9. Digitale Wasserzeichen

9.1 Sicherheitsaspekte

9.2 Wasserzeichen: Historie, Arten und Anwendungen

9.3 Robuste Wasserzeichen

9.4 Offene Probleme

9.5 Fragile Wasserzeichen

Der Foliensatz zu diesem Kapitel beruht auf einem Vortrag von Frau Dr. Jana Dittmann, FhG IPSI, bei der ich mich für die Überlassung der Unterlagen vielmals bedanke.

9.1 Sicherheitsaspekte

Security

Maßnahmen, die beabsichtigte Angriffe auf Rechner, gespeicherte und übertragene Daten sowie Kommunikationsbeziehungen verhindern

Safety

Maßnahmen, die die Auswirkungen unbeabsichtigter Ereignisse, die zu einem Ausfall oder Beschädigung von Rechnern, gespeicherten oder übertragenen Daten und Kommunikationsbeziehungen führen, vermindern

Datenschutz (privacy)

Schutz von personenbezogenen Daten vor unberechtigtem Zugriff

Datensicherheit

Sicherung von Daten vor Verlust (zum Beispiel durch Sicherungskopien)

Beispiel-Szenario

Ein Fotograph findet seine Fotos in einer digitalen Bilddatenbank im Internet, wo sie zum Verkauf angeboten werden. Er ist nicht in der Lage, seine Urheberschaft zu beweisen und Lizenzrechte durchzusetzen, da das digitale Bildmaterial keinen Hinweis auf ihn als Urheber enthält.

Sicherheitsaspekt: Digitale Wasserzeichen

Sicherheitsaspekt	**Kurzbeschreibung**
Zugriffsschutz	Kontrolle des Systemzuganges und Zugriffsbeschränkungen auf Systemfunktionen und Datenbestände
Authentizität	Nachweis der Identität des Urhebers/Autors und des Datenmaterials. Es wird eine Authentifizierung vorgenommen und die Authentizität bestätigt.
Vertraulichkeit	Verhindert, dass unberechtigte Dritte auf Daten zugreifen können
Integrität	Erbringt den Nachweis, dass die Daten unverändert vorliegen
Nachweisbarkeit	Prüfung der Authentizität und Integrität der Daten auch von berechtigten Dritten, so dass die Verbindlichkeit der Kommunikation gewährleistet wird

Urheberrecht (Copyright)

Gegenstand

Werke der Literatur, Wissenschaft und Kunst

Schutz des geistigen Eigentums und der Art und Weise der Gestaltung, vor

- unbefugter wirtschaftlicher Verwertung des Werkes
- Verletzung der ideellen Interessen am Werk.

Rechte

- Urheberpersönlichkeitsrecht (Nennung, Authentizität)
- Verwertungsrechte (Vervielfältigungs-, Verbreitungs- und Senderechte -> Verfolgbarkeit)
- Schranken des Urheberrechts: Allgemeininteresse

Schutz durch das Urheberrecht (copyright) trifft auch auf Computerprogramme und technische Dokumentationen zu.

9.2 Wasserzeichen: Historie, Arten und Anwendungen

9.2.1 Historie

Wasserzeichen sind fest verbunden mit dem Datenmaterial.

Geschichte: Wasserzeichen in Papier, Teppichen, Geldscheinen ...

Digitale Wasserzeichen: steganographische Verfahren

- sichtbare, unsichtbar-robuste oder unsichtbar-fragile Markierungen
- Ziele
    - **Authentizität:** Urheberschutz (copyright protection)
    - **Integrität:** Nachweis, dass das Dokument nicht manipuliert wurde

9.2.2 Arten von Wasserzeichen

- Sichtbare Wasserzeichen zum Urheberschutz oder zur Annotation von Dokumenten mit Meta-Daten
- Unsichtbar-robuste Wasserzeichen
  - Einbringung von versteckten Botschaften
  - Einbringung von Copyright-Informationen, Authentizität (“Copy Control Watermark“)
  - Einbringung von Metainformationen
- Unsichtbar-fragile Wasserzeichen zum Integritätsnachweis (Nachweis, dass nichts gefälscht wurde)

Steganographie

Steganographie: Verwendung unsichtbarer “Wasserzeichen“ für geheime Nachrichten

Idee

Geheime Nachrichten in harmlosen Nachrichten so verbergen, dass ein Dritter nicht erkennt, dass eine geheime Nachricht vorhanden ist.

Beispiel für eine steganographische Nachricht

Ein einfaches Prinzip am Beispiel eines Urlaubsgrußes:

> Liebe Kolleginnen! Wir genießen nun endlich unsere Ferien auf dieser Insel vor Spanien. Wetter gut, Unterkunft auch, ebenso das Essen. Toll!
> Gruß, J. D.

Algorithmus

- Buchstaben bis zum nächsten Leerzeichen zählen
- Anzahl ungerade ergibt eine binäre 0, Anzahl gerade ergibt eine binäre 1.
- Entstehende Binärzahlen in Gruppen zu acht Bits als ASCII-Zeichen interpretieren.

