Iriserkennung

Der Vergleichstest für den Iriscode

Nils Timm

PADERBORN
CENTER FOR
PARALLEL
COMPUTING

Gliederung

- Motivation
- Iriscode
- Vergleichstest
  - *Hamming Distanz*
  - *Der Test im Feldversuch*
- Statistisches Modell
  - *Herleitung des Modells*
  - *Gültigkeit des Modells*
- Verbesserung des Vergleichstests
  - *Orientierungsinvarianz*
  - *Anpassung des statistischen Modells*
- Schluss
  - *Zusammenfassung*
  - *Ausblick*

Motivation

- Eines der wichtigsten Ziele bei biometrischer Identifikation
  - Minimale Fehlerwahrscheinlichkeiten bei Vergleichstests

→ Funktionsweise des Vergleichstests für Iriscode

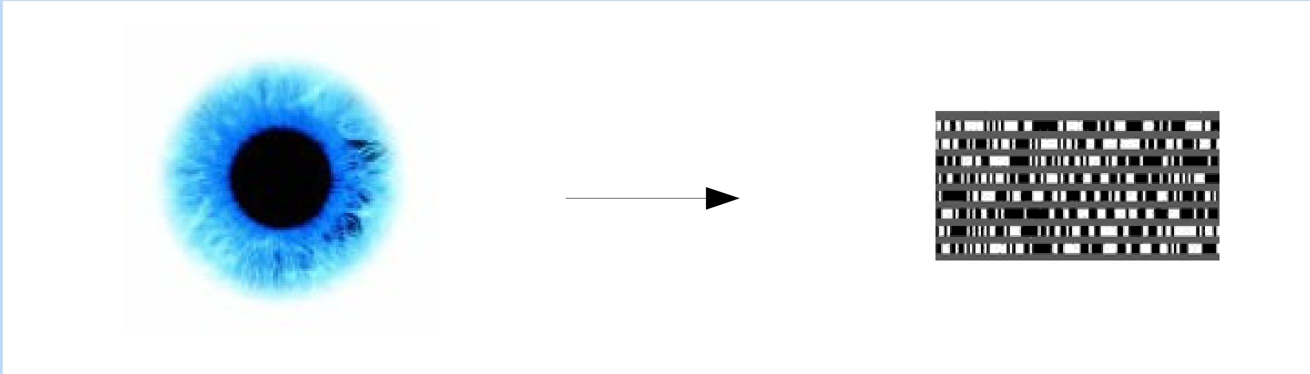
→ Resultate des Vergleichstests im Feldversuch

→ Herleitung eines gültigen statistischen Modells

  → Ermöglicht Fehlerwahrscheinlichkeiten zu prognostizieren

  → Finden einer optimalen Akzeptanzschwelle

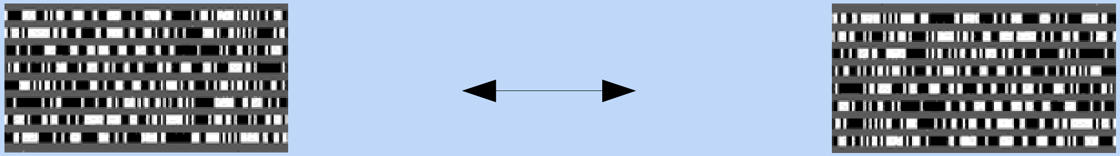
Iriscode



- Irismuster
  - ➔ *Iriscode (2048-Bit Vektor)*

- Verdeckte Bereiche der Iris
  - *Durch Augenlider, Glanzlichtreflexionen, Kontaktlinsenränder*
  - ➔ *Maske (2048-Bit Vektor)*

Vergleichstest
Hamming Distanz
Der Test im Feldversuch



- Prüfung auf statistische Unabhängigkeit von 2 Iriscodes

  ≈ Verschiedenheit der Irismuster

- Realisiert durch Berechnung der Hamming-Distanz (HD)

