

# Fingerprinting auf Basis der Geometrischen Struktur von Videos

Dima Proefrock, Mathias Schlaueg, Erika Müller

Universität Rostock, Institut für Nachrichtentechnik, Richard Wagner Str. 31, 18119 Rostock, {dima.proefrock, mathias.schlaueg, erika.mueller}@uni-rostock.de

## 1. Einleitung

Mit steigender Bandbreite privater Internetanschlüsse steigt auch die Attraktivität, aktuelle Filme vor ihrem Erscheinen im Free-TV als Download zu kaufen (z.B. In2Movies, T-Online) oder sie über ein Video on Demand System (z.B. Maxdome, One4Movie) zu beziehen. Der Zugang zu digitalen Kopien aktueller Filme kann allerdings auch von Raubkopierern ausgenutzt werden. Aktuelle Sicherheitsmaßnahmen, wie Verschlüsselungstechnologien und Set-Top-Boxen, können in der Regel leicht umgangen werden oder sind mit starker Qualitätsreduzierung verbunden.

Eine Alternative bietet Fingerprinting (siehe [1]). Beim Fingerprinting wird in jede verkaufte digitale Kopie eines Videos ein individueller Fingerabdruck eingebettet. Taucht eine illegale Kopie des Videos auf, kann mit Hilfe des Fingerabdruckes der Raubkopierer identifiziert werden (Traitor Tracing). Der Fingerabdruck wird direkt in die Videodaten eingebettet und muss unter anderem robust gegenüber Formatkonvertierungen (z.B. MPEG2  $\rightarrow$  H.264), Skalierung und starker Kompression sein.

In dieser Arbeit wird ein neuartiges Fingerprinting-Verfahren vorgestellt. Der Fingerprint wird als Folge von Bits in die geometrische Struktur eines Videos eingebettet. Dabei wird gezeigt, dass die geometrische Struktur eines Videos sehr robust gegenüber Kompression ist. Die Herausforderungen bestehen darin, die Struktur zu erfassen, sie gezielt zu verändern und dabei die Videoqualität zu erhalten.

Ein wesentliches Element bei der Strukturerkennung ist die Bestimmung der räumlichen Position einzelner Kanten. Es wird gezeigt, wie mit Hilfe des Normierten Schwerpunktes eine Entkoppelung zwischen der Erfassung einer Kante und ihrer räumlichen Position möglich ist. Da eine Kante nicht direkt erfasst werden muss, entfallen aufwendige Verfahren der Bildverarbeitung. Es wird gezeigt, wie mittels herkömmlicher Verzerrungsverfahren (Warping bzw. Resampling (siehe [2])) die Kantenpositionen manipuliert und somit die Fin-

gerprint-Bits eingebettet werden. Dabei wird das Video als 3-Dimensionaler Raum verstanden, in dem die Fingerprint-Bits adaptiv (in Abhängigkeit des Videoinhaltes) verteilt werden. Die eingebrachten Verzerrungen sind lokal begrenzt und die Verzerrungsstärke ist relativ klein. Dadurch kann eine ausreichend gute Videoqualität gewährleistet werden. Es wird gezeigt, dass der Fingerprint sehr robust ist. Ein geringer Rechenaufwand ist ein wesentlicher Punkt bei der Integration des Verfahrens in bereits bestehende Download bzw. Streaming Systeme. In dieser Arbeit wird eine Möglichkeit aufgezeigt, wie viele individuelle Videokopien mit sehr geringem Rechenaufwand generiert werden können.

Diese Arbeit gliedert sich in folgende Kapitel. Nach der Einleitung folgt im 2. Kapitel die Beschreibung der indirekten Parametrisierung der geometrischen Struktur von Bilddaten. In Kapitel 3 wird gezeigt, wie die geometrische Struktur verändert wird, um ein Fingerprint-Bit einzubetten. Im folgenden Kapitel wird ein Ansatz zur Vermeidung von Artefakten erläutert. Im 5. Kapitel wird ein Konzept vorgestellt, mit dem sich komprimierte Videos mit individuellen Fingerabdrücken mit geringem Rechenaufwand generieren lassen.

## 2. Indirekte Parametrisierung der geometrischen Struktur

Um die geometrische Struktur von Bildern oder Videos als Einbettungsdomain zu nutzen, muss diese parametrisiert werden. Es kann zwischen der direkten und der indirekten Parametrisierung unterschieden werden. Bei der direkten Parametrisierung werden Kanten, Ecken und Objekte innerhalb eines Bildes erfasst und durch Parameter beschrieben. Die Verfahren zur Erfassung sind rechenintensiv und, in Kombination mit Fingerprinting, fehleranfällig. Die blockbasierte indirekte Parametrisierung liefert für die geometrische Struktur eines Blockes Parameter, ohne dass Kanten, Ecken und Objekte erfasst werden müssen.