Ergebnis

- erste 8 Wörter 01010011, ASCII ‘S‘
- nächste 8 Wörter ergeben 01001111, ASCII ‘O‘
- letzte 8 Wörter wieder 01010011, ASCII ‘S‘

Der Urlaubsgruß verbirgt den versteckten Hilferuf **“SOS“**.

Steganographie-Techniken

Substitutionale Steganographie

Ersetzen einer verrauschten oder für das Auge oder Ohr nicht wahr zu nehmenden Komponente der digitalen Nachricht durch eine geheime Nachricht.

Konstruktive Steganographie

Nachbildung von Geräuschsignalen basierend auf dem Modell des Originalgeräuschs.

Digitale Wasserzeichen

Zusätzliche Informationen in einem digitalen Dokument zum Nachweis der Urheberschaft, zur Verfolgung von einzelnen Kundenkopien oder zur Integration von Metadaten in Bild, Video, Audio, 3D-Modell oder Software.

Technische Herausforderungen

- Entwicklung von Markierungsverfahren
- Durchführung von Robustheitstests
- Entwicklung von Watermarking-Tools für die verschiedenen Dokumententypen

---

Anforderungen an Wasserzeichen

- **Transparenz**: Verhinderung von Qualitätsverlust im Original-Dokument
- **Robustheit** gegen Transformationen und Angriffe (Entfernen, Ersetzen durch ein fremdes Wasserzeichen)

9.3 Robuste Wasserzeichen

9.3.1 Bild

Verfahren

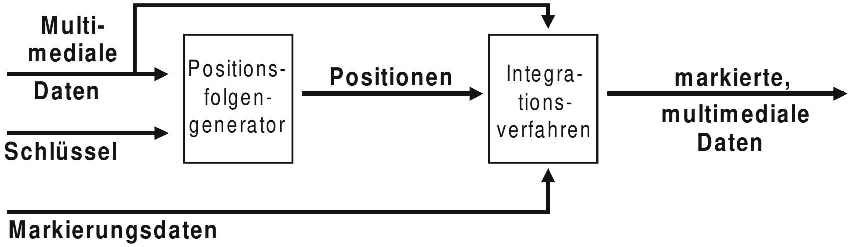
- Bildbereich: Veränderungen an einzelnen Bildpunkten, z. B. Amplituden-Modulation im Blaukanal
- Frequenzbereich: Veränderungen der DCT-Koeffizienten, Wavelet-basierte Verfahren

Kommerzielle Produkte

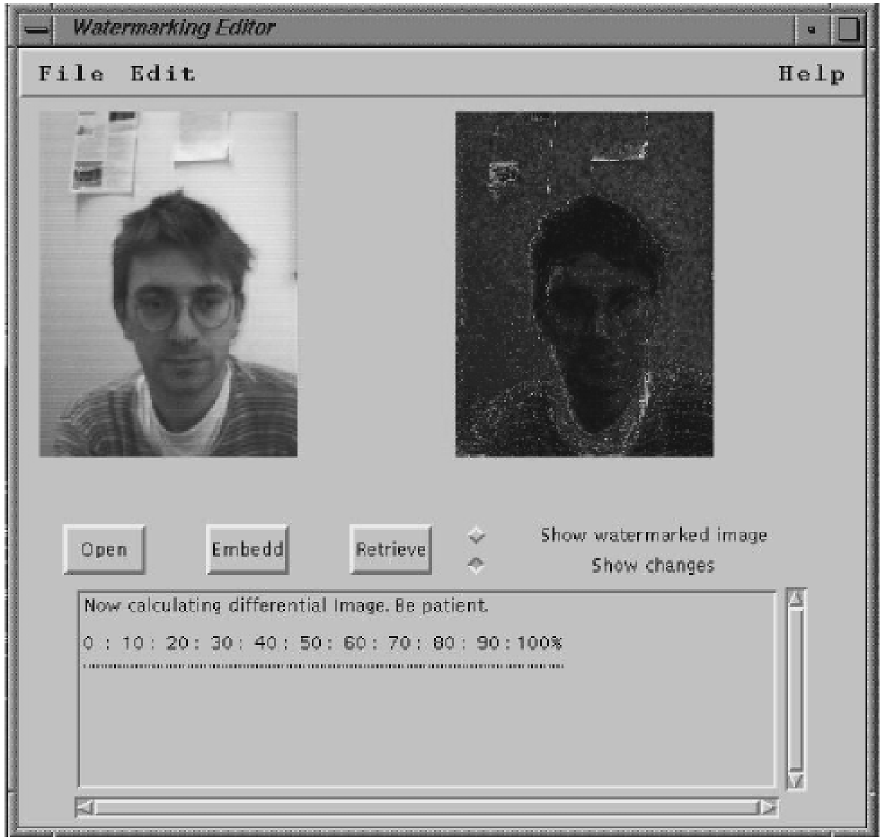
- z. B. Signum Technologies, Digimark Technologies (Adobe Photoshop)

Integration von Text oder Binärfolgen in einem Bild als genau spezifiziertem Rauschmuster.

