



Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG





Agenda

1. Histogram of Oriented Gradients (HOGs)
 - 1.1. Gradientenberechnung
 - 1.2. Histogramme
 - 1.3. Normalisierung
2. Feature Compression
3. Descriptor Design
 - 3.1. Gradient Histogram Binning
 - 3.2. Cell Configuration
 - 3.3. Distanzmaße
4. Descriptor Compression
5. Descriptor Matching
 - 5.1. Compressed Domain Matching
 - 5.2. Accelerated Search Strategiens
6. Vergleich
7. Fazit

1. Histogram of Oriented Gradients (HOGs)

- Darstellung von Aussehen und Form von Objekten durch die Verteilung der lokalen Intensität/ Anordnung von Kanten
1. Gamma-/Farb-Normalisierung
 2. Gradientenberechnung
 3. Unterteilung in Zellen
 4. Berechnung des Histogramms der Kantenorientierung
 5. Ermittlung des Histogramms für einen Block
 6. Normalisierung der Zellen --> Robustheit
 7. Merkmalvektor aus allen Histogrammen

1.1 Gradientenberechnung

$$x - \text{Gradient} = 120 - 70 = 50$$

$$y - \text{Gradient} = 100 - 50 = 50$$

$$\text{Gradient Magnitude} = \sqrt{50^2 + 50^2} = 70,1$$

$$\text{Gradient Angle} = \tan^{-1}\left(\frac{50}{50}\right) = 45^\circ$$

	100	
70		120
	50	



(a) Eingangsbild



(b) Gradientenbild in x-Richtung

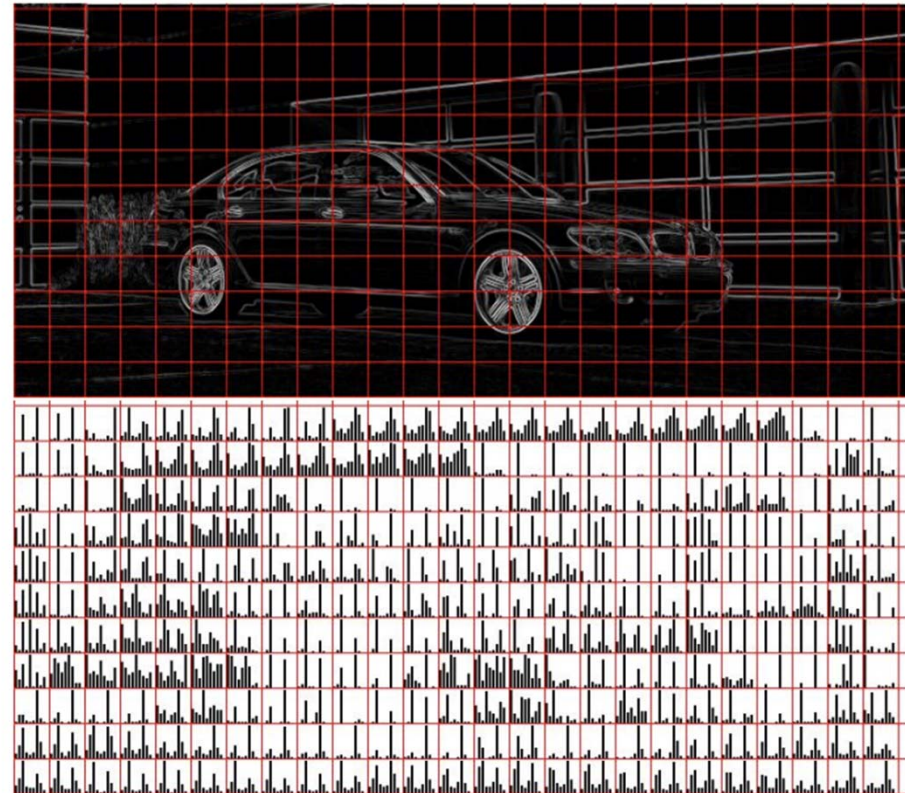


(c) Gradientenbild in y-Richtung

Quelle: Neumann, 2016, s.27

1.2 Berechnung der Histogramme pro Zelle

- Bild wird in Zellen unterteilt
- Für jede Zelle wird das Histogramm der gerichteten Gradienten erstellt
- Die Richtung bestimmt den Bin (Gradient Angle)
- Der Betrag bestimmt das Gewicht (Gradient Magnitude)