Hamming-Distanz

- Maß für Unterschiedlichkeit von Zeichenketten
  - In diesem Fall Iriscode Vektoren

- bitweiser Vergleich

- liefert Anzahl nicht-übereinstimmender Bitstellen

  ≈ Unterschiedlichkeit eines Irismuster-Paares

Hamming-Distanz

- Formel:

$$HD = \frac{\|(codeA \otimes codeB) \wedge maskA \wedge maskB\|}{\|maskA \wedge maskB\|}$$

- codeA, codeB
  - Bitvektoren der Irismuster A und B
- maskA, maskB
  - Maskenvektoren von A und B

Hamming-Distanz

- Formel:

$$HD = \frac{\left\| \left(\boldsymbol{codeA} \otimes \boldsymbol{codeB} \right) \wedge maskA \wedge maskB \right\|}{\left\| maskA \wedge maskB \right\|}$$

- nicht-übereinstimmende Bitstellen von A und B

Hamming-Distanz

- Formel:

$$HD = \frac{\|(codeA \otimes codeB) \wedge \boldsymbol{maskA} \wedge \boldsymbol{maskB}\|}{\|maskA \wedge maskB\|}$$

- in beiden Iriscodes gültige Bitstellen

Hamming-Distanz

- Formel:

$$HD = \frac{\left\|\left(codeA \otimes codeB\right) \wedge maskA \wedge maskB\right\|}{\left\|maskA \wedge maskB\right\|}$$

- nicht-übereinstimmende, gültige Bitstellen

Hamming-Distanz

- Formel:

$$HD = \frac{\|(codeA \otimes codeB) \wedge maskA \wedge maskB\|}{\|maskA \wedge maskB\|}$$

- Normierung des Ergebnisvektors auf Anzahl Bits

Hamming-Distanz

- Formel:

$$HD = \frac{\left\| (codeA \otimes codeB) \wedge maskA \wedge maskB \right\|}{\left\| maskA \wedge maskB \right\|}$$

- Normierung der HD auf einen Wert zwischen 0 und 1
  - 0: perfekte Übereinstimmung
  - 1: völlige Unterschiedlichkeit

Hamming-Distanz

- einfache boolesche Operatoren $\otimes$, $\wedge$
- gut in Hardware zu realisieren
- Implementierung in parallelen Blöcken bei Durchmusterung großer Datenbanken möglich
- zeiteffizient implementierbar

Kombinatorik der HD

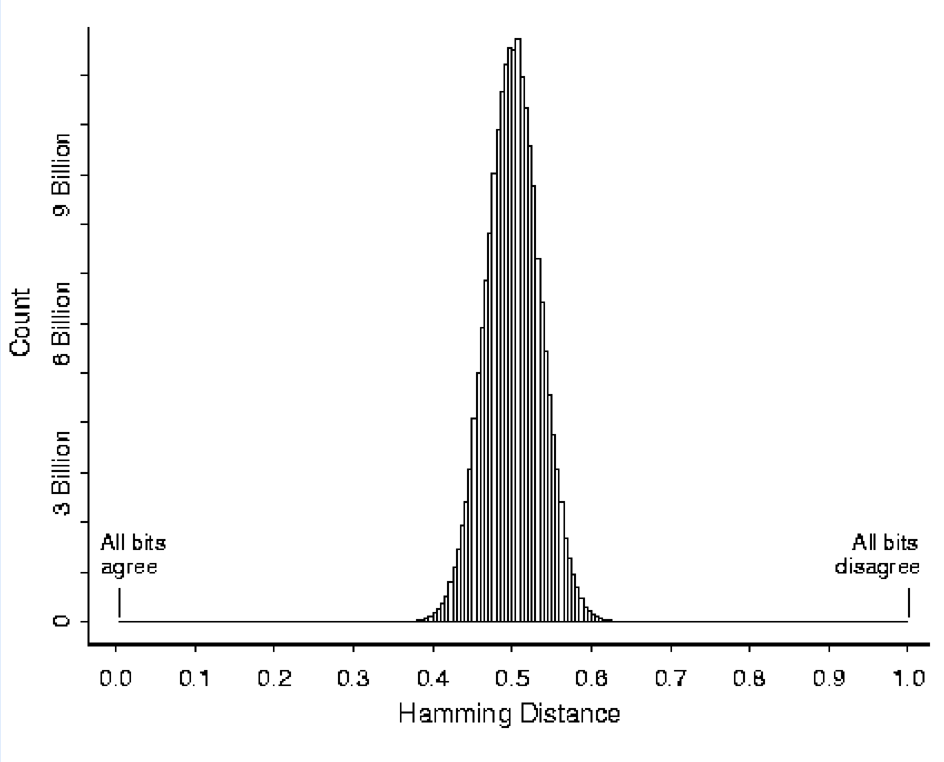
- Eigenschaft den Iriscodes
  - Jedes Bit mit gleicher Wahrscheinlichkeit 0 oder 1