Einfügen von digitalen Wasserzeichen


Multi-
mediale
Daten
Schlüssel
Positions-
folgen-
generator
Positionen
Integra-
tions-
verfahren
markierte,
multimediale
Daten
Markierungsdaten

Verfahren im Blaukanal (1)

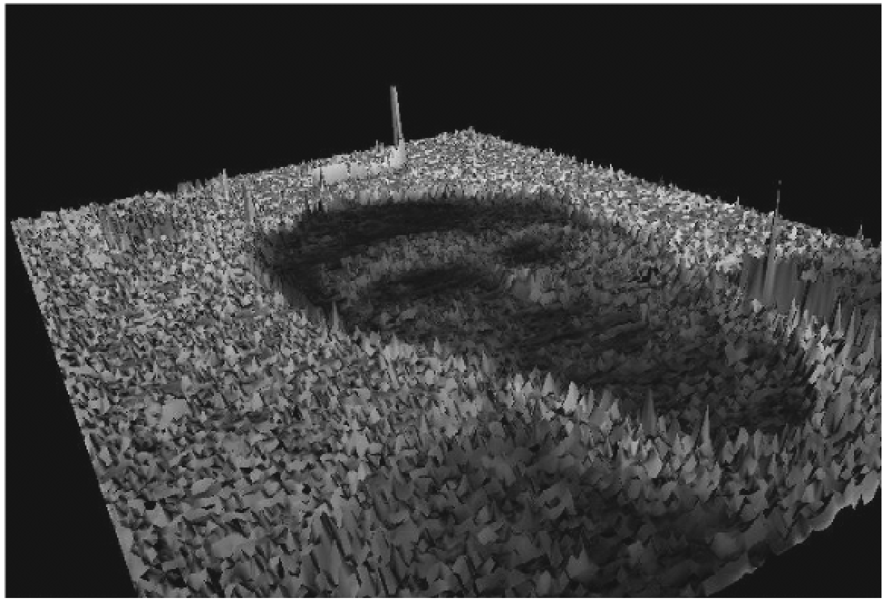

Watermarking Editor
File
Edit
Help
Open
Embedd
Retrieve
Show watermarked image
Show changes

```
Now calculating differential image. Be patient.
0 : 10 : 20 : 30 : 40 : 50 : 60 : 70 : 80 : 90 : 100%
```

Verfahren im Blaukanal (2)

3D-Watermark

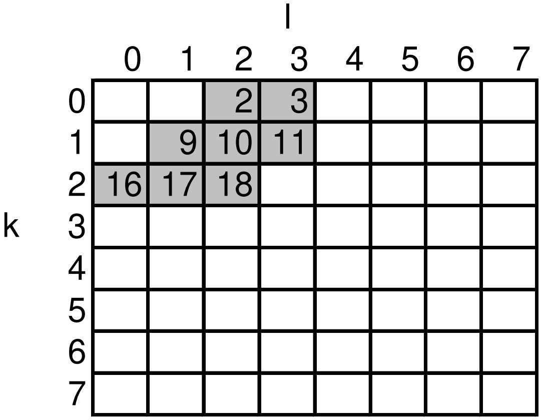
Ein bei FhG/IPSI in Darmstadt entwickeltes Verfahren zur Analyse von Wasserzeichenverfahren für Bilddaten.



Frequenzraumverfahren von Zhao/Koch (1)

Einbetten der Informationen

- Bitfolge c_i, i=1,...,n soll in ein Bild eingebettet werden
- Bit c_i wird durch ein **Verhältnis** zwischen drei Frequenzkoeffizienten im mittleren Bereich eines DCT-Blockes dargestellt
- Es werden acht Positionen im DCT-Block für das Verfahren verwendet, auf denen die Varianz der Graustufen-Werte erfahrungsgemäß nicht allzu groß ist.



k \ l	0	1	2	3	4	5	6	7
0			2	3				
1		9	10	11				
2	16	17	18					
3								
4								
5								
6								
7								

Frequenzraumverfahren von Zhao/Koch (2)

Mögliche Dreierkombinationen

Set No.	(k_1,l_1)	(k_2,l_2)	(k_3,l_3)
1	2 (0,2)	9 (1,1)	10 (1,2)
2	9 (1,1)	2 (0,2)	10 (1,2)
3	3 (0,3)	10 (1,2)	11 (1,3)
4	10 (1,2)	3 (0,3)	11 (1,3)
5	9 (1,1)	2 (0,2)	10 (1,2)
6	2 (0,2)	9 (1,1)	10 (1,2)
7	9 (1,1)	16 (2,0)	2 (0,2)
8	16 (2,0)	9 (1,1)	2 (0,2)
9	2 (0,2)	9 (1,1)	16 (2,0)
10	9 (1,1)	2 (0,2)	16 (2,0)
11	10 (1,2)	17 (2,1)	3 (0,3)
12	17 (2,1)	10 (1,2)	3 (0,3)
13	10 (1,2)	3 (0,3)	17 (2,1)
14	3 (0,3)	10 (1,2)	17 (2,1)
15	9 (1,1)	16 (2,0)	17 (2,1)
16	16 (2,0)	9 (1,1)	17 (2,1)
17	10 (1,2)	17 (2,1)	18 (2,2)
18	17 (2,1)	10 (1,2)	18 (2,2)