Quelle: Neumann, 2016, s.28

1.3 Normalisierung

- Zellen werden in Blöcke zusammengefasst
 - Zelle gehört zu mehreren Blöcken
- Gradienten werden normalisiert
- Normalisierung führt zu einer größeren Robustheit gegenüber Veränderungen
- Alle normalisierten Blöcke zusammengefasst ergeben den HOG-Deskriptor



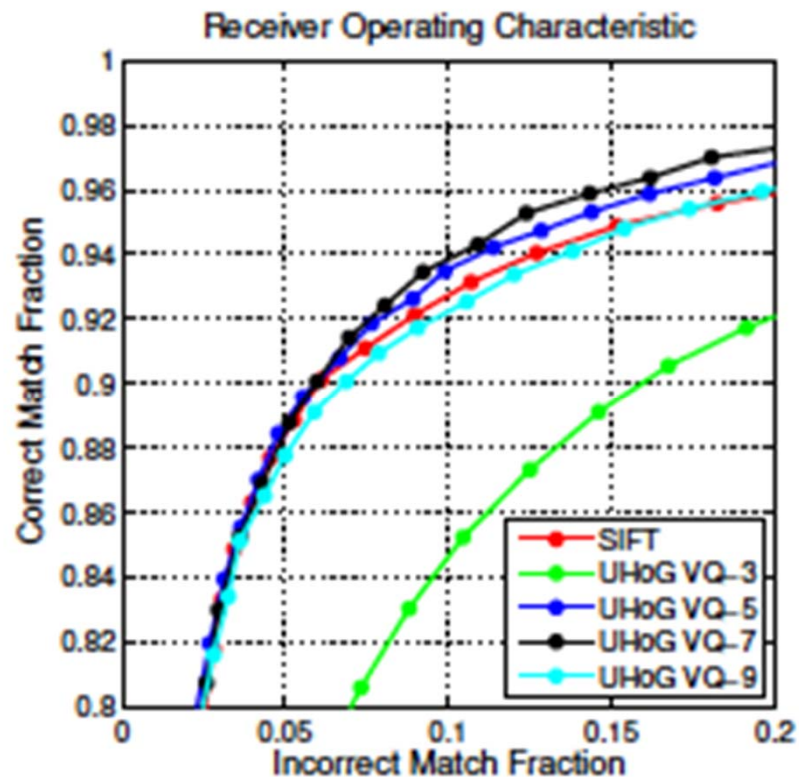
2. Feature Compression

- Übertragungs- und Speicherprobleme beeinträchtigen Berechnungsgeschwindigkeit
- Server-Side Storage
 - Datenbasis verkleinern
- Application Latency
 - Systemlatenz verkürzen in dem weniger Bits gesendet werden können
- Data Transmission
 - Weniger Daten über Wireless Channels und Backhaul Links übertragen

3. Descriptor Design

- Evaluation mit Receiver Operating Characteristic (ROC)
 - Distanz zwischen jedem Deskriptorpaar
 - Histogramme für matching pairs und non-matching pairs
 - ROC Kurve zeigt Sensitivität (y-Achse) und Spezifität (x-Achse)