- Keine Korrelation zwischen verschiedenen Irismustern

→ Erwartungwert für HD zweier verschiedener Irismuster ist *0,5*

Der Test im Feldversuch

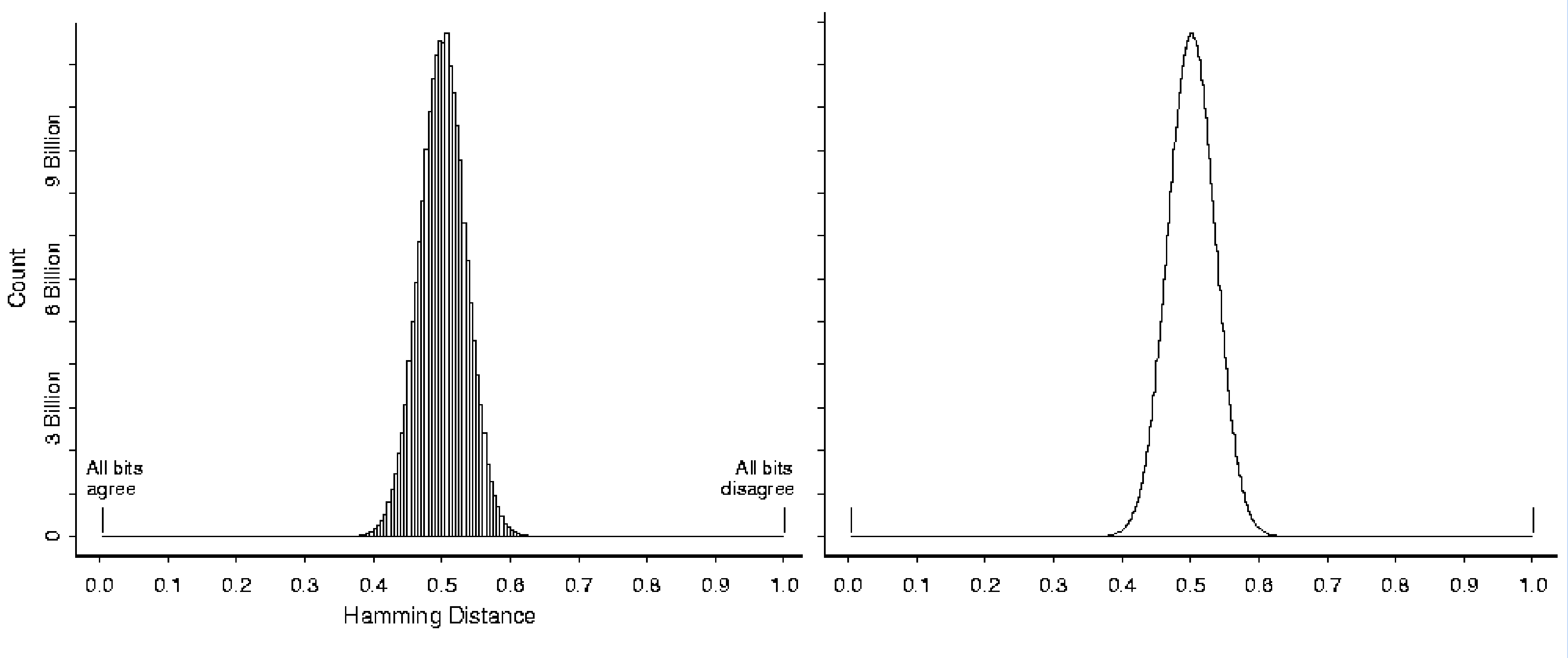
- Feldversuch mit 200 Mrd. Vergleichen verschiedener Irismusterpaarungen:



  - HD Mittelwert: 0,499
  - Standardabweichung: 0,034

Der Test im Feldversuch

- Verteilung entspricht Binomialfunktion mit 249 Freiheitsgraden
  ≈ *Bernoulli Experiment mit 249 Münzwürfen*



Der Test im Feldversuch

- Warum nur 249 Freiheitsgrade und nicht 2048?
  - Interne Korrelationen innerhalb eines Irismusters
  - 249 Bitteilmengen gegenseitig unabhängig

Statistisches Modell
Gültigkeit des Modells

Statistisches Modell

- Bernoulli-Experiment:
  - Trefferwahrscheinlichkeit p = 0,499 (HD Erwartungswert)
  - Standardabweichung $\sigma = 0{,}034$
  - ➔ Wurfanzahl $N = p(1-p)/\sigma^2 = 249$ (Freiheitsgrade)

Statistisches Modell

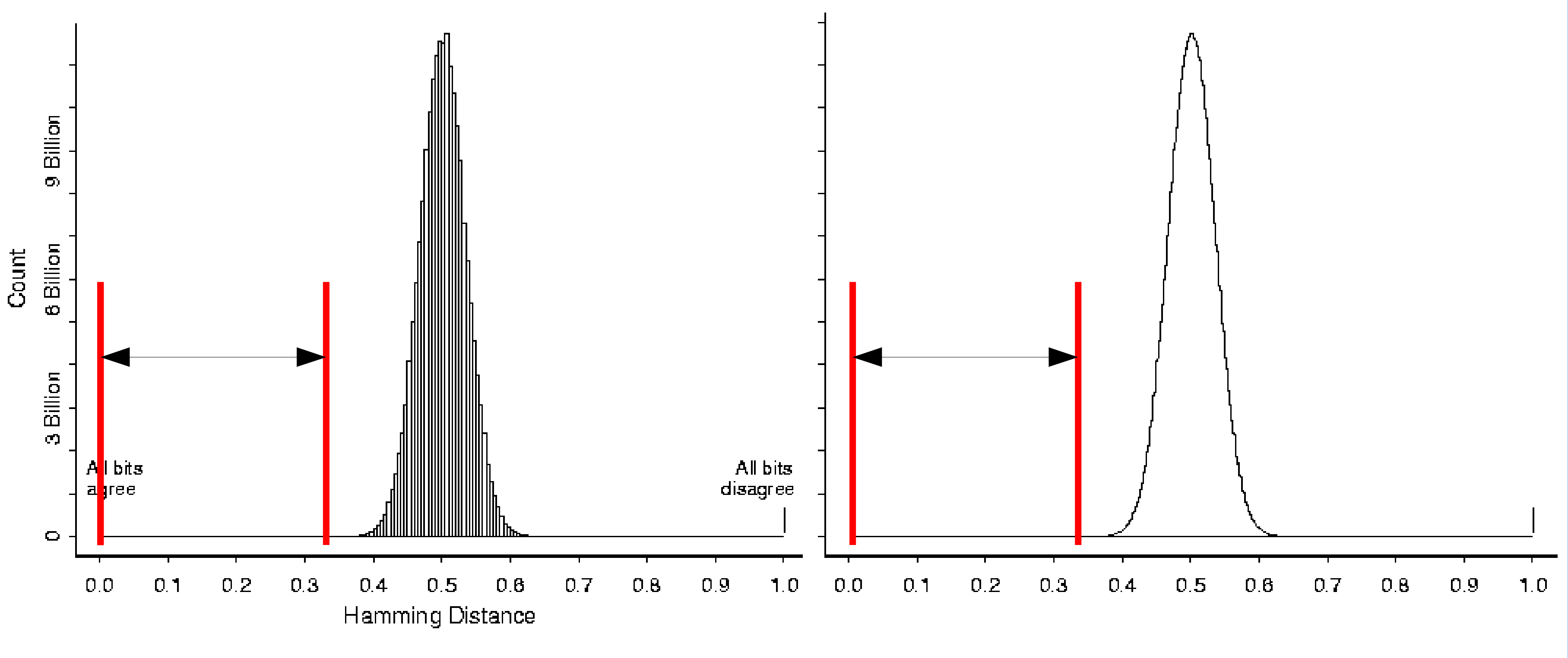
- Binomialfunktion:

$$f(x)=\frac{N!}{m!(N-m)!}p^m(1-p)^{(N-m)}$$

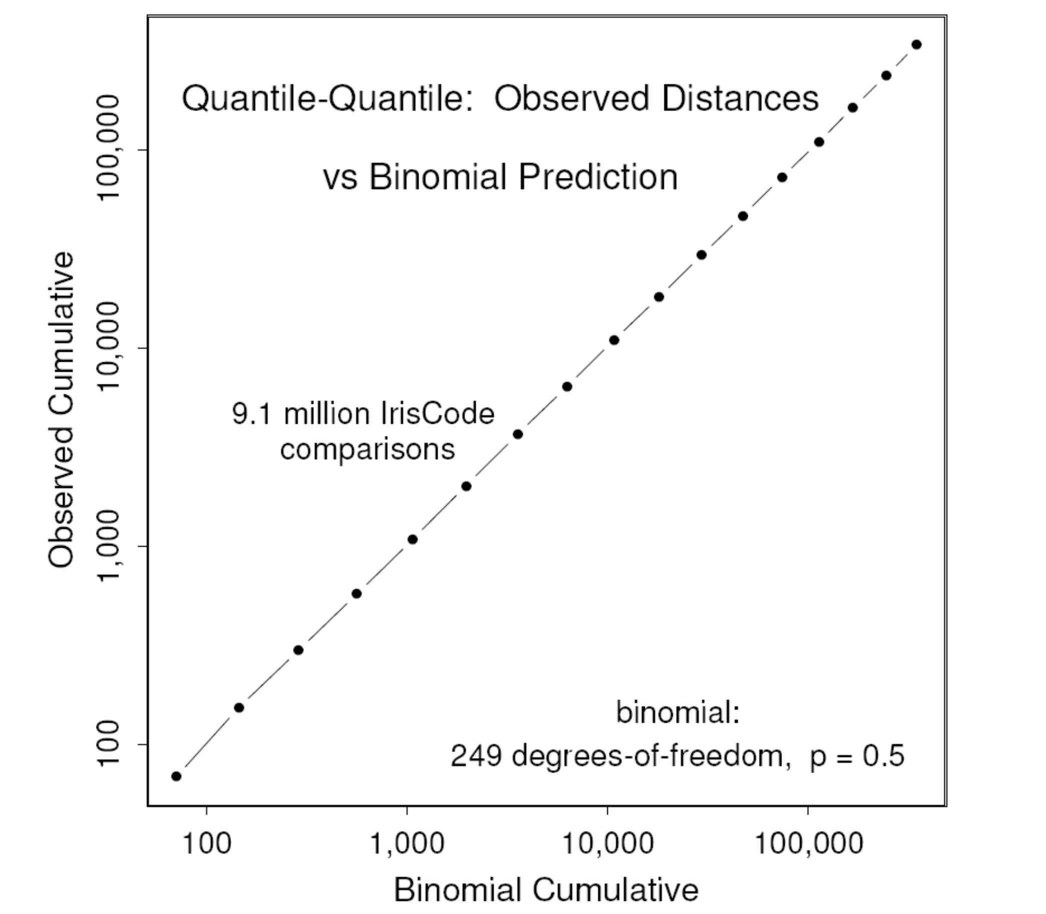
  - mit x = m/N Erfolgsanteil der N Versuche
    - ➔ x entspricht der Hamming Distanz
    - ➔ f(x) entspricht der W'keit, dass ein Vergleichstest HD = x liefert

