

# Nutzung der Nachbarkorrelation für Steganalyse mit erhöhter Zuverlässigkeit

Andreas Westfeld

Technische Universität Dresden  
Institut für Systemarchitektur  
westfeld@inf.tu-dresden.de

**Zusammenfassung:** Die Vielzahl aufgebotener Steganalysemethoden zeigt bei niedriger Änderungsdichte und kleinen Trägermedien Schwächen. Die hier vorgestellte Methode ist anwendbar, aber nicht beschränkt auf LSB-Steganographie. Sie tastet Medien entlang einer raumfüllenden Kurve ab. Beispielsweise kann die Abtastung entlang einer Hilbertkurve die Korrelation benachbarter Farb- oder Helligkeitswerte eines Bildes besser ausnutzen, um steganographische Nachrichten mit niedriger Änderungsdichte zuverlässiger nachzuweisen. Ein Vergleich von fünf verschiedenen Abtastreihenfolgen zeigt, dass rekursive Ansätze am leistungsfähigsten sind. Unter Verwendung einer umfangreichen Datenbank von kleinen Bildern wird die erhöhte Zuverlässigkeit der modifizierten steganalytischen Tests nachgewiesen. Auch die Visualisierung steganographischer Größen kann von der Anwendung inverser raumfüllender Kurven profitieren.

## 1 Einführung

Steganographie ist die Kunst und Wissenschaft der unsichtbaren Kommunikation. Ihr Ziel ist die Übertragung von Informationen, die unsichtbar in ein Trägermedium eingebettet werden. Das Ziel der Steganalyse ist, steganographische Änderungen in Trägermedien zu entdecken.

Die Trägerdaten können zum Beispiel Bild- und Audiodateien sein, welche gewöhnlich lokale Korrelationen aufweisen. Örtlich bzw. zeitlich benachbarte Stichprobenwerte (Bildpunkte, Frequenzkoeffizienten usw.) weisen dann kleinere Unterschiede auf als weiter entfernt liegende.

Während die ersten steganalytischen Methoden nur Histogramme betrachteten und diese Korrelationen gänzlich ignorierten (z. B. Chi-Quadrat-Angriff [WP00]), berücksichtigen neuere Methoden die Nachbarschaft von Stichprobenwerten (beispielsweise die Paaranalyse [FGS03], der RS-Angriff [FGD01a] und der Probepaarangriff [DWW02]). Jedoch werden auch bei den neueren Methoden die Korrelationen nicht optimal betont, weil die Stichprobenwerte zeilenweise eingesammelt werden. Im vorliegenden Beitrag soll gezeigt werden, dass rekursive, raumfüllende Kurven mehrdimensionale Korrelationen am besten ausnutzen können und die Leistungsfähigkeit der bekannten steganalytischen Techniken

maximieren. Andere Forscher haben bereits mit unterschiedlichen Arten der Gruppierung von Bildpunkten experimentiert. Jessica Fridrich schlug vor, im RS-Angriff quadratische Gruppen von je  $2 \times 2$  Bildpunkten statt längliche  $4 \times 1$  Bildpunkte zu verwenden [FGD01a]. Andrew Ker arbeitete daran weiter und schlug vor, nicht benachbarte Bildpunkte von der Paaranalyse auszuklammern und die Abtastumgebung für den RS-Angriff zu vergrößern [Ker04].

LSB-Steganographie ist die am stärksten erforschte und bekannteste steganographische Methode. Gerade deshalb eignen sich Angriffe darauf besonders gut als Beispiel zur Demonstration der Leistungsfähigkeit von raumfüllenden Kurven, obwohl deren Nutzen nicht auf den Nachweis dieses Verfahrens beschränkt ist.

Der nächste Abschnitt bietet eine Übersicht, welche Transformationen mit raumfüllenden Kurven möglich sind. Abschnitt 3 vergleicht die herkömmliche zeilenweise Abtastung mit vier verschiedenen Optimierungen. Wir werden feststellen, dass die rekursiven Abtastpfade am besten abschneiden. Am Beispiel der Paaranalyse beschreibt Abschnitt 4 die Anwendung raumfüllender Kurven. Das letzte Kapitel fasst zusammen.

## 2 Anwendung raumfüllender Kurven

Eine raumfüllende Kurve bildet ein eindimensionales Intervall auf eine zweidimensionale Fläche ab (flächenfüllend) oder auf einen dreidimensionalen Raum. David Hilbert, ein deutscher Mathematiker, erfand eine einfache raumfüllende Kurve, die als Hilbertkurve bekannt ist und ein Quadrat füllt. Die Hilbertkurve kann einfach durch eine initiale Zeichenkette  $L$  und die folgenden Umformungsregeln implementiert werden [Sau88]

$$\begin{aligned}
 L &\longrightarrow +RV - LVL - VR + \\
 R &\longrightarrow -LV + RVR + VL - \\
 V &\longrightarrow \text{gehe einen Bildpunkt vorwärts} \\
 + &\longrightarrow 90^\circ \text{ nach rechts} \\
 - &\longrightarrow 90^\circ \text{ nach links}
 \end{aligned}$$

Diese Regeln terminieren nach einer festgelegten Rekursionstiefe.