3.1 Gradient Histogram Binning

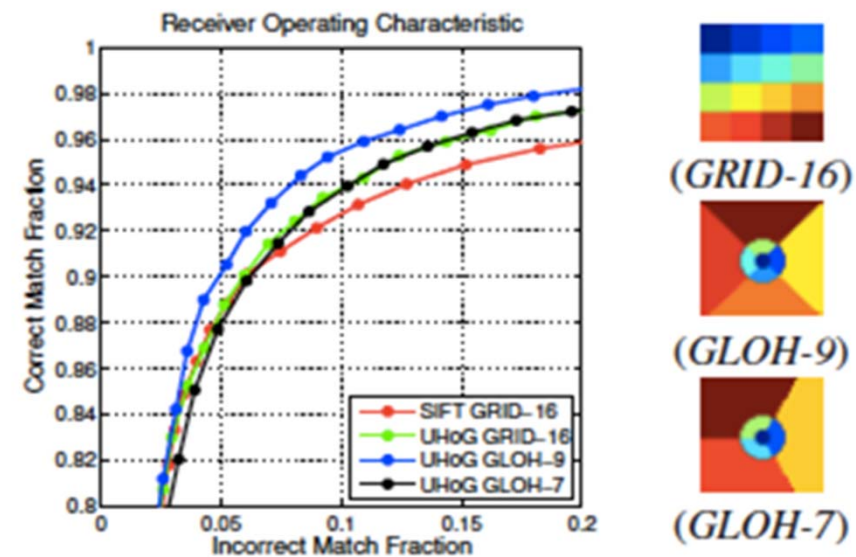


Quelle: Chandrasekhar V. et al, 2009, s.2506

- **Möglichkeiten**
 - VQ-3
 - VQ-5
 - VQ-7
 - VQ-9
- **Ergebnis**
 - Konfiguration mit 5 Bins erzielt ähnliche Performance wie SIFT
 - **VQ-5**

3.2 Cell Configuration

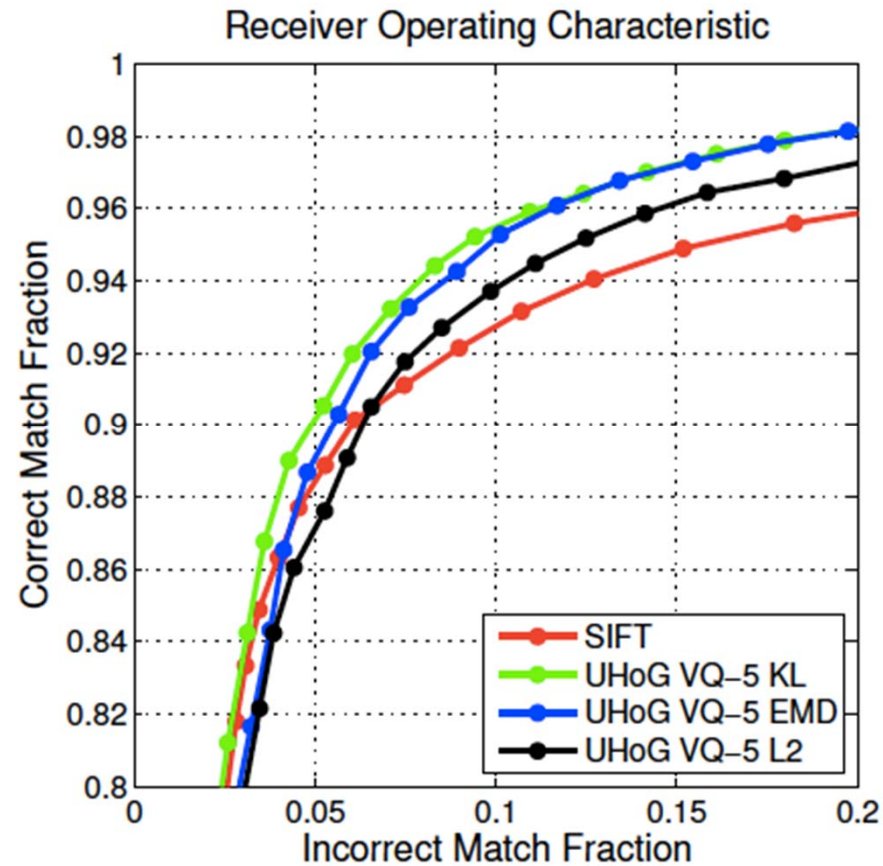
- Weniger Zellen um den kleinstmöglichen Feature Descriptor zu erreichen
- **Möglichkeiten**
 - GRID-16
 - GLOH-9
 - GLOH-7
- **Ergebnis**
 - GLOH-9 performt besser als GRID-16 mit 44% Dimensionsreduktion
 - **GLOH-9**