Gültigkeit des statistischen Modells

- Kumulation der Werte im linken Auslaufbereich
  - f(x): W'keit HD = x bei unterschiedlichem Irispaar
  - F(x): W'keit HD ≤ x bei unterschiedlichem Irispaar



Gültigkeit des statistischen Modells



- Linearität zwischen kumulierten Werten des Feldversuchs und der Binomialverteilung bestätigen Gültigkeit des Modells

Beispielberechnungen

- Wahrscheinlichkeit für HD ≤ 0,333 bei einem Paar verschiedener Irismuster
  - 1 : 16 Mio.

- Wahrscheinlichkeit für HD ≤ 0,300 einem Paar verschiedener Irismuster
  - 1 : 10 Mrd.

- (Kleinste tatsächlich aufgetretene HD bei einem Paar verschiedener Irismuster: 0,334)

→ HD ≤ 0,3 weist stark auf dieselbe Identität hin

Verbesserung des Vergleichstests
Orientierungsinvarianz
Anpassung des statistischen Modells

Verbesserung des Vergleichstests

- Robustheit der Iriserkennung
- Darstellung muss invariant sein gegenüber

  (1) Abbildungsgröße der gesamten Iris
  (2) Pupillengröße innerhalb der Iris
  (3) Drehbewegung des Augapfels
  (4) Kamerawinkel
  (5) Orientierung der Iris bzw. Kopfneigung