### 2.1 Ein Bild in einen Vektor umwandeln

Abbildung 1 zeigt die bekannte Hilbertkurve für die Rekursionstiefe 1, 2 und 3. Die fette Nummerierung folgt der Kurve und die kursive den Zeilen. Beide werden später zur Erläuterung von Permutationen gebraucht. Mit der ersten Kurve können  $2 \times 2$ , mit der zweiten  $4 \times 4$  und mit der dritten  $8 \times 8$  Bildpunkte abgetastet werden. Für größere Bildformate muss die Rekursionstiefe entsprechend vergrößert werden. Generell kann festgestellt werden, dass für ein Bild mit den Abmessungen  $x$  und  $y$  (in Bildpunkten) die Rekursionstiefe

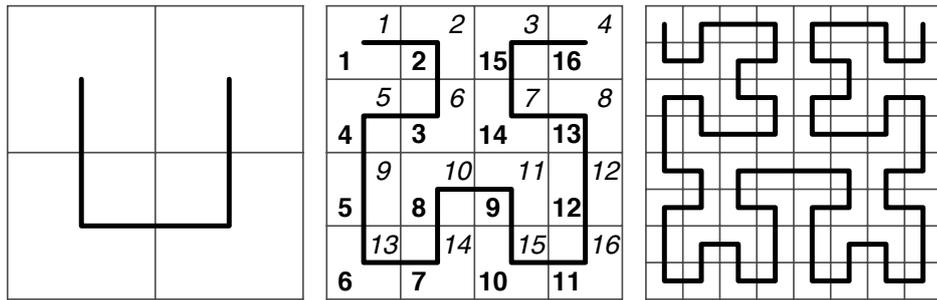


Abbildung 1: Hilbertkurven für die Rekursionstiefe 1, 2 und 3

$\lceil \log_2 \max(x, y) \rceil$  gebraucht wird. Ein  $800 \times 600$ -Bild erfordert beispielsweise eine Hilbertkurve mit einer Tiefe von 10, die ausreichend für bis zu  $1024 \times 1024$  Bildpunkte ist. Die Positionen der Hilbertkurve außerhalb des Bildes werden ignoriert und tragen nicht zum resultierenden Vektor von Bildpunkten bei.

Abbildung 2 zeigt eine zweite Variante der Hilbertkurve. Mit der Rekursionstiefe 1, 2 bzw.

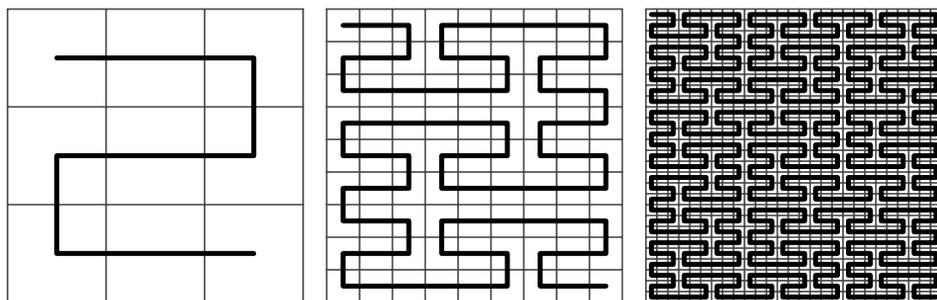


Abbildung 2: Variante 2 der Hilbertkurve für die Tiefen 1, 2 und 3

3 können  $3 \times 3$ ,  $9 \times 9$  bzw.  $27 \times 27$  Bildpunkte abgetastet werden. Auch hier muss für größere Bildformate die Rekursionstiefe entsprechend vergrößert werden. Für ein Bild mit den Abmessungen  $x$  und  $y$  (in Bildpunkten) wird die Hilbertkurve 2 mit einer Rekursionstiefe  $\lceil \log_3 \max(x, y) \rceil$  gebraucht. Ein  $800 \times 600$ -Bild erfordert beispielsweise eine Hilbertkurve 2 mit einer Tiefe von 7, die ausreichend für bis zu  $2187 \times 2187$  Bildpunkte ist.

Am Beispiel eines  $6 \times 4$ -Bildes zeigt Abbildung 3, wie die erste Variante der Hilbertkurve mit der Rekursionstiefe 3 bzw. die zweite Variante mit der Rekursionstiefe 2 an das Bild angepasst wird, indem außerhalb des rechteckigen Formats liegende Positionen der Kurven ausgelassen werden.