Quelle: Chandrasekhar V. et al, 2009, s.2507

3.3 Distanzmaße

- Möglichkeiten
 - L_2 – Norm
 - KL Divergence
 - EMD
- Ergebnis
 - KL Divergence



Quelle: Chandrasekhar V. et al, 2009, s.2507

SIFT, SURF, UHoG

$$\mathcal{D}_{SIFT}(i) = \sum_{(d_x, d_y) \in \Omega_i} \sqrt{d_x^2 + d_y^2} P_{D_x, D_y}(d_x, d_y)$$

$$\mathcal{D}_{UHoG}^i = [\hat{P}_1^i, \hat{P}_2^i, \dots, \hat{P}_N^i],$$

$$\mathcal{D}_{SURF}(1) = \sum_{d_x} \sum_{d_y} P_{D_x, D_y}(d_x, d_y) |d_x|$$

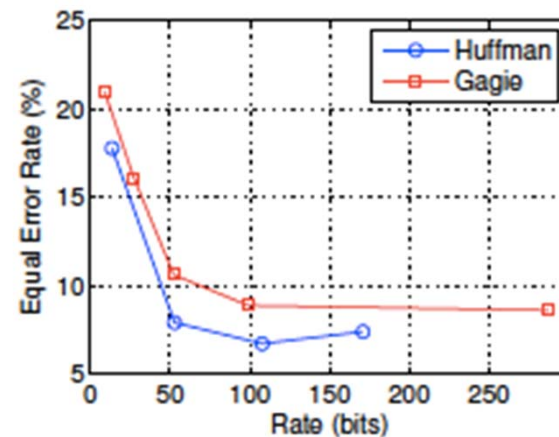
$$\mathcal{D}_{SURF}(2) = \sum_{d_x} \sum_{d_y} P_{D_x, D_y}(d_x, d_y) d_x$$