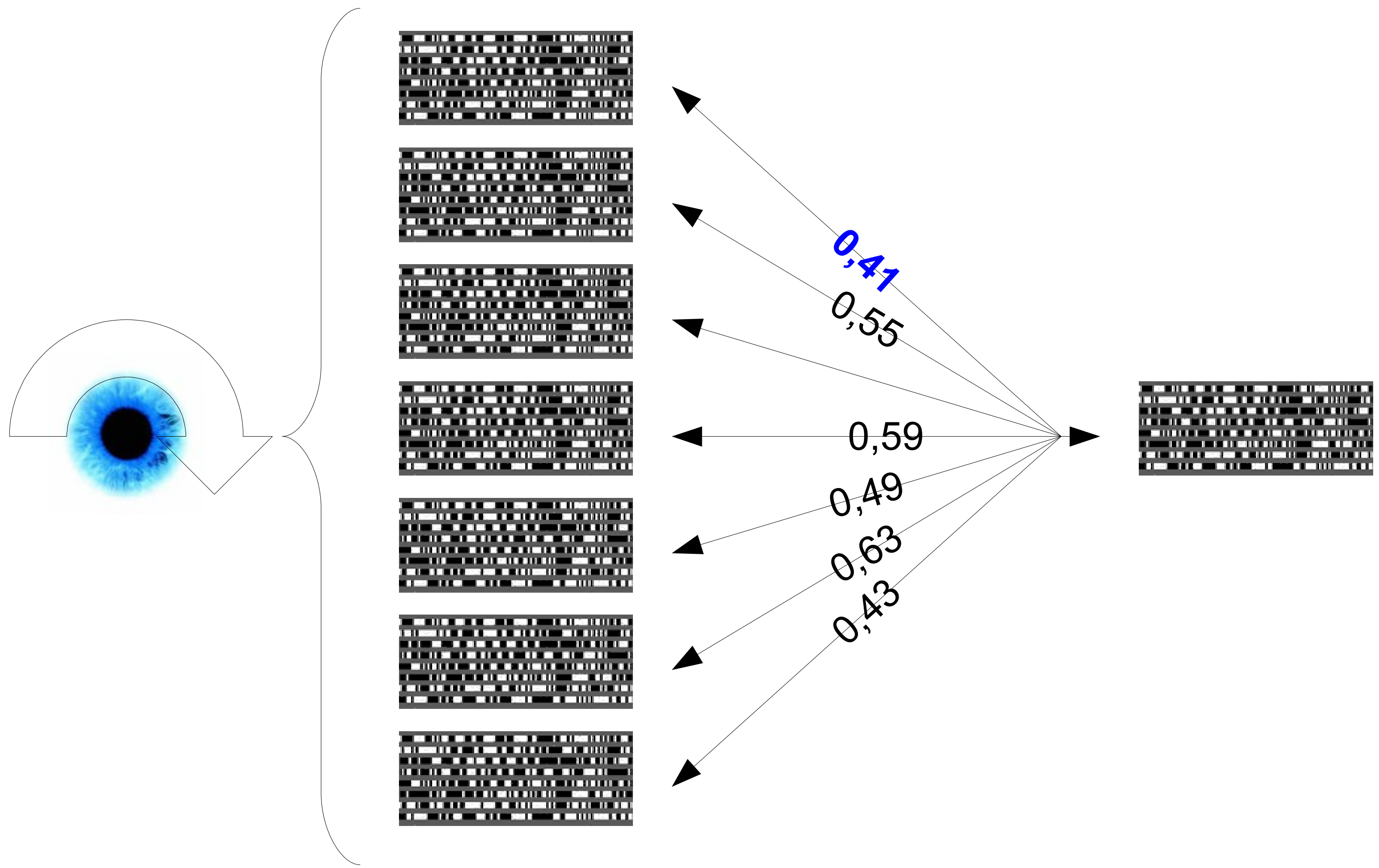
- (1) - (4) werden durch Algorithmus gewährleistet
- (5) lässt sich durch Anpassung des Vergleichstests erreichen

Orientierungsinvarianz

- Orientierung der Iris bei Erzeugung des Iriscodes beschrieben durch Winkelvariable θ
- Zyklische Weiterschaltung von θ
- Vergleich des Iriscodes in verschiedenen diskreten Orientierungen
- HD-Wert der besten Übereinstimmung wird beibehalten

Orientierungsinvarianz

- Üblicherweise Vergeichstests in 7 Orientierungen



Anpassung des statistischen Modells

- Verteilungsfunktion für HD verschiedener Irispaare — $f_0(x)$
- Kumulation von $f_0(x)$ Wahrscheinlichkeit für HD ≤ x — $F_0(x)$
- Gegenwahrscheinlichkeit, HD > x — $1-F_0(x)$
- Wahrscheinlichkeit HD > x bei *n* unabhängigen Prüfungen — $[1-F_0(x)]^n$
- Wahrscheinlichkeit HD ≤ x bei Verwendung der geringsten HD aus *n* Vergleichen — $F_n(x)=1-[1-F_0(x)]^n$
- Neue Verteilungsfunktion durch Ableitung von $F_n(x)$ — $f_n(x)=nf_0(x)[1-F_0(x)]^{n-1}$

