

VO „Bildverarbeitung: Biometrische Verfahren“

A. Uhl

Fachbereich Computerwissenschaften
Universität Salzburg



28. Juni 2006

Personalia: A. Uhl

Email-Adresse: uhl@cosy.sbg.ac.at.

Basis-URL: <http://www.cosy.sbg.ac.at/staff/andreas.uhl.html>.

Büro: FB Computerwissenschaften, Zi. 1.11, Jakob-Haringer Str. 2, Salzburg-Itzling.

Telefonnummer (Büro): (0662) 8044-6303.

Telefonnummer (Sekretariat): (0662) 8044-6715.

Formalia

LVA-URL: <http://www.cosy.sbg.ac.at/~uhl/student.html>.

Abhaltezeit der LVA: Mittwoch 9³⁰– 11⁰⁰.

Termine: 8.3. (9:30 Vorbesprechung), dann wöchentlich

Abhalteort der LVA: Hörsaal T02

Vorwort

Willkommen zur Lehrveranstaltung BILDVERARBEITUNG: BIOMETRISCHE VERFAHREN. Es handelt sich um **keine** Überblicksveranstaltung sondern um eine Vorlesung die sich nahe an aktuellen Forschungsthemen bewegt.

Biometrie hat sich zu einem recht großen Gebiet entwickelt, die Ausrichtung dieser LVA richtet sich weitgehend nach meinen eigenen Forschungskompetenzen, die im Bereich Multimedia Signalverarbeitung (i.e. hier Bildverarbeitung) angesiedelt sind.

Es gibt hier im Haus bisher keinerlei Forschungs- und Publikationsaktivität im Bereich Biometrie. In Österreich gibt es bis auf TU Graz kaum Aktivitäten.

Diese LVA soll ein erster Schritt in eine diesbezügliche Richtung sein. Im Lehrbereich kann eine weitere Vertiefung im WS 2006/2007 im Rahmen des SE Visual Computing und Multimedia vorgenommen werden.

Organisatorisches

- Prüfung: Schriftliche Klausur zur Vorlesung. Termin nach Wunsch. Im PS Bearbeitung/Präsentation eines Projektes in Gruppen.
- Skriptum auf meiner Webseite www.cosy.sbg.ac.at/~uhl/student.html zum Download. Im Skriptum wird konsequent Copyright verletzt. Dank an folgende Personen für Material aus ihren Vortragsunterlagen: Andre Drygajlo (EPFL Lausanne), Jean-Luc Dugelay (Institut EuroCom, Sophia-Antopolis), Anil Jain (Michigan State Univ.), Arun Ross (West Virginia University), Josef Scharinger (Univ. Linz), Claus Vielhauer (Univ. Magdeburg), and many more.

Inhalte

- Einführung in Biometrische Verfahren
- Kurzdiskussion von Verfahren mit nicht-Bildverarbeitungen Methoden: Voice Recognition, Keystroke dynamics, Geruch, akustische Ohr Erkennung, On-Line Unterschriftserkennung, Herzrhythmus, Schädelresonanz, EEG - NWA I ...
- Augen: Iris & Retina
- Fingerabdrücke
- Handbiometrie (Palmprint, Hand- und Fingergeometrie, Venegeometrie, Grip pattern, Nagelbett und Nagelrückseite)
- Gesichtserkennung, Lippenbewegung, Ohrgeometrie
- Gang
- Zähne

Literatur

- Bücher (Semesterapparat)

- ★ Biometric Systems (Wayman, Jain, Maltoni, Maio, Springer Verlag 2005)
- ★ Biometric User Authentication for IT Security (Vielhauer, Springer 2006)
- ★ Handbook of Fingerprint Recognition (Maltoni, Maio, Jain, Prabhakar, Springer 2005)
- ★ Handbook of Face Recognition (Li, Jain, Springer 2004)
- ★ Palmprint Authentication (Zhang, Kluwer 2004)

- Zeitschriften

- ★ IEEE Transactions on Information Forensics and Security
- ★ IEEE Transactions on Pattern Analysis and Machine Intelligence
- ★ IEEE Transactions on Image Processing
- ★ Weitere Journals im Bereich Sicherheit & Bild/Videoverarbeitung, Sprachverarbeitung,
- ★ NICHT: Biometrics !!!