$$\mathcal{D}_{SURF}(3) = \sum_{d_x} \sum_{d_y} P_{D_x, D_y}(d_x, d_y) |d_y|$$

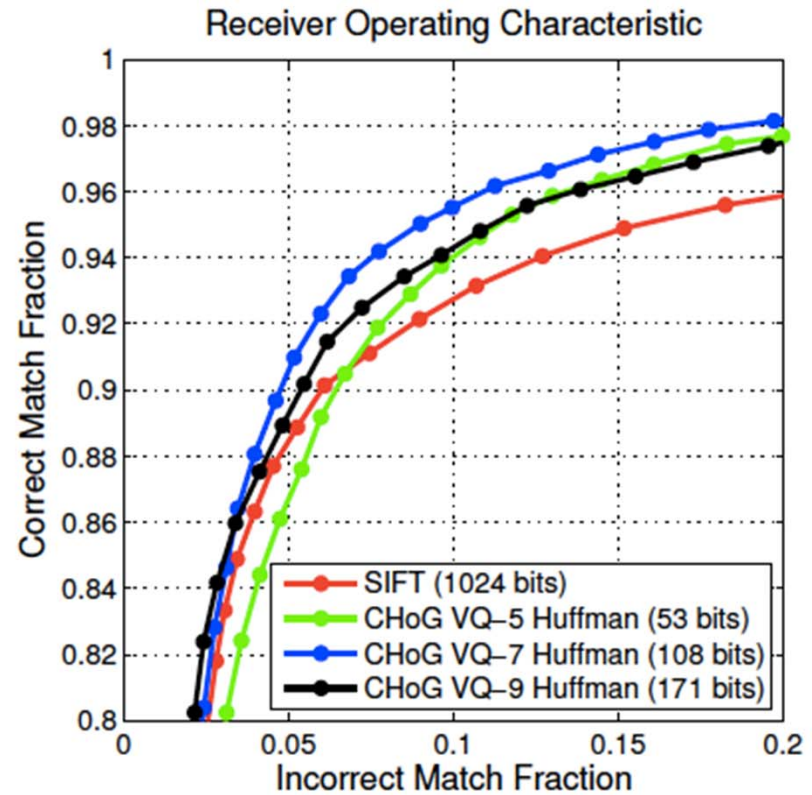
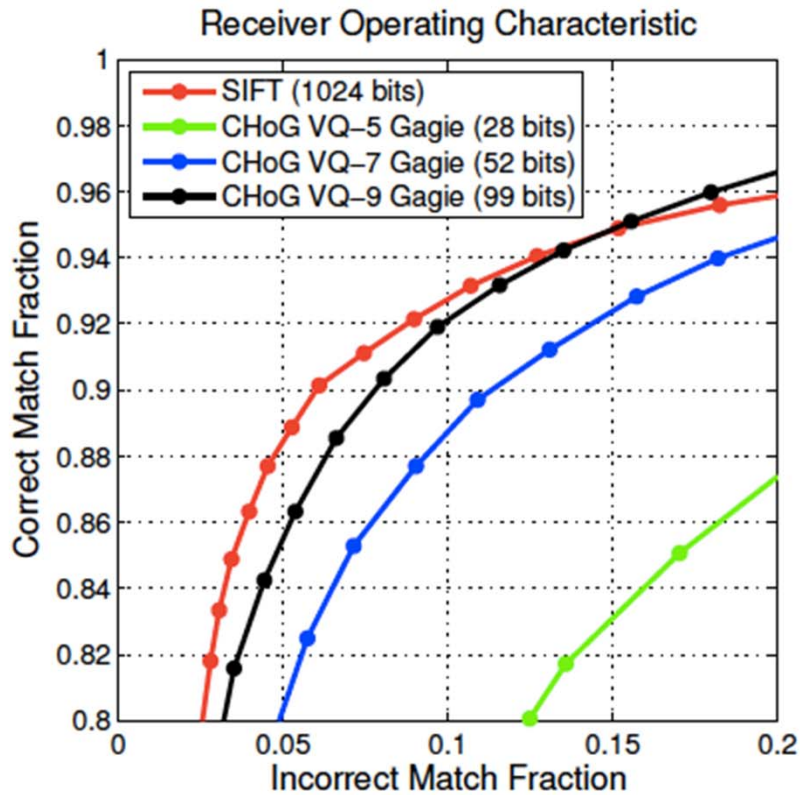
$$\mathcal{D}_{SURF}(4) = \sum_{d_x} \sum_{d_y} P_{D_x, D_y}(d_x, d_y) d_y$$

4. Descriptor Compression

- Histogramm als eine Baumstruktur darstellen um die Datenmenge zu reduzieren
- Vergleich von Huffmann Tree Coding und Gagie Tree Coding
 - Gagie erreicht SIFT bei einer VQ-9 Konfiguration mit 99 bits
 - Huffmann erreicht SIFT schon bei einer VQ-5 Konfiguration mit 53 bits
 - Huffmann weist bei gleicher bit rate eine niedrigere EER auf