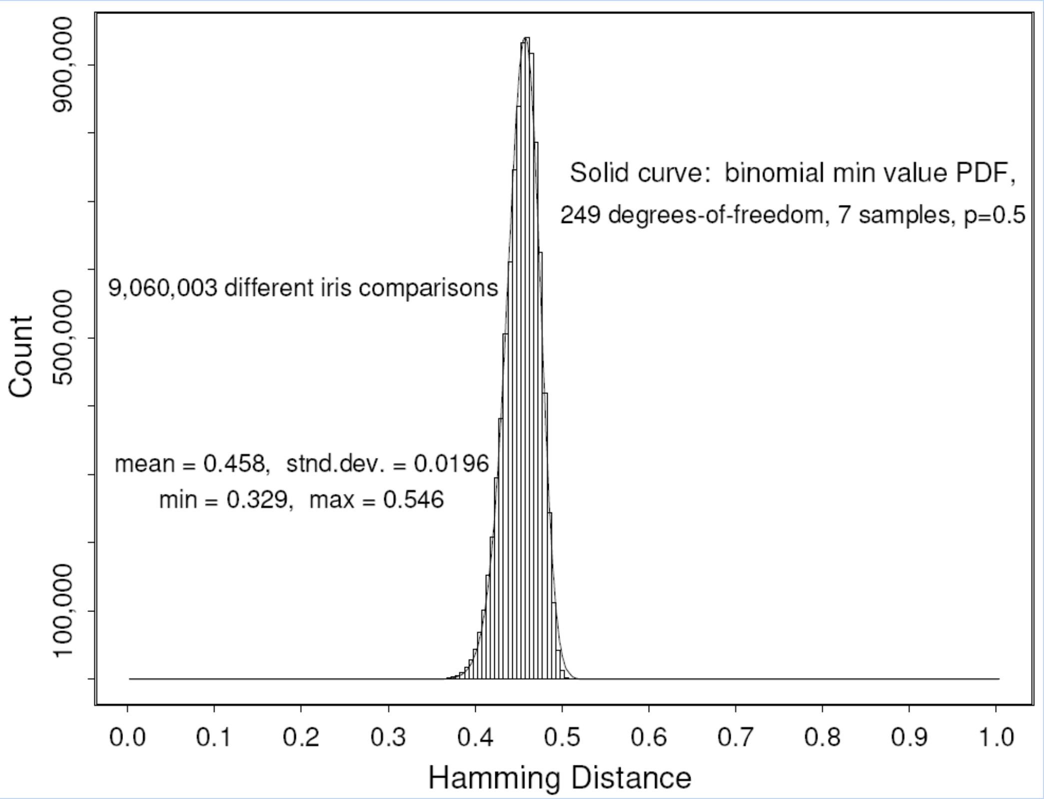
Vergleich mit Feldversuch

- Feldversuch durchgeführt mit Vergleichen in n=7 Orientierungen, jeweils beste Übereinstimmung beibehalten

- Ausgezeichnete Übereinstimmung der Verteilung des Feldversuchs mit neuer Verteilungsfunktion $f_7(x)$

- Linksverschiebung der Verteilung, neuer Mittelwert HD = 0,458

- Kleinste tatsächlich aufgetretene HD: 0,330



Zusammenfassung
Ausblick

Zusammenfassung

- Vergleichstest durch Ermittlung der Hamming Distanz zweier Iriscodes
- Der Vergleichstest für **unterschiedliche** Irismuster im Feldversuch ergibt eine Binomialverteilung der HD
- Herleitung eines gültigen statistischen Modells
- Verbesserung der Iriserkennung durch Orientierungsinvarianz
- Modell ermöglicht Prognosen für W'keiten von Hamming Distanzen unterschiedlicher Irispaare

Ausblick

- W'keit einer Falschübereinstimmung in Abhängigkeit der Akzeptanzschwelle lässt sich anhand des Modells vorhersagen

- Statistisches Modell als Grundlage um optimale Akzeptanzschwellen festzulegen

- Wahl der Akzeptanzschwelle sollte abhängig sein von
  - Datenbankgröße
  - HD Werten bei Vergleichen von Iriscodes identischer Irismuster
  - → Thema des nachfolgenden Vortrags

Fragen

?