Im folgenden Beispiel bringen wir die Bildpunkte eines  $4 \times 4$ -Bildes in Hilbertreihenfolge

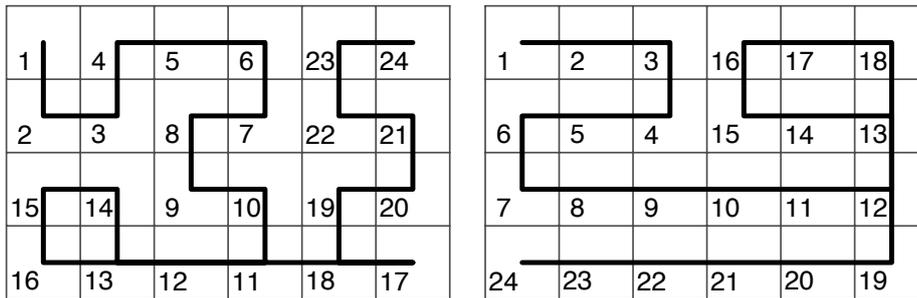


Abbildung 3: Abtastreihenfolge bei Formatanpassung mit Variante 1 und 2 der Hilbertkurve

(vergleiche Abbildung 1, Mitte). Zuerst sammeln wir die Bildpunkte zeilenweise in einen Vektor. (Üblicherweise haben die Bildpunkte diese Reihenfolge bereits in der Bilddatei.) Dann wenden wir die folgende Permutation<sup>1</sup> an: (1, 2, 15, 16, 4, 3, 14, 13, 5, 8, 9, 12, 6, 7, 10, 11). Der Vorteil ist, dass das Bild „nestweise“ abgetastet wird, d. h. in kleinen, dicht beieinander liegenden Gruppen von Bildpunkten. Nahe aufeinander folgend abgetastete Werte weisen nach der Permutation eine stärkere Korrelation auf.

## 2.2 Einen Vektor in ein Bild umwandeln

Der umgekehrte Fall dient der Visualisierung von Vektoren als Fläche, um eine augenscheinliche Untersuchung zu ermöglichen, auch wenn sich der Zusammenhang zu Bildpunkten nicht mehr herstellen lässt oder dieser nie bestand (wie beispielsweise bei Audiodateien). Dafür nehmen wir zum Beispiel einen Vektor der Länge 16 und visualisieren ihn in Hilbertreihenfolge. Zuerst wenden wir die folgende inverse Permutation<sup>2</sup> an: (1, 2, 6, 5, 9, 13, 14, 10, 11, 15, 16, 12, 8, 7, 3, 4). Der entstehende Vektor wird als  $4 \times 4$ -Bild angezeigt.

Bei der Darstellung ist zu beachten, dass die Hilbertkurve nur in den seltensten Fällen vollständig durch Elemente des Vektors genutzt wird. Für einen Vektor der Länge  $n$  benötigen wir  $\lceil \log_2 \sqrt{n} \rceil$  als Rekursionstiefe (bzw.  $\lceil \log_3 \sqrt{n} \rceil$  für die Variante 2 der Hilbertkurve).

## 3 Vergleich verschiedener Abtastvarianten

Wir werden im Folgenden die Eignung der verschiedenen Abtastvarianten für die Steganalyse untersuchen. Dabei sind verschiedene Fragen zu beantworten: Warum nützt eine möglichst starke Korrelation der Steganalyse? Sind rekursive Abtastprinzipien anderen

<sup>1</sup>fette Nummerierung in Abbildung 1 zeilenweise gelesen

<sup>2</sup>kursive Nummerierung in Abbildung 1 entlang der Kurve gelesen

überlegen? Welche der beiden Varianten der Hilbertkurve ist besser?

Die erste Frage ist leicht zu beantworten: Die Bildern innewohnende Struktur bedingt eine Korrelation benachbarter Bildpunkte. Steganographie, insbesondere das Ersetzen niederwertigster Bits durch unabhängige Nachrichtenbits, zerstört bestimmte Strukturen im Bild. Steganalyse mit Statistik höherer Ordnung verwendet die Korrelation benachbarter Bildpunkte, um den Strukturverlust nachzuweisen und dadurch die Anwendung zu erkennen.

Je stärker die Korrelation vor dem Einbetten ist, desto klarer können wir steganographische Spuren entdecken. Die Korrelation hängt von der Abtastreihenfolge ab, insbesondere vom räumlichen (euklidischen) Abstand der Bildpunkte. Die herkömmliche zeilenweise Abtastung ist jedoch bezüglich des Abstands nicht optimal.

Die zeilenweise Abtastung ist die verbreitetste Variante. Um zeilenweise abzutasten brauchen die Werte oft nur nacheinander aus der Bilddatei gelesen zu werden, eventuell nach einem Dekompressionsschritt. Der Nachteil ist, dass an jedem Zeilenende ein Rücksprung zum Beginn der nächsten Zeile erfolgt (siehe Abbildung 4). Zwischen Ende und Anfang

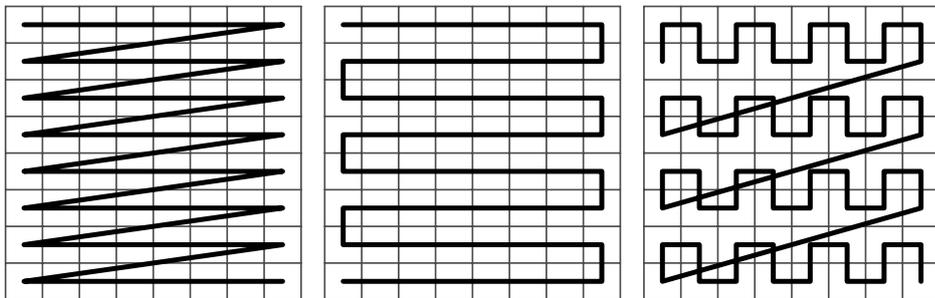


Abbildung 4: Zeilenweise, Slalom- und Mäanderabtastung

einer Zeile ist die Korrelation nur sehr gering.