Basis der Parametrisierung ist der normierte Schwerpunkt (Normed Centre of Gravity – NCG) eines Blockes. Der NCG liefert die Schwerpunktkoordinaten  $x, y \in \mathbb{R}[0, n]$  eines Blockes der Größe  $n \times n$  und die Schwerpunktsstärke

$$L \in \mathbb{R} \left[ 0, k \cdot \sum_{i=1}^{\frac{n}{2}} \sin \left( \frac{\pi}{n} + (i-1) \cdot \left( \frac{2 \cdot \pi}{n} \right) \right) \right] \text{ wobei } k \text{ die Anzahl der Grauwerte pro Bild-}$$

punkt ist. Die Schwerpunktkoordinaten werden auf ein adaptives Quantisierungsgitter projiziert und ergeben den Positionskennwert  $s \in \mathbb{R}[0, 1]$ . Der genaue Vorgang der Berechnung von  $L$  und  $s$  ist in [3] zu finden. Das vereinfachte Prinzip der Berechnung wird in Abb. 1 gezeigt.

Der Positionskennwert  $s$  ist die Abbildung der geometrischen Struktur des Blockes auf einen Parameter. Über die Schwerpunktsstärke  $L$  lässt sich die Robustheit der geometrischen Struktur bzw.  $s$  gegenüber verlustbehafteter Kompression abschätzen. Dadurch ist es möglich, für die Einbettung des Fingerprints gezielt robuste Blöcke eines Videos auszuwählen.

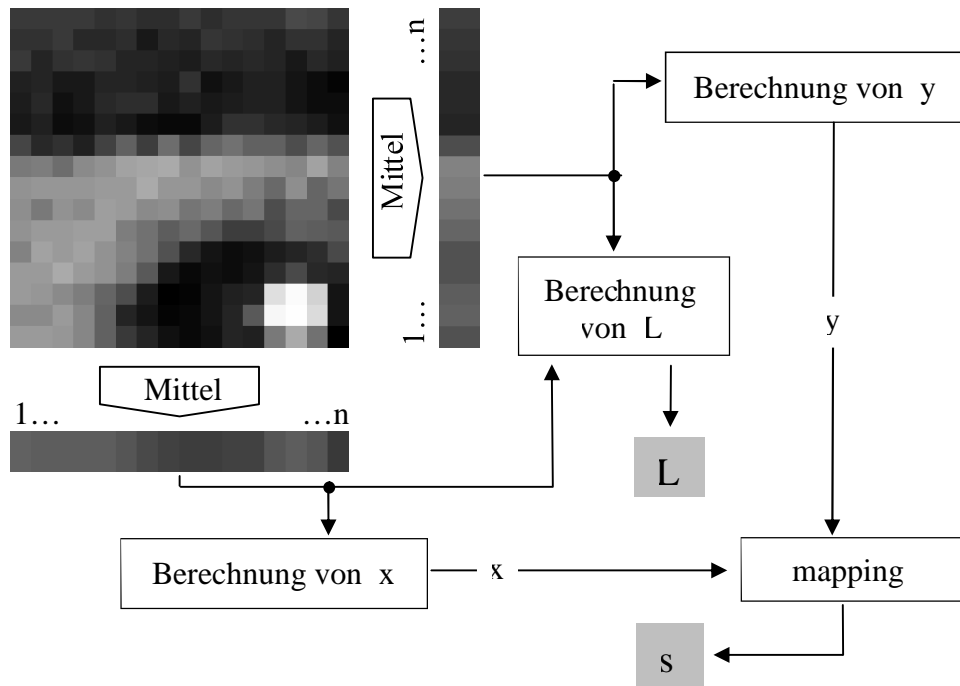


Abb. 1: Parametrisierung der geometrischen Struktur eines Blockes.

### 3. Einbettung eines Fingerprint-Bits

Zur Einbettung eines Fingerprint-Bits wird die geometrische Struktur innerhalb eines Blockes durch geometrische Verzerrungen gezielt verändert. Es hat sich gezeigt, dass ein einheitliches Verzerrungsfeld dafür geeignet ist. Das Feld wird mit den Parametern  $v_x, v_y \in \mathbb{R}[v_{\min}, v_{\max}]$  beschrieben. Zur Einbettung eines Bitwertes '0' bzw. '1' wird der Block so verzerrt, dass der resultierende Positionskennwert  $s$  gleich 0 bzw. 1 ist. Zur Berechnung von  $v_x$  und  $v_y$  wird ein iterativer gradientenbasierter Suchalgorithmus verwendet. Ein Beispiel für die Einbettung eines Bits wird in Abb. 2 gezeigt.