Frequenzraumverfahren von Zhao/Koch (3)

Einbetten der Informationen

- Pseudo-zufällige Auswahl einer Dreierkombination aus den acht Frequenzen (Auswahl von Y_A, Y_B, Y_C)
- Vergleich mit der maximal änderbaren Distanz D; wenn überschritten: kennzeichnen als “invalid“
- ansonsten Einbringen des Frequenzmusters (Y_A, Y_B, Y_C) für 0 oder 1 in die Positionen in der Matrix.

Bit	**1**	**1**	**1**	**1**	**0**	**0**	**0**	**0**	**In**	**va**	**lid**
$\mathbf{Y_A}$	H	H	M	M	L	L	M	M	H	L	M
$\mathbf{Y_B}$	H	M	H	M	L	M	L	M	L	H	M
$\mathbf{Y_C}$	L	L	L	L	H	H	H	H	M	M	M

152	0	4	0	0	0	0	0
0	0	23	0	0	0	0	0
0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Frequenzraumverfahren von Zhao/Koch (4)

Einbetten der Bits in die drei Koeffizienten:

- Wenn c_i=1: ändere (Y_A, Y_B, Y_C) so ab, dass $Y_A > Y_C + D$ und $Y_B > Y_C + D$.
- Wenn c_i=0: ändere (Y_A, Y_B, Y_C) so ab, dass $Y_A + D < Y_C$ und $Y_B + D < Y_C$.

Je höher D ist, desto zuverlässiger ist das eingebettete Bit, aber auch desto sichtbarer!

Bit	**1**	**1**	**1**	**1**	**0**	**0**	**0**	**0**	**In**	**va**	**lid**
$\mathbf{Y_A}$	H	H	M	M	L	L	M	M	H	L	M
$\mathbf{Y_B}$	H	M	H	M	L	M	L	M	L	H	M
$\mathbf{Y_C}$	**L**	**L**	**L**	**L**	**H**	**H**	**H**	**H**	**M**	**M**	**M**

152	0	**4**	0	0	0	0	0
0	0	**23**	0	0	0	0	0
0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Frequenzraumverfahren von Zhao/Koch (5)

Auslesen der Information

- pseudo-zufällige Auswahl der n Blöcke aus dem Bild und Koeffizientenkombinationen in den Blöcken
- Lese Y_A, Y_B und Y_C aus
- Prüfe auf “invalid“
- Wenn $Y_A >= Y_C$ und $Y_B >= Y_C$ return $c_i = 1$
- Wenn $Y_A <= Y_C$ und $Y_B <= Y_C$ return $c_i = 0$

Bit	**1**	**1**	**1**	**1**	**0**	**0**	**0**	**0**	**In**	**va**	**lid**
Y_A	H	H	M	M	L	L	M	M	H	L	M
Y_B	H	M	H	M	L	M	L	M	L	H	M
Y_C	**L**	**L**	**L**	**L**	**H**	**H**	**H**	**H**	**M**	**M**	**M**

152	0	**4**	0	0	0	0	0
0	0	**23**	0	0	0	0	0
0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Frequenzraumverfahren von Zhao/Koch (6)

Verfahrens-Vorteile

- visuelle Abschätzung durch “invalid“
- Parameter D bestimmt die Robustheit
- kompressionsrobust (mittlere Koeffizienten)

Verfahrens-Nachteile

- Verschieben von Blöcken zerstört das Wasserzeichen
- Skalierung des Bildes zerstört das Wasserzeichen, da im Original-Algorithmus die Pseudo-Zufallsfolge von der Größe des Bildes (x- und y-Dimension) abhängt
- Sichtbarkeitsparameter (Parameter D) nicht adaptiv an den Bildinhalt (Kanten!)
- Empfindlich gegen Ausschnittbildung
- Empfindlich gegenüber geometrischen Transformationen: Skalierung, Rotation, ...
- Kopieren des Wasserzeichens möglich

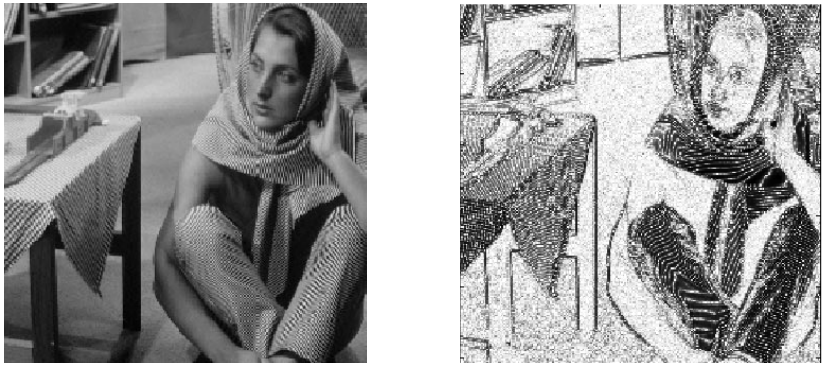
Referenz

Jian Zhao and Eckhard Koch: *Embedding robust labels into images for copyright protection*. In: Proc. of the International Congress on Intellectual Property Rights for Specialized Information, Knowledge and New Technologies. Vienna, Austria, August 1995.
http://www.mediasec.com/know/papers.html