Die zweite Variante in Abbildung 4 kehrt nach jeder abgetasteten Zeile die Richtung um und fährt so im Slalom bis zum Ende des Bildes fort. Dabei ist der Abstand zum vorhergehenden abgetasteten Wert stets ein Bildpunkt.

Aber auch die Slalomvariante hat einen Nachteil, den sie sich mit der zeilenweisen Abtastung teilt: In einer Linie vorwärts zu gehen bedeutet, sich auf dem kürzesten Weg von den vorhergehenden Abtastwerten zu entfernen. Je schneller die Entfernung größer wird, desto schneller sinkt auch die Korrelation zu diesen Werten. Um sich langsamer von den vorhergehenden Abtastwerten fort zu bewegen, ist eine mäanderförmige Gangart zweckdienlich.

Die Paaranalyse [FGS03] tastet ein Bild mehrfach ab und sammelt Werte in Abhängigkeit von den Bildpunkten in einen Vektor. Abschnitt 4 geht genauer auf die Details der Paaranalyse ein. Abbildung 5 zeigt die Verteilungsdichte der euklidischen Abstände von Bildpunkten, wie sie aufeinander folgend durch die Paaranalyse ausgewertet werden. 536 Graustufenbilder mit  $341 \times 454$  Bildpunkten trugen zu jeder der fünf Kurven bei. Der ma-

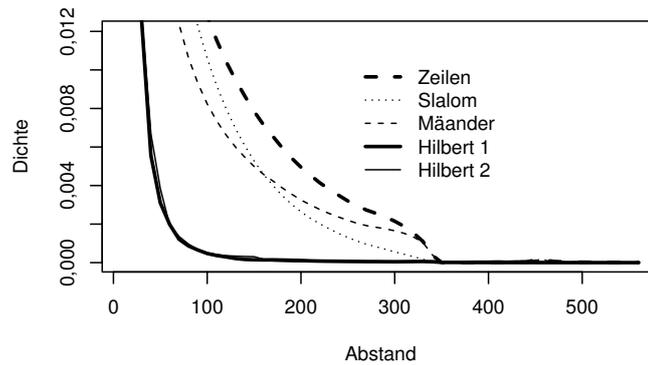


Abbildung 5: Paaranalyse: Verteilung der euklidischen Abstände

ximal mögliche Abstand zweier Bildpunkte ist also  $\sqrt{341^2 + 454^2} = 567.8$ . Das Abtasten entlang einer der Hilbertkurven verschiebt den Schwerpunkt der Dichte zu den kürzeren Abständen. Ein direktes Histogramm der Abstände sähe sehr „gefiedert“ aus, da im Histogramm benachbarte Abstände wie z. B.  $2, \sqrt{5}, \sqrt{8}$  und  $3$  mit stark von einander abweichender Wahrscheinlichkeit auftreten. Um einen Vergleich der Verteilungen zu erleichtern, wurde deren Dichte unter Verwendung eines dreieckigen Fensters mit einer Breite von 20 ermittelt. Tabelle 1 enthält die Anzahl kurzer Abstände und den mittleren Abstand für die verschiedenen Abtastreihenfolgen.

Tabelle 1: Anzahl kleiner Abstände und mittlerer Abstand für verschiedene Abtastreihenfolgen

	= 1	≤ 3	≤ 6	≤ 10	mittlerer Abstand
Zeilen	24 406 584	36 031 059	43 343 341	48 659 685	35,01
Slalom	25 758 228	38 960 877	47 007 036	52 768 368	24,60
Mäander	28 842 260	45 392 457	52 791 135	57 743 123	24,76
Hilbert 1	29 384 613	52 513 152	66 189 063	73 689 384	5,19
Hilbert 2	29 612 378	51 551 806	65 517 080	72 905 013	5,66

In einem weiteren, nun folgenden Experiment wurden 536 kleine Graustufenbilder mit  $213 \times 284$  Bildpunkten untersucht. Gemessen wurde die Autokorrelation in der Folge der abgetasteten Werte  $a_1, \dots, a_n$ . Für den Abstand  $d = 1, \dots, 300$  wurde die Korrelation zwischen den Folgen  $a_1, \dots, a_{n-d}$  und  $a_{1+d}, \dots, a_n$  ermittelt. Die 536 Kurven für jede der fünf verschiedenen Abtastvarianten wurden gemittelt und sind in Abbildung 6 dargestellt. Klarer Sieger ist die Abtastung nach der Hilbertkurve in der ersten Variante, dicht gefolgt von der zweiten. Wahrscheinlich ist ein geringeres Flächenwachstum je Rekursionsebene von Vorteil, zumindest für die vorliegenden Bilder. Die zeilenweise Abtastung hat erwartungsgemäß die am stärksten fallende Korrelation. Deutlich sichtbar ist die Periodizität der Kurve. Nach einem Abstand von einer Zeilenlänge ( $d = 213$ ) hat die Kurve ein