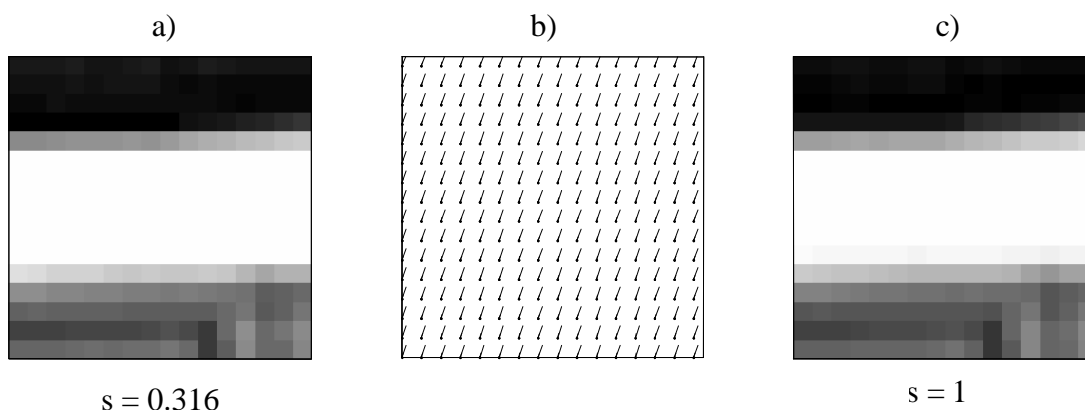
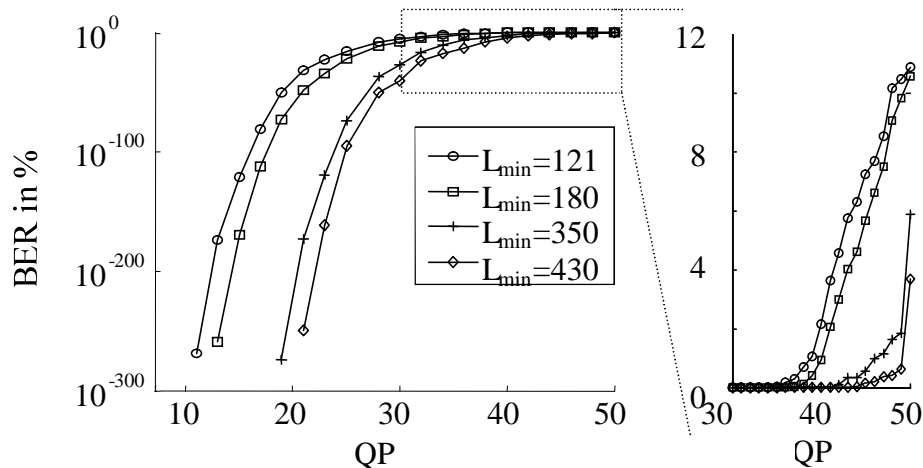


Abb. 2: Beispiel für die Einbettung eines Bits '1' durch Verzerrung mit a) Original-Block, b) Verzerrungsfeld (Verstärkt um Faktor 3) und c) Block mit eingebettetem Bit.

Die Robustheit des eingebetteten Bits hängt von der verwendeten maximalen Einbettungsstärke (Verzerrungsstärke) und der Eignung der geometrischen Struktur (Schwerpunktstärke) eines Blockes ab. Im folgenden Fall wird die Robustheit gegenüber H.264/AVC Kompression untersucht. Die Verzerrungsstärke beträgt  $v_{\min} = -1$  und  $v_{\max} = 1$ . Es werden Blöcke mit einer Schwerpunktstärke  $> L_{\min}$  verwendet. QP ist ein Parameter der H.264/AVC Kompression. Ein höheres QP resultiert in einer stärkeren Kompression.