Visuelle Wahrnehmung

Idee

Einfügen des Wasserzeichens in Bildbereiche, in denen das menschliche Wahrnehmungssystem die Modifikation möglichst wenig bemerkt. Hier: im weißen Hintergrund des rechten Bildes.



9.3.2 Ton

Ziele

- Verfolgung illegaler Kopien
- Feststellung des rechtmäßigen Urhebers bzw. Kundens

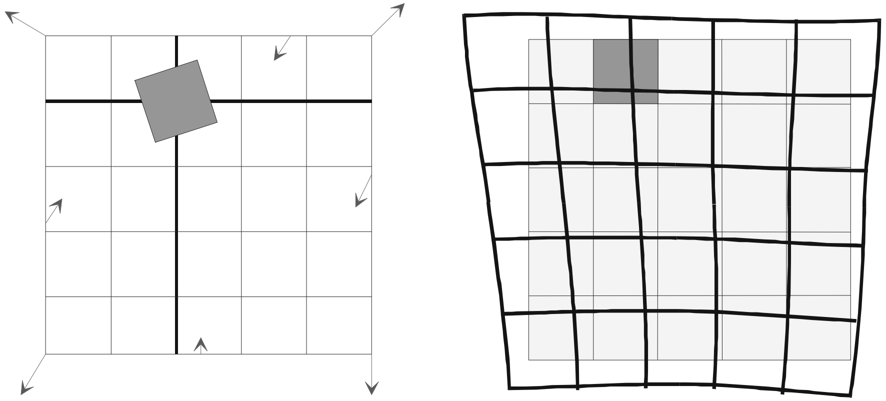
Methoden

- Lautstärken-Modulation
- Modulation des Rauschsignals
- Gefahr: Kompression erfasst und verdirbt die Wasserzeicheninformation (z. B. bei Ausnutzung des Verdeckungseffekts in MP3)

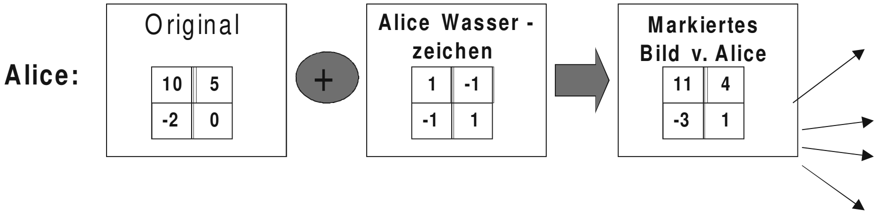
9.4 Offene Probleme

StirMark ist ein Verfahren zum Löschen von Wasserzeichen in Dokumenten. Es simuliert sogenannte Resampling-Prozesse, ähnlich wie Ausdrucken und erneutes Einscannen. Es werden kleine, zufällig ausgewählte, geometrische Operationen ausgeführt, wie Verzerren, Skalieren, Rotieren oder Resampling mit Interpolation.

StirMark löscht fast alle Wasserzeichen der ersten Generation, ohne die Dokumente sichtbar sehr zu verfälschen.



Mehrfachmarkierung



Alice:

Original

10	5
-2	0

+

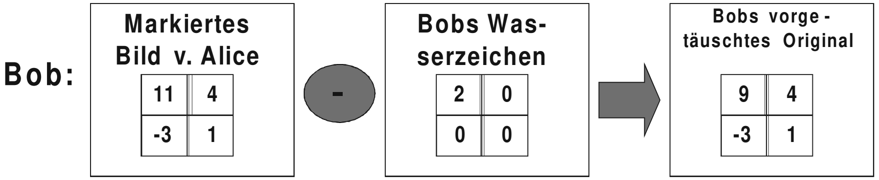
Alice Wasserzeichen

1	-1
-1	1

Markiertes Bild v. Alice

11	4
-3	1

Man betrachte den Angriff: **Mehrfachmarkierung** (Multiple Ownership Attack)



Bob:

Markiertes Bild v. Alice

11	4
-3	1

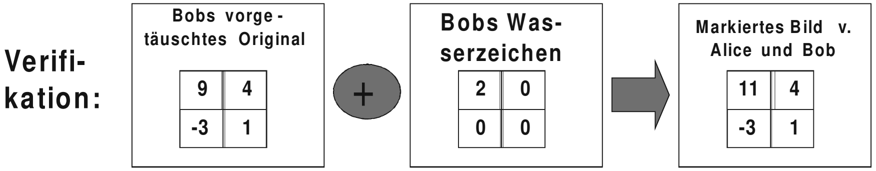
-

Bobs Wasserzeichen

2	0
0	0

Bobs vorgetäuschtes Original

9	4
-3	1



Verifikation:

Bobs vorgetäuschtes Original

9	4
-3	1

+

Bobs Wasserzeichen

2	0
0	0

Markiertes Bild v. Alice und Bob

11	4
-3	1

Digital Fingerprinting

Idee

Einbringen von kundenspezifischen Informationen in jedes ausgelieferte Exemplar des Dokuments.