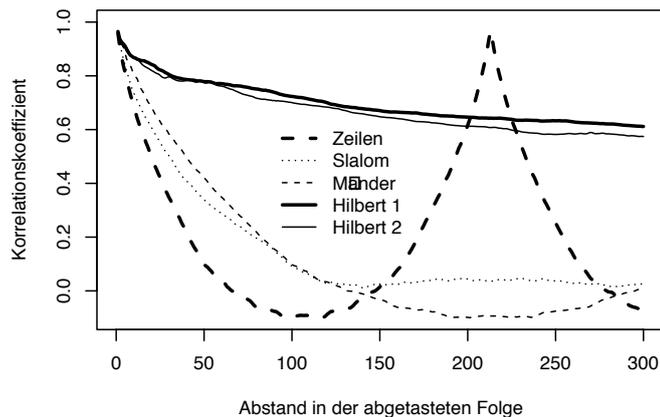


Abbildung 6: Mittelwerte der Autokorrelationskoeffizienten über 536 Testbilder ( $213 \times 284$  Bildpunkte) bei fünf verschiedenen Abtastreihenfolgen

Maximum in der Korrelation. Dann beträgt der Abstand genau ein Pixel an der Längsseite zweier benachbarter Zeilen. Die Slalomvariante fällt zunächst stärker als der Mäander, übertrifft ihn aber weiter hinten – offenbar wegen seines Zeilensprungs, den sie selbst vermeidet. Slalom und Mäander haben eine Periode von zwei Zeilenlängen, denn nach  $d = 426$  Bildpunkten ist der linke Bildrand wieder erreicht.

#### 4 Beispiel Paaranalyse

Eine häufig angewendete Methode zum Nachweis von LSB-Steganographie ist die Paaranalyse. Sie zeichnet sich dadurch aus, dass sie auch bei gleichmäßiger, zufälliger Spreizung der Änderungen über das gesamte Trägermedium die Anwendung von Steganographie erkennt und darüber hinaus die Länge der eingebetteten Nachricht abschätzen kann [FGS03]. Wir wenden diese Methode hier zur Demonstration auf Bilder mit Graustufenpalette an. Wir gehen davon aus, dass jeder Bildpunkt einen von 256 verschiedenen Grauwerten darstellt und diese jeweils in einem Byte kodiert.

Die Paaranalyse konstruiert binäre Vektoren  $z_n$ , ( $n = 0, 1, \dots, 127$ ) für Grauwertpaare  $(2n, 2n + 1)$ , die sich nur im niederwertigsten Bit unterscheiden, d. h.  $\{(0, 1), (2, 3), \dots, (254, 255)\}$ . Zur Erfassung aller dieser Grauwertpaare wird das Bild 128-mal abgetastet. Die Abtastung der Bildpunkte erfolgt in der Reihenfolge, in der die Werte in der Bilddatei abgelegt sind, d. h. zeilenweise. Analog werden weitere 128 Vektoren  $z'_n$  für  $(2n + 1, 2n + 2)$  konstruiert, die dualen Grauwertpaare  $\{(1, 2), (3, 4), \dots, (255, 0)\}$ . Das Bild wird dann zeilenweise abgetastet:

- Wenn der Wert  $2n$  auftritt, wird  $z_n$  um eine 0 erweitert.
- Wenn der Wert  $2n + 1$  auftritt, wird  $z_n$  um eine 1 erweitert.
- Wenn der Wert  $2n + 1$  auftritt, wird  $z'_n$  um eine 0 erweitert.
- Wenn der Wert  $2n + 2$  auftritt, wird  $z'_n$  um eine 1 erweitert.

Dann werden diese jeweils 128 binären Vektoren kombiniert zu zwei Vektoren:

$$\begin{aligned} z &= z_0 z_1 z_2 z_3 \dots z_{127} \text{ und} \\ z' &= z'_0 z'_1 z'_2 z'_3 \dots z'_{127} \end{aligned}$$

( $z'_{127}$  bezieht sich auf die Werte 255 und 0).

Bei einer Überprüfung dieser Vektoren zeigt sich, dass bei der Abtastung eines Bildes ohne eingebettete Nachricht überwiegend homogene Paare (00 und 11) in beiden Vektoren vorhanden sind. Bei Abtastung eines Bildes mit steganographischer Nachricht hingegen nimmt der Anteil inhomogener Paare (01 und 10) in  $z$  zu, während die Paare von  $z'$  zumeist homogen bleiben.

Die erhöhte Zuverlässigkeit wurde experimentell nachgewiesen. Dazu wurden 536 Graustufenbilder von einer digitalen Kamera getestet. Um mögliche Einflüsse der ursprünglichen JPEG-Kompression auszuschließen [FGD01b], sind die Bilder vor dem Test im Verhältnis 5:1 ( $341 \times 454$ ) und 8:1 ( $213 \times 284$ ) verkleinert worden.