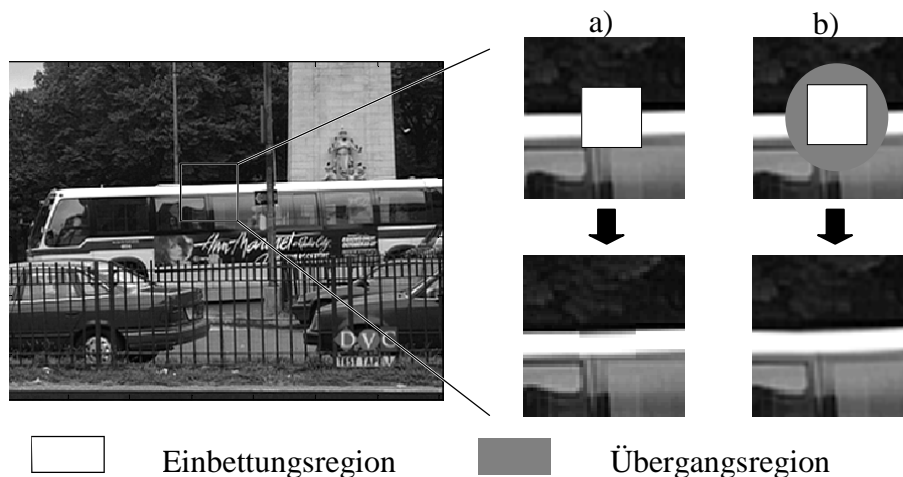


**Abb. 3: Robustheit der eingebetteten Bits gegenüber H.264/AVC Kompression**

Wie in Abb. 3 gezeigt wird, sind die eingebetteten Bits sehr robust gegen starke verlustbehaftete Kompression. So ist die Bit Error Rate bei  $QP = 45$  (resultierendes PSNR  $< 25$  dB) nur 0.16%.

#### 4. Vermeidung von Artefakten

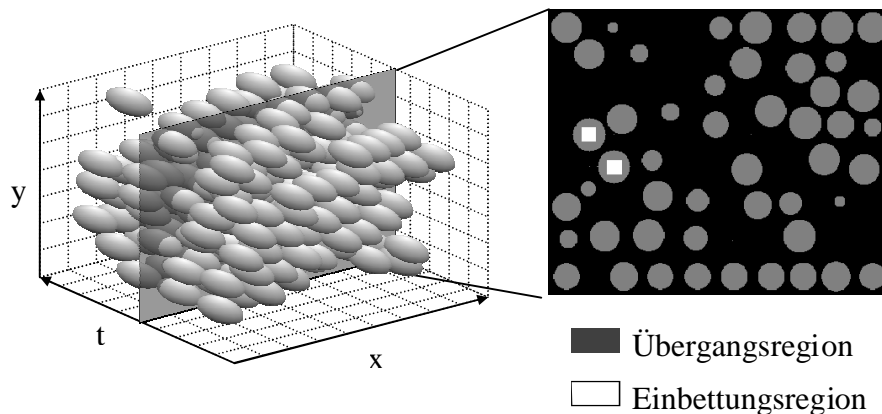
Wird der Inhalt eines Blockes geometrisch verzerrt, können, insbesondere zwischen verzerrten und unverzerrten Blöcken, Artefakte entstehen. Zur Vermeidung von Artefakten definieren wir um die verzerrten Blöcke eine Übergangsregion.



**Abb. 4: Beispiel für Artefaktvermeidung mit a) ohne und b) mit Übergangsregion**

Die Übergangsregion bildet einen Ellipsoid mit dem Radius  $r$ , in dessen Zentrum sich der Block befindet. Die Verzerrung des Blockes wird auf die Übergangsregion übertragen und linear mit dem Abstand zum Ellipsoidenrand verringert (Beispiel siehe Abb. 4).

Um möglichst viele Bits innerhalb des Videos einzubetten, müssen die Ellipsoide innerhalb des 3-dimensionalen Videoraumes (2 räumliche Dimensionen  $x$ ,  $y$  und eine zeitliche Dimension  $t$ ) möglichst dicht zusammenliegen. Dabei dürfen sie sich nicht überschneiden. Da nicht alle Blöcke als Bit-Träger in Frage kommen, können die Ellipsoide nicht in gleichmäßigen Strukturen angeordnet werden, für die bereits optimale Packungsdichten berechnet wurden (siehe [4]). Um dieses Problem zu lösen, wurde ein effizienter Algorithmus entwickelt, der die Ellipsoide möglichst dicht aber mit einem Mindestabstand in dem Videoraum anordnet. Eine genaue Beschreibung ist in [5] zu finden. Ein Beispiel für die Anwendung des Algorithmus wird in Abb. 5 gezeigt. Es wird der Querschnitt (Frame 75) durch die Ellipsoide des Videos „Bus“ gezeigt.