- Verfolgung kundenspezifischer Kopien
- Erkennen illegaler Weiterverbreitung

Problem mit herkömmlichen Wasserzeichen-Algorithmen:

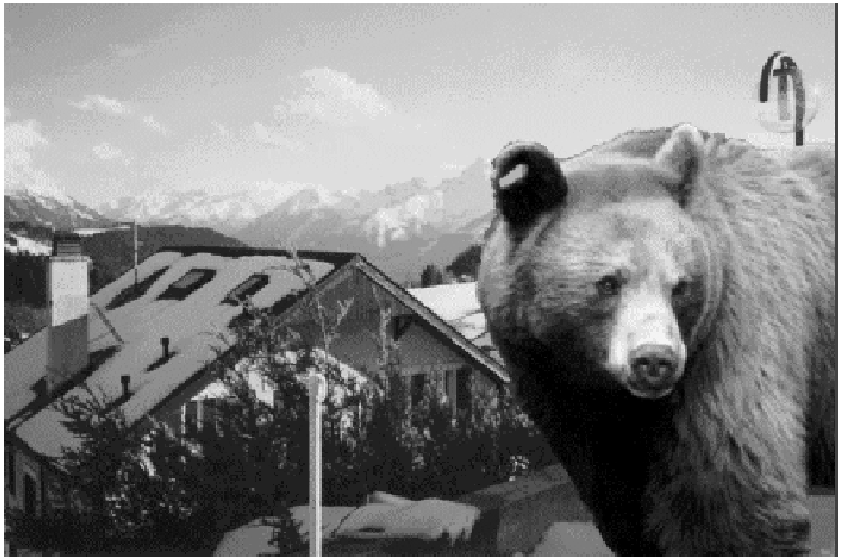
- Da unterschiedliche Informationen eingebracht werden, entstehen unterschiedliche Kopien
- Verfahren skaliert nicht gut für große Kundenzahlen
- Kunden können zusammenarbeiten und Wasserzeichen durch Differenzbildung angreifen: “Koalitionsangriff“

9.5 Fragile Wasserzeichen

Ziel

Schutz von Dokumenten vor unerkannten inhaltlichen Veränderungen. Dazu dienen “zerbrechliche“ (content-fragile watermarks“)