Das Diagramm in Abbildung 7 zeigt die ROC-Kurve (Receiver Operating Characteristic), die die Wahrscheinlichkeit Fehler erster und zweiter Art ( $\alpha$  und  $\beta$ , false positives und false negatives) gegenüberstellt. Die gestrichelte Kurve ergibt sich mit unveränderter Paaranalyse (zeilenweise Abtastung), wenn nur 5 % der maximal möglichen Kapazität ausgeschöpft sind. Die durch Volllinie dargestellte Kurve ergibt sich bei der modifizierten Paaranalyse (Abtastung in Hilbert-1-Reihenfolge) für dieselben Bilder.

Die Güte der steganalytischen Methode wird dabei auf zwei Weisen beurteilt: 1. Die von der ROC-Kurve und der Diagonale begrenzte Fläche muss möglichst groß sein. Die Zuverlässigkeit  $\rho$  (das Doppelte dieser Fläche) steigt von 0,58 (gestrichelte Linie) auf 0,73 (Volllinie). 2. Die Wahrscheinlichkeit für einen Fehler erster Art (false positive) sollte höchstens 5 % sein bei einer Wahrscheinlichkeit für einen Fehler zweiter Art (false negative) von 50 %. Sie sinkt von 10,3 % auf 3,9 %. Weitere Ergebnisse sind in Tabelle 2 zusammengefasst.

## 5 Visualisierung

Es gibt einen visuellen Angriff, der die steganographischen Werte optisch darstellt. Dieser Angriff funktioniert jedoch nur, wenn das untersuchte Bild Sättigungsbereiche oder allmähliche Graustufenübergänge aufweist. Neuere mathematische Angriffe erreichen Ergebnisse von hoher Genauigkeit, sind jedoch nicht so anschaulich wie der visuelle Angriff.

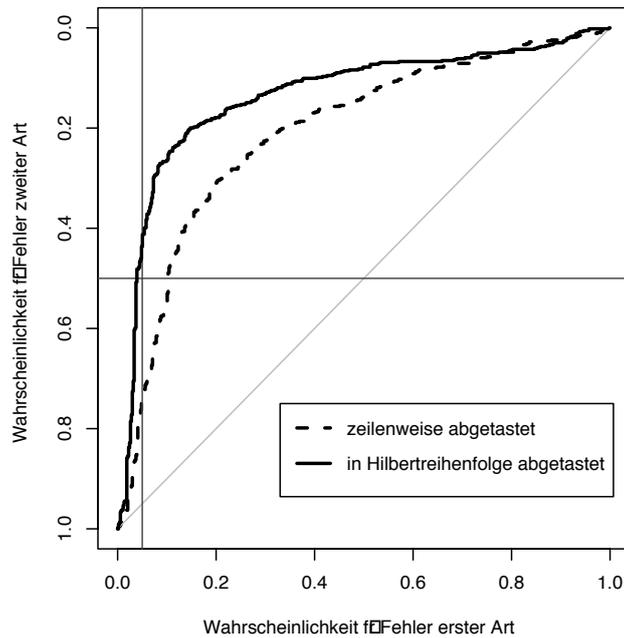


Abbildung 7: ROC-Kurve (Receiver Operating Characteristic), für  $2 \times 536$  Testbilder ( $213 \times 284$  Bildpunkte)

Tabelle 2: Erhöhung der Zuverlässigkeit durch raumfüllende Kurven

Kapazität (%)	341 $\times$ 454 Bildpunkte				213 $\times$ 284 Bildpunkte			
	$\rho$		$\alpha(\beta = 0,5)$		$\rho$		$\alpha(\beta = 0,5)$	
	zw	RFK	zw	RFK	zw	RFK	zw	RFK
1	0,2124	0,2626	0,3321	0,2761	0,1456	0,1989	0,3750	0,3284
2	0,4059	0,4774	0,1922	0,1437	0,2727	0,3794	0,2854	0,1996
3	0,5537	0,6383	0,1157	0,0690	0,3873	0,5292	0,2146	0,1157
4	0,6681	0,7481	0,0728	0,0410	0,4934	0,6463	0,1474	0,0616
5	0,7572	0,8188	0,0485	0,0280	0,5842	0,7308	0,1026	0,0392
10	0,9407	0,9680	0,0093	0	0,8479	0,9155	0,0317	0,0112
15	0,9871	0,9920	0	0	0,9356	0,9785	0,0112	0,0019
20	0,9966	0,9973	0	0	0,9631	0,9903	0,0075	0,0019
50	1	1	0	0	0,9936	0,9960	0,0019	0,0019

$\alpha(\beta = 0,5)$ : Wahrscheinlichkeit für einen Fehler erster Art (false positive) bei einer Wahrscheinlichkeit für einen Fehler zweiter Art von 50 %

$\rho$ : Zuverlässigkeit (doppelte Fläche zwischen ROC-Kurve und Diagonale)

zw steht für zeilenweise Abtastung

RFK steht für Abtastung nach raumfüllender Kurve (Hilbert-1-Kurve)

In diesem Abschnitt werden wir sehen, dass auch mathematische Angriffe – z. B. für didaktische Zwecke – visuell unterstützt werden können.