**Abb. 5: Ellipsoide des Videos „Bus“.**

Bei einer Bildgröße von  $352 \times 288$ , einer Blockgröße von  $n=16$  und einem Ellipsoiddurchmesser von 12 bis 20 Bildpunkten wird so eine durchschnittliche Kapazität von 4 bis 2 Bits/Frame erreicht. Ein größerer Ellipsoiddurchmesser führt zu einer verringerten Bildung von Artefakten senkt aber auch die Kapazität bzw. den möglichen Payload des Fingerprints.

## 5. Effiziente Generierung komprimierter Videokopien

Zur Generierung der komprimierten Kopie eines Videos, inklusive einem individuellen Fingerprint, wird GOP-Interleaving (GOP – Group of Pictures) genutzt, um den Rechenaufwand zu reduzieren. Dafür wird aus dem Originalvideo jeweils eine Kopie erzeugt, in der die Fingerprint-Bitfolgen die Werte '0' bzw. '1' haben. Anschließend werden beide Kopien komprimiert. Durch die Kombination von GOP's aus beiden Kopien können neue Kopien mit unterschiedlichen Fingerprint-Bitfolgen erzeugt werden. Der Rechenaufwand für diesen Schritt ist sehr gering. Der Einbettungsprozess und die Kompression muss jeweils nur zweimal durchgeführt werden. Das Prinzip wird in Abb. 6 gezeigt.

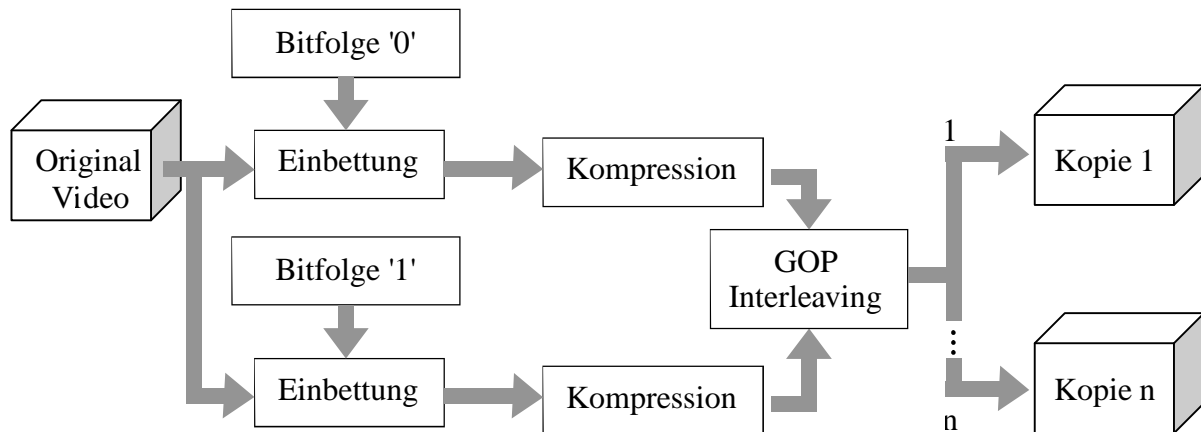


Abb. 6: Generierung von Videokopien mittels GOP-Interleaving

## 6. Zusammenfassung

Fingerprinting ist eine Möglichkeit, den Ursprung illegaler Kopien von Videos zu identifizieren. In dieser Arbeit wurde ein Verfahren vorgestellt, das die geometrische Struktur eines Videos als Fingerprint-Einbettungsdomain nutzt. Es wurde gezeigt, dass die eingebettete Bitfolge sehr robust gegen starke verlustbehaftete Kompression ist. Mit dem vorgestellten Verfahren kann das Auftreten von Artefakten verhindert werden. Das Verfahren kann ohne großen Kostenaufwand in bestehende Download bzw. Streaming Systeme integriert werden.

## 7. Literatur

- [1] Kundur D., Karthik K. (2004) Video Fingerprinting and Encryption Principles for Digital Rights Management. In Proc. of IEEE, Vol. 92, no. 6, pp. 918–932
- [2] Dodgson N. A. (1997) Quadratic Interpolation for Image Resampling. IEEE Trans. on Image Processing, Vol.6, pp.1322-1326
- [3] Präfrock D., Schlaueg M., Müller E. (2006) Video Watermarking by Using Geometric Warping Without Visible Artifacts. Proc. of Information Hiding (8th IH 2006)
- [4] Schürmann A. (2002) Dense Ellipsoid Packings. Journal of Discrete Mathematics, vol. 247, n 1-3, pp. 243-249
- [5] Präfrock D., Schlaueg M., Müller E. (2007) Geometric Warping Watermarking Extended Concerning Geometric Attacks and Embedding Artifacts. Proceedings of 9th ACM Multimedia and Security Workshop