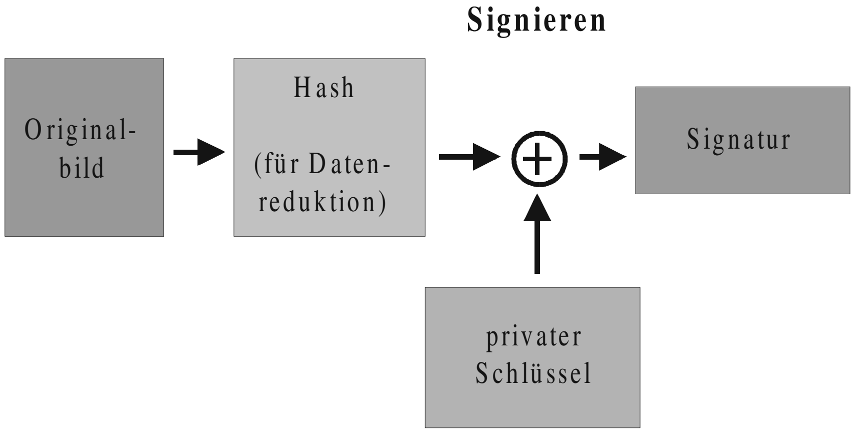
Beispiel



3. Attacks

original warp mosaic

9.5.1 Erster Ansatz: Digitale Signatur

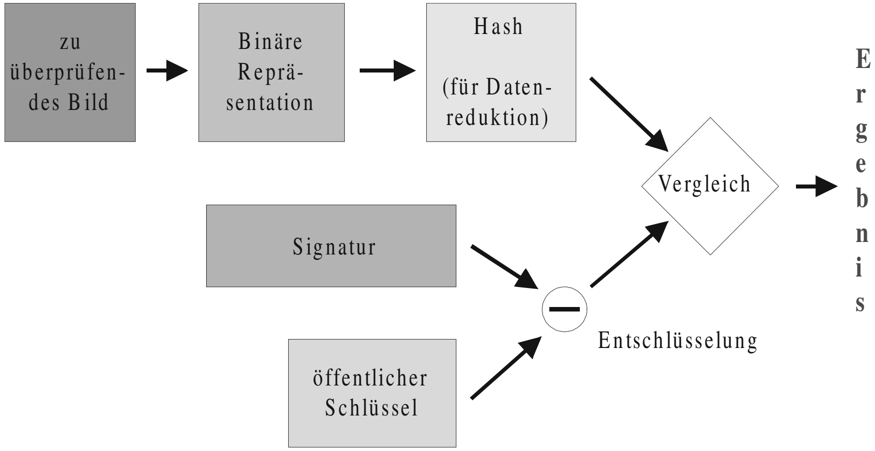


Signieren

Original-bild → Hash (für Datenreduktion) → ⊕ → Signatur

privater Schlüssel → ⊕

Prüfen der digitalen Signatur



zu überprüfendes Bild

Binäre Repräsentation

Hash

(für Datenreduktion)

Vergleich

Ergebnis

Signatur

Entschlüsselung

öffentlicher Schlüssel

Digitale Signatur: Probleme

☑ Authentizität und Integrität

zulässige Bildveränderungen:

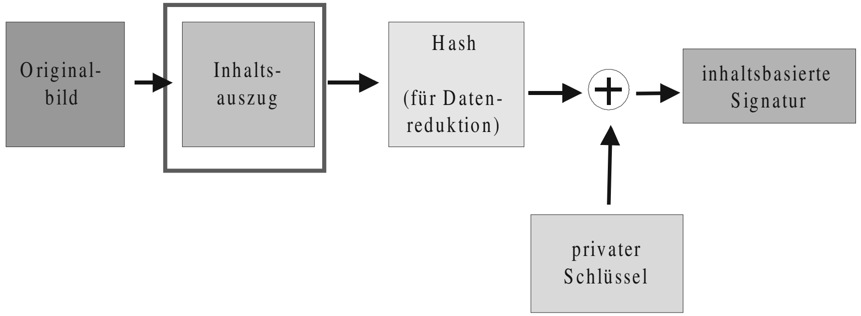
- ☐ leichte Übertragungsfehler
- ☐ Rauschen Quantisierung/
- ☐ Kompression
- ☐ Auflösungsverringerung
- ☐ Skalierung
- ☐ Farbformatkonvertierung
- ☐ Gamma-Vorverzerrung...

zu erkennende Manipulationen:

- Entfernen von Bildelementen
- Verschieben, Verändern von Positionen zueinander
- Veränern der Eigenschaften
- Veränderung der Szenenlichtverhältnisse

Digitale Signatur, inhaltsbasiert

Ein Problem der bisher diskutierten Ansätze ist, dass Rauschen auf dem Übertragungsweg, Skalierung zur Adaption an Bandbreiten und Filter-Operationen zur Qualitätsverbesserung oft das Wasserzeichen zerstören, während für den Betrachter/Hörer eine Veränderung der Qualität kaum feststellbar ist. Deshalb sucht man nach Verfahren, die das Wasserzeichen bzw. die Signatur nur in den **Kernaussagen** des Bildes/Tonsignals unterbringen (**inhaltsbasierte Verfahren**).


Original-bild

Inhalts-auszug

Hash

(für Daten-reduktion)

+

inhaltsbasierte Signatur

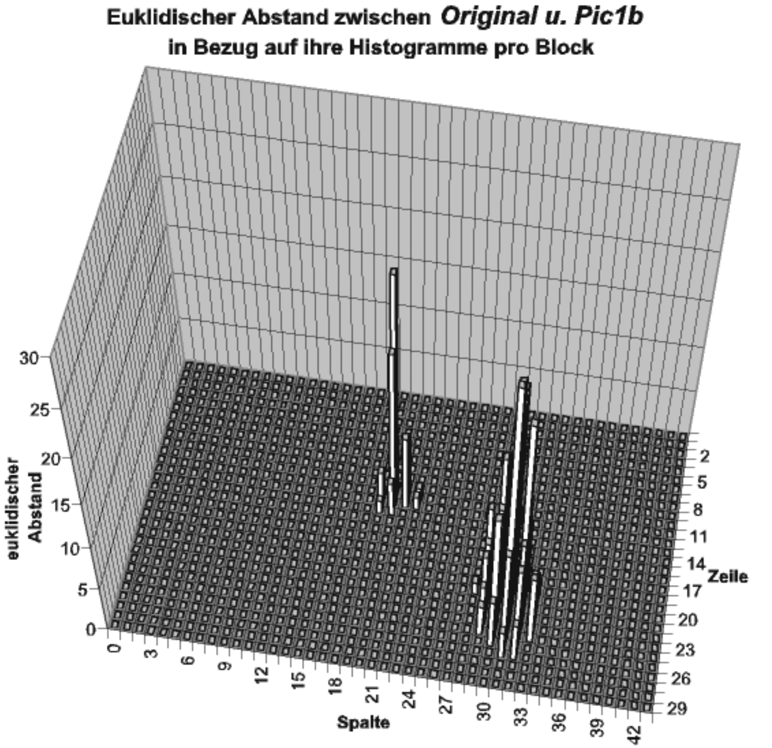
privater Schlüssel

Bestimmung eines Inhaltsauszugs

Beispiele für die Auswahl von "**Kernaussagen**" für Bilder und Videos:

- DC-Koeffizienten pro Block + Vergleich der DC-Koeffizienten
- Vorzeichen der Differenz der DC-Koeffizienten aufeinanderfolgender Blöcke + Vergleich der Vorzeichen (diese hängen kaum von Rauschen oder Filtern ab)
- Intensitätshistogramme pro Block + euklidischer Abstand
- Intensitätshistogramme pro Block + Mittelwertvergleich und Varianzen
- Extraktion der Bildkanten + Vergleich der Kantenbilder

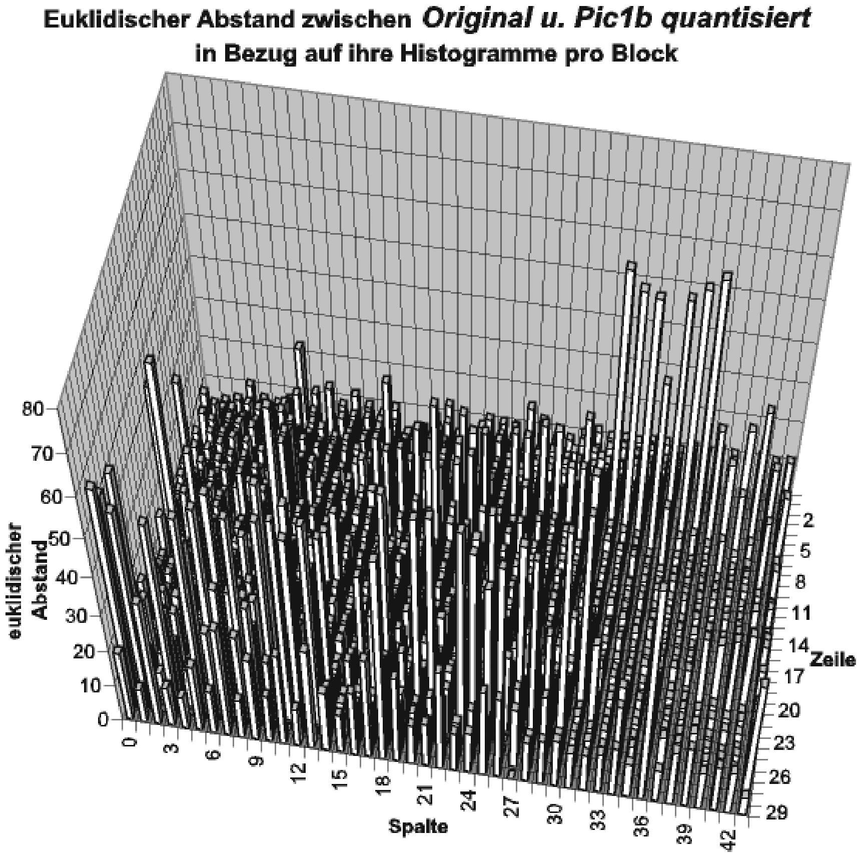
Problem Inhaltsauszug (1)



Euklidischer Abstand zwischen Original und manipuliertem Bild

$$\left[\sum_{i=1}^{n}(x_i - y_i)^2\right]^{\frac{1}{2}}$$

Problem Inhaltsauszug (2)



Euklidischer Abstand zwischen Original und manipuliertem sowie quantisiertem Bild

Problem Tauglichkeit Inhaltsauszug

DCT-Ansätze → **Probleme bei der Skalierung/Quant.**

Histogramm-Ansätze → **Schwellwerte nötig!**

Kantenbilder-Ansätze → **Probleme bei der Skalierung: Blockansatz**

“Zugelassene” Bildveränderungen:
Farbkonvertierungen, teilweise Übertragungsfehler, teilweise “Quantisierung”

9.5.2 Zweiter Ansatz: Inhaltsauszug als Wasserzeichen

Idee

Inhaltsauszug als fragiles Wasserzeichen in den Datenstrom einbringen

Verfahren

- Schwellwertbasis
- Inhaltsauszug

Unsicherheit

- fragil versus robust gegen zugelassene Veränderungen wie Skalierung und Kompression
- korrekte Manipulationsdetektion?

Fragile Wasserzeichen sind erst am Beginn der Forschung.

Zusammenfassung

Wasserzeichen

- Pragmatischer Ansatz für Copyrightschutz
- Möglichkeit der Verfolgung illegaler Kopien und Markierung einzelner Exemplare mit kundenspezifischen Informationen (Fingerprinting)
- Von großer Bedeutung für das WWW!

Forschungsbedarf

- Robustheit (Retrieval), Robustheitsgarantien: Zertifizierte Wasserzeichenverfahren zum **Herkunftsnachweis**
- fragile Wasserzeichen zum **Integritätsnachweis**
- Anwendung auf Audio, 3D-Szenen, Software?