Wie bereits in Abschnitt 2.2 beschrieben, können die Vektoren  $z$  und  $z'$  visualisiert werden, was wir hier am Beispiel der Bilder mit 0 %, 50% und 100% Einbettung (siehe Abbildung 8) nachvollziehen werden.



Abbildung 8: Bild ohne Nachricht, mit 50 % und 100 % Kapazitätsausnutzung

Der visuelle Angriff nach [Wes99] hilft hier bei der Unterscheidung zwischen Trägermedium und Steganogramm nicht weiter (siehe Abbildung 9).

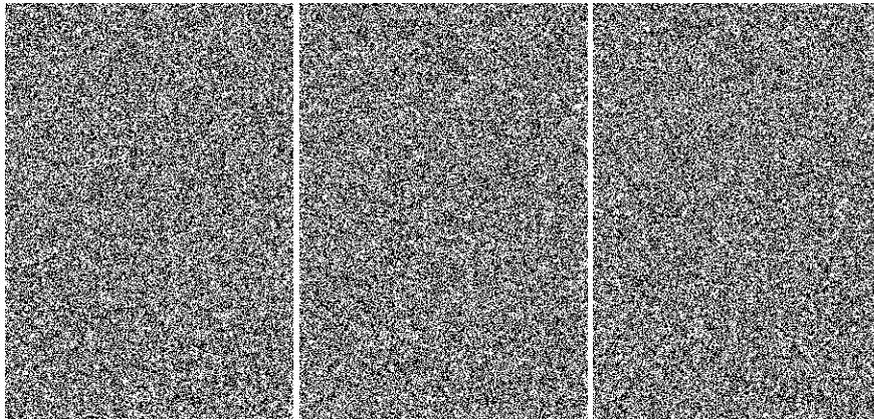


Abbildung 9: Visueller Angriff nach [Wes99] auf Abbildung 8 schlägt fehl

Dennoch können wir in Abbildung 10 sehen, dass die beiden Vektoren für das Trägermedium ohne eingebettete Nachricht eine ähnliche Struktur aufweisen. Um überschüssige

Positionen zu vermeiden, wurde hier nur der Teil der Hilbert-1-Kurve verwendet, der auch beim Abtasten verwendet wurde. Die Ausgabe wurde also analog zu Abbildung 3 auf das ursprüngliche Bildformat eingeschränkt. Nach dem Einbetten einer Nachricht wird die Struktur in  $z$  schwächer und in  $z'$  stärker. Der Anteil homogener Paare in  $z'$  steigt von 53,0% auf 56,5% im Beispielbild. In Hilbertreihenfolge abgetastet steigt er von 53,6% auf 58,0%. Nach dem Einbetten der maximal möglichen Nachricht hat  $z$  einen Anteil von 50,0% homogenen Paaren, gleichgültig ob zeilenweise oder in Hilbertreihenfolge abgetastet. Jedoch ist der Unterschied zwischen  $z$  und  $z'$  deutlicher, wenn die Stichprobenwerte in Hilbertreihenfolge abgetastet werden.