Quelle: Chandrasekhar V. et al, 2009, s.2510



Quelle: Chandrasekhar V. et al, 2009, s.2509

→ CHoG VQ-5 Huffman

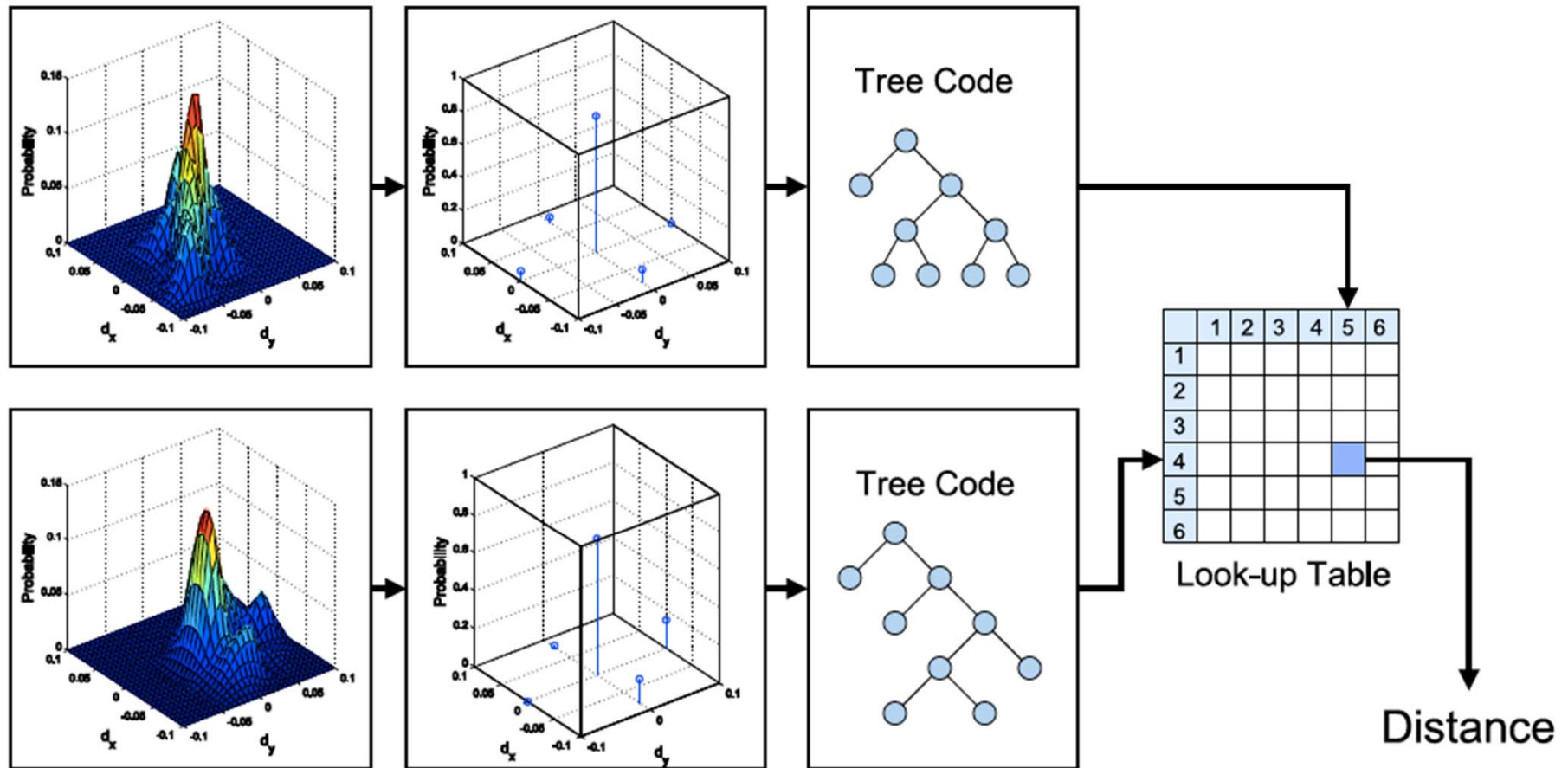


5. Descriptor Matching

- mit den komprimierten Deskriptoren arbeiten
- beim Vergleich muss der Deskriptor nicht zu erst dekomprimiert werden --> niedrigere Komplexität
- 10x schnellere Nearest Neighbor Berechnung

