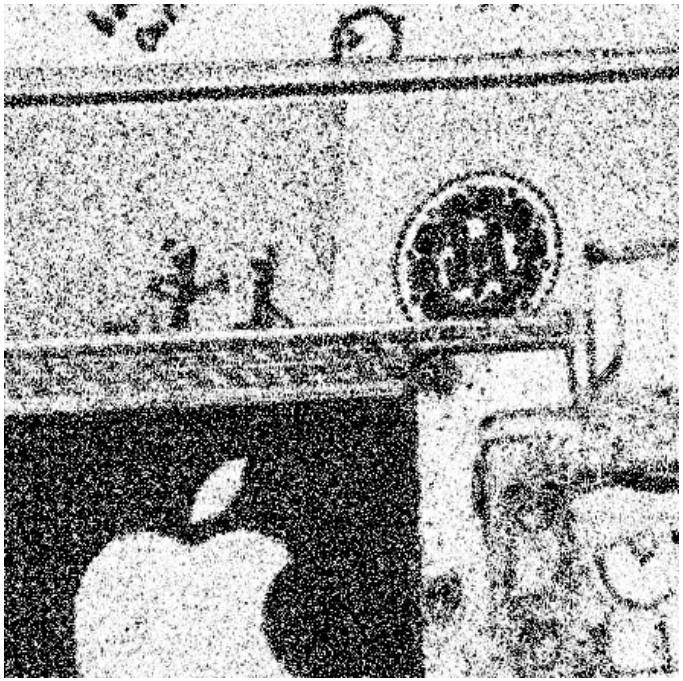
Ausblick (3)

Vergrößerung um 5%

Ohne Angriff



Mit Angriff



Weitere Verfahren (1)

- Verfahren auf Basis von Modellen der Bildentstehung:
  - Analyse der chromatischen Abberation (Johnson & Farid, 2006)
  - Kamera als optischer Tiefpass (Ng et al., 2004)
  - Analyse des Sensorrauschens (Lukáš et al., 2005)
  - Analyse der CFA-Interpolationsartefakte (Popescu & Farid, 2005)
  - Analyse der Kamera-Übertragungsfunktion (Hsu & Chang, 2006; Lint et al. 2005)

Weitere Verfahren (2)

- Verfahren auf Basis von (impliziten) Modellen des Bildinhaltes:
  - Detektion inkonsistenter Belichtungsverhältnisse (Johnson & Farid, 2005)
  - Detektion doppelter Bildteile (Farid, 2006; Fridrich et al., 2003; Popescu & Farid, 2004)
  - Lokales Rauschen (Popescu & Farid, 2004)
  - Doppelte JPEG-Kompression (Popescu & Farid, 2004; Lukáš & Fridrich, 2003)

Fazit

- Bereits zahlreiche Verfahren zur Detektion von Bildfälschungen, welche ohne Wissen zum Originalbild auskommen (blind, passiv)
- Häufig basiererend auf vereinfachenden Annahmen

  ⇒ nicht immer geeignet für "Real World"-Szenarios
- Kombinierter Einsatz mehrerer Verfahren macht die Erstellung überzeugender Fälschungen jedoch bereits ungleich schwerer
- Wichtig: Auseinandersetzung mit komplexen Angriffen zur Entwicklung besserer Detektionsverfahren!

Quellen

- Digitalkamera-Identifikation:

  J. Lukàš, J. Fridrich, M. Goljan: Determining Digital Image Origin Using Sensor Imperfections. Proceedings of SPIE, Vol. 5685, 249-260, 2005.

  http://www.ws.binghamton.edu/fridrich/publications.html

- Detektion von Re-Sampling:

  A.C. Popescu, H. Farid: Exposing Digital Forgeries by Detecting Traces of Re-sampling. IEEE Transactions on Signal Processing, 53(2), 758-767, 2005.

  http://www.cs.dartmouth.edu/farid/publications