Konferenzen, Tagungen, Workshops

Beachte Terminologie: Biometrics vs. Biometric, Biometry !

- IAPR International Conference on Biometrics ICB:
<http://www4.comp.polyu.edu.hk/~icba/>
- Biometrics Symposium 2006 (<http://www.citer.wvu.edu/bsym2006/>)
- Biometrics 2006 (<http://www.biometrics.elsevier.com/>)
- Workshop on Multimodal user Authentication (<http://mmua.cs.ucsb.edu/>)
- IEEE Computer Society Workshop on Multi-modal Biometrics (@ IEEE CVPR
http://www.vislabs.ucr.edu/Multi_Modal_Biometrics/)
- SPIE Biometric Technology for Human Identification (@ SPIE Defense and Security)
- Annual Summerschool (Alghero, Sardinien):
<http://www.computer-vision.191.it/Biomet-School.html>

Inhalt

- Biometrie Einführung,
- Nicht-visuelle Biometrie,
- Offline Unterschriftserkennung,
- Iris Recognition,
- Fingerprints,
- Palmprint, Hand- und Fingergeometrie,
- Face Recognition,
- Retina Scan,
- Gait Recognition.

Table of Contents: Biometrie Einführung

- Grundlagen
- Biometrische Merkmale
- Biometrische Systeme: Systemmodell
- Bewertung von Biometrischen Systemen
- Security und Privacy von Biometrischen Systemen

Begriffe: Was ist Biometrie ?

Begrifflich: aus dem Griechischen – “bios” für Leben, “metros” für messen, Maß

Im weitesten Sinn bezeichnet Biometrie die statistische Analyse von biologischen Beobachtungen und Phänomenen.

Biometrics: automatische Erkennung von Individuen basierend auf biologischen/physiologischen oder verhaltensbasierten Charakteristika

Anthropometrie: Messtechnische verfahren für den menschlichen Körper und seine Teile, z.B. Forensische Antropometrie dient zur Identifikation von Kriminellen mit solchen Messverfahren.

Erkennung der Identität stützt sich klassischerweise auf drei Methoden:

- Besitz-basiert: Identifikation durch etwas, das man hat (z.B. Dokument, Token, Smartcard)
- Wissens-basiert: Identifikation durch etwas, das man weiss (z.B. Passwort, PIN)
- Biometrics-basiert: Identifikation durch etwas, das man ist (menschlicher Körper, biometrischer Identifier)

Nachteile traditioneller Identitätserkennung

Besitztümer können verloren, gestohlen, oder geraubt werden. Wissen kann vergessen, erraten oder in Erfahrung gebracht werden (schlechte Passwörter, schreiben von PINs auf Smartcard!).

Traditionelle Methoden können daher zwischen echter und falscher Identität oft nicht unterscheiden.



“Sorry about the odor. I have all my passwords tattooed between my toes.”

Anwendungsgebiete: Medical Biometrics

Im Bereich Bioinformatik (definiert als Anwendung der Informatik in Biologie und Medizin) wird der Begriff Biometrics oft synonym verwendet für “Biomedical Data Sciences”. In Analogie zu anderen biometrischen Disziplinen werden auch hier Messdaten von biologischen oder medizinischen Phänomenen gesammelt, allerdings mit gänzlich anderer Zielrichtung.

Hier geht es nicht um Identifikation eines *Individuums* sondern eher um statistische Evaluierung von grossen *Populationen*, z.B. Klassifikation von Genen, Proteinen und Krankheiten (z.B. tritt diese oder jene genetische Anomalie auf, so findet sich in 20% der Fälle jene klinische Symptomatik).

Entsprechend anders sind auch die Methoden.

Anwendungsgebiete: Forensic Biometrics

Diese Verfahren sind der eigentliche Ursprung der Biometrie wie wir sie heute kennen. Die Ziele und Methoden sind zwar ähnlich wie im Bereich User Authentication, es gibt aber einen grossen Unterschied:

- User Authentication: biometrische Verfahren dienen einem Benutzer um seine Authentizität zu beweisen
- Forensics: biometrische Verfahren werden von einer anderen Person eingesetzt, um die Zielperson zu identifizieren. Die Perspektive ist hier also eine andere.

Wichtige Beispiele hier sind Fingerabdrücke oder Videoüberwachung kombiniert mit Gesichtserkennung, und alle weiteren Methoden die CSI zur Täteridentifizierung einsetzt.

Anwendungsgebiete: Convenience Biometrics – HCI

Im Bereich HCI geht es darum, durch die Erkennung des Benutzers die Leistung und Genauigkeit der Interfaces zu erhöhen oder einfach die Benutzerfreundlichkeit zu erhöhen.

- Wird der Benutzer erkannt, können z.B. seine sprachlichen Eigenheiten bei Spracherkennung eingesetzt werden und die Erkennungsrate wird erhöht.
- Wird der Benutzer erkannt, können seine persönlichen Einstellungen und Konfigurationen geladen werden.

Hier sind die Methoden die selben wie im Bereich User Authentication, nur sind die Zeile etwas anders.

Anwendungsgebiete: Security Biometrics – User Authentication

Hier geht es darum die Identität des Benutzers zu bestimmen oder zu bestätigen. Ziele von Authentifizierung sind logischer und physischer Zugang zu einer Infrastruktur (access control) oder das Binden von digitaler Information an eine Identität (information authentication).

Access control Beispiele: Fingerprint Scanner an der Haustür, biometrisch abgesicherter Zugang zu Speichermedien, Benutzer Authentifizierung für Rechner oder Rechnernetzwerke,

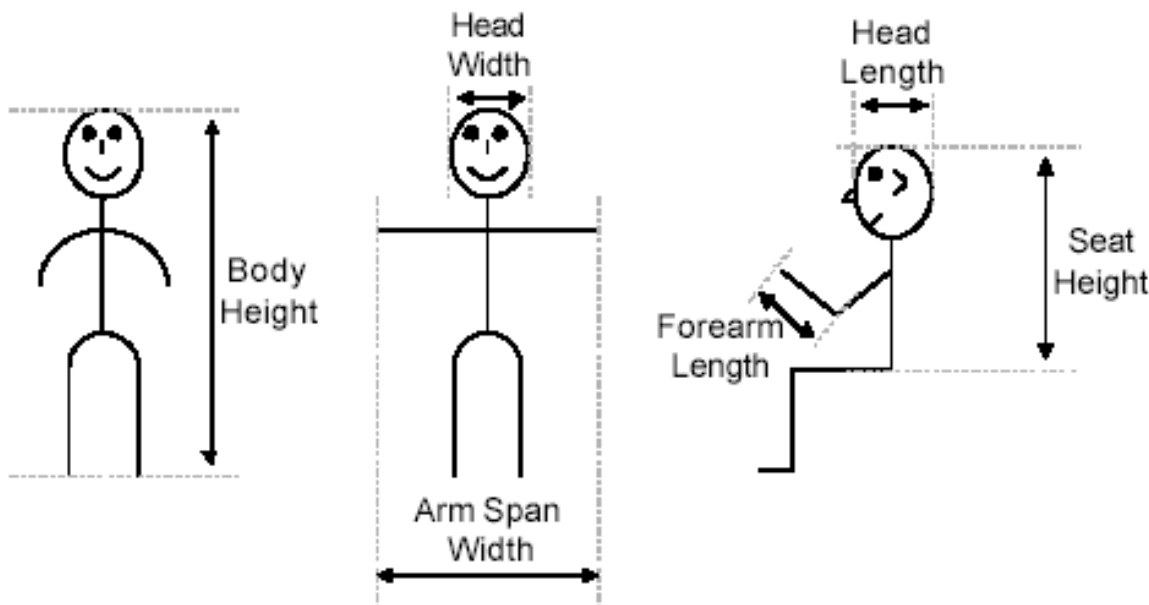
Information authentication Beispiele: Elektronische Signaturen für Dokumente (Biometric Hash), Copyright Schutz für Multimedia Daten, Generierung von Schlüsselmaterial für kryptographische Verfahren, **ACHTUNG:** hier gibts einen wichtigen Unterschied – man muss irgendwie auf eindeutige Bits kommen, Schwellwertverfahren reichen hier nicht !

Im Gegensatz zu anderen Anwendungsgebieten wird Biometrie von den Nutzern hier mit vollem Bewusstsein eingesetzt, um ein bestimmtes Ziel zu erreichen.

Geschichtliches: Alfonse Bertillon

Das Markieren von Gefangenen durch Tätowierungen wurde 1832 in Frankreich abgeschafft. Dadurch entstand das Problem dass Wiederholungstäter identifiziert werden mussten. Das resultierende System war das erste das eine systematische wissenschaftliche Methode zur Personenidentifikation einsetzte.

Neben Körperabmessungen (in Bins klassifiziert) und Augenfarbe wurden charakteristische Bewegungen und spezielle Hauteigenschaften (lokal wie global) archiviert. Nach dem Fall "Will West" durch Fingerabdrücke abgelöst (geschichtliches dazu siehe dort).



Will West's Bertillon Measurements

178.5; 187.0; 91.2; 19.7; 15.8; 14.8; 6.6; 28.2; 12.3; 9.7



William West's Bertillon Measurements

177.5; 188.0; 91.3; 19.8; 15.9; 14.8; 6.5; 27.5; 12.2; 9.6; 50.3

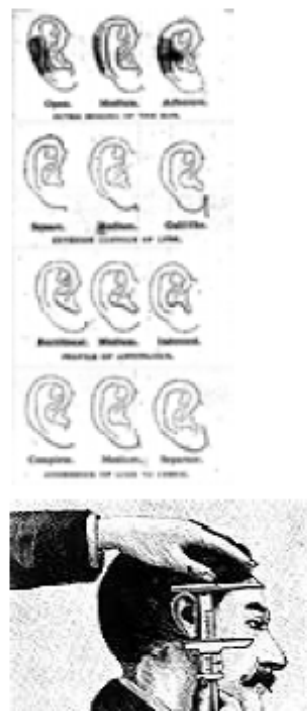
Bertillonage (1882)



(M) (L)

Height 1m 75	Head girth 58.5	L Foot 26.5	Class 188-18	Age 36	Sex M
Weight 75	Head width 16	L Hand 16.5	Index 11	Married age 32	
Forearm 52.5	Chest width 84.5	L Hand 15.5	Forearm 16.5	Profession 188-18	
Forearm 52.5	Forearm girth 32.5	L Hand 16.5	Forearm 16.5	Profession 188-18	

STATE OF NEW YORK.
Office of Superintendent of State Prisons,
BUREAU OF IDENTIFICATION,
Cantons, Albany.



Biometric Authentication: Grundlagen

Ein biometrisches Authentifizierungsverfahren (und wir beschränken uns im Folgenden auf diese Applikation) ist im wesentlichen ein Mustererkennungssystem das eine Person erkennt, indem es die Authentizität einer speziellen physiologischen oder verhaltensorientierten Charakteristik feststellt. Je nach Applikationszusammenhang kann ein biometrisches System ein Verifizierungs- oder Identifizierungssystem sein.

Verifizierung: es wird die Frage beantwortet “bin ich der der ich vorgebe zu sein?”. Die zu verifizierende Person gibt die behauptete Identität an und das aufgenommene biometrische Merkmal (engl. trait oder characteristic) wird mit dem unter dieser Identität in der Datenbank abgespeicherten Merkmal verglichen. Es wird ein 1 zu 1 Vergleich durchgeführt und die Antwort des Systems bestätigt oder verneint die behauptete Identität.

Identifikation: es wird die Frage beantwortet “wer bin ich?”. Ein Individuum wird erkannt indem die gesamte Datenbank nach einem passenden Merkmal abgesucht wird. Es wird ein 1 zu N Vergleich durchgeführt und die Antwort des Systems besteht aus der Identität des Individuums oder einer Fehlschlagmeldung.

Biometric Verification vs. Identification I

Beispiel: Zahlenschloss; ein vierstelliges Zahlenschloss hat 10000 mögliche Einstellungen.

Verification: mein Schloss hat die Kombination 2463. Wenn ich ein Schloss finde und ich will verifizieren dass es meines ist kann ich nachsehen, ob die Kombination 2463 ist. Wenn die Kombination stimmt, ist es wahrscheinlich meines. Die Wahrscheinlichkeit, dass