5.1 Compressed Domain Matching

- Berechnung und Speicherung der Distanzen zwischen den komprimierten Distributionen/ Verteilungen
- Baum Index hilft uns beim schnelleren Berechnen der Distanzen zwischen den Deskriptoren
- Distanzen nur noch in Tabellen nachschlagen spart so viel Zeit und Komplexität, dass man effizientere Vergleichsmasse verwenden kann (KL, EMD) ohne zusätzliche Komplexität

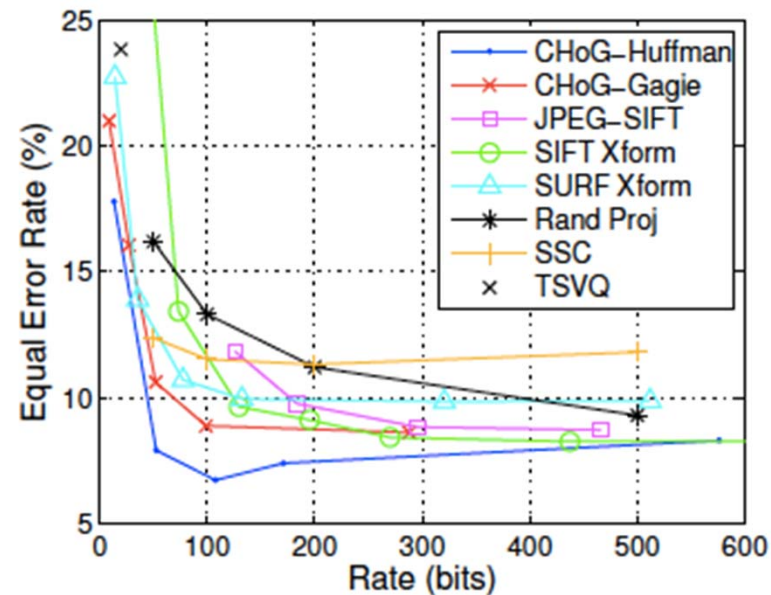


Quelle: Chandrasekhar V. et al, 2009, s.2510

5.2 Accelerated Search Strategies

- Nearest-Neighbor Queries
- CHoG mit 45-dim Deskriptor mit GLOH-9, VQ-5 und Huffmann Tree
Compression Ca. 10x schneller als mit SIFT
 - Niedrigere Dimensionalität
 - Look-up Tabellen für schnellere Distanzberechnung

6. Vergleich



Quelle: Chandrasekhar V. et al, 2009, s.2511

	CHoG-Huffmann	JPEG	Random Projection	Boosting SSC	Transform Coding	TSVQ
EER ~8%	53 bits	295 bits	500 bits	-	SIFT Xform: 270 bits SURF Xform: 133 bits	-

7. Fazit

- Framework für die Berechnung Deskriptoren mit einer 20x Reduktion in Bit Rate, niedriger Komplexität und Matching-Beschleunigung
- Distanzberechnung zwischen komprimierten Deskriptoren ohne Dekodierung
- CHoG zeigt eine bessere Performance als andere Ansätze zur Feature Compression bei niedrigerer oder gleicher Bit Rate
- Dadurch werden Anwendungen ermöglicht, die durch Bandbreite, Speicherplatz oder Latenzzeiten begrenzt werden



Quellen

- Neumann D., Fahrzeugetfassung mithilfe von Stereo-Vision sowie HOG-, LBP-, und Haar-Feature-basierter Bildklassifikatoren, 2016
- Dalal N., Triggs B., Histograms of Oriented Gradients for Human Detection, 2005
- Thang Vu N., Oberkörperdetektion mittels Histogram of Oriented Gradients und Haar-Feature Kaskaden, 2008
- Chandrasekhar V. et al, CHoG: Compressed Histogram of Gradients A Low Bit-Rate Feature Descriptor, 2009

Unter Quantisierung versteht man die Bewertung der Helligkeit (Intensität) eines Pixels mittels einer festgelegten Grauwert- bzw. Farben-Menge, z.B. natürliche Zahlen von 0 bis 255.