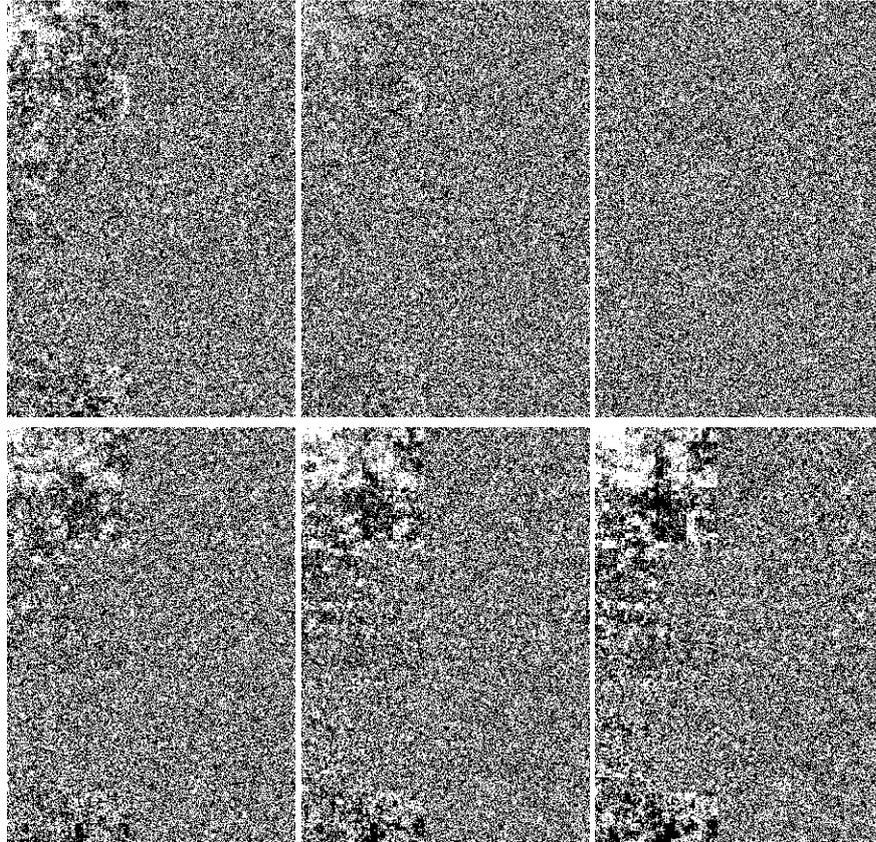


Abbildung 10: Visualisierung der Vektoren  $z$  (oben) und  $z'$  (unten), die bei der Paaranalyse von Abbildung 8 gewonnen werden

Diese Methode der Darstellung ist nicht nur auf die Analyse von Bildern beschränkt. Es ist ebenso möglich,  $z$  und  $z'$  z. B. von Audiodateien zu visualisieren. Obwohl Audiodateien eindimensional sind, lassen sich die resultierenden Vektoren der Paaranalyse in Hilbertreihenfolge darstellen. Eventuell überschüssige Positionen der Kurve sollten im Vektor mit einem anderen, optisch vom Rest zu unterscheidenden Wert aufgefüllt werden.

## 6 Zusammenfassung

Es konnte experimentell gezeigt werden, dass die Korrelation benachbarter Bildpunkte durch rekursive Abtastreihenfolgen besser ausgenutzt werden kann. Dabei werden Fehler erster Art auf ungefähr 1/3 verringert im Vergleich mit der Abtastung in der Reihenfolge, in welcher die Werte in der Bilddatei auftreten. Abtasten entlang einer raumfüllenden Kurve erhöht die Zuverlässigkeit insbesondere auch für kleine Bilder und geringe Änderungsdichten, die gewöhnlich schwer zu erkennen sind.

Steganalytische Größen können durch die inverse Anwendung einer raumfüllenden Kurve visualisiert werden. Obwohl ein früherer visueller Angriff [Wes99] den Unterschied zwischen Trägermedium und Steganogramm nicht aufzudecken vermochte, zeigt die vorgeschlagene Visualisierung die Strukturunterschiede zwischen den von der Paaranalyse generierten Vektoren  $z$  und  $z'$  klar und deutlich.

## Literatur

- [DWW02] Sorina Dumitrescu, Xiaolin Wu und Zhe Wang. Detection of LSB Steganography via Sample Pair Analysis. In Fabien A. P. Petitcolas (Hrsg.), *Information Hiding. Fifth International Workshop*, LNCS 2578, S. 355–372, Berlin Heidelberg, 2002. Springer-Verlag.
- [FGD01a] Jessica Fridrich, Miroslav Goljan und Rui Du. Reliable Detection of LSB Steganography in Grayscale and Color Images. In *Proceedings of the ACM Workshop on Multimedia and Security*, S. 27–30, Ottawa, Canada, 2001.
- [FGD01b] Jessica Fridrich, Miroslav Goljan und Rui Du. Steganalysis Based on JPEG Compatibility. In Andrew G. Tescher, Bhaskaran Vasudev und Jr. Bove, V. Michael (Hrsg.), *Proceedings of SPIE, Multimedia Systems and Applications IV, Volume 4518*, S. 275–280, Denver, CO, 2001.
- [FGS03] Jessica Fridrich, Miroslav Goljan und David Soukal. Higher-order Statistical Steganalysis of Palette Images. In Edward J. Delp und Ping Wah Wong (Hrsg.), *Proceedings of SPIE, Security and Watermarking of Multimedia Contents V. Volume 5020 of Proc. SPIE.*, S. 178–190, 2003.
- [Ker04] Andrew D. Ker. Improved Detection of LSB Steganography in Grayscale Images. In Jessica Fridrich (Hrsg.), *Information Hiding. Sixth International Workshop*, LNCS 3200, S. 97–115, Berlin Heidelberg, 2004. Springer-Verlag.
- [Sau88] Dietmar Saupe. A Unified Approach to Fractal Curves and Plants. In Heinz-Otto Peitgen und Dietmar Saupe (Hrsg.), *The Science of Fractal Images*, S. 273–286, New York NY, 1988. Springer-Verlag New York, Inc.
- [Wes99] Andreas Westfeld. Angriffe auf steganographische Systeme. In Rainer Baumgart, Kai Rannenberg, Dieter Wähler und Gerhard Weck (Hrsg.), *Verlässliche Informationssysteme. IT-Sicherheit an der Schwelle des neuen Jahrtausends*, DuD Fachbeiträge, S. 263–286, Braunschweig Wiesbaden, 1999. Vieweg.
- [WP00] Andreas Westfeld und Andreas Pfitzmann. Attacks on Steganographic Systems. In Andreas Pfitzmann (Hrsg.), *Information Hiding. Third International Workshop*, LNCS 1768, S. 61–76, Berlin Heidelberg, 2000. Springer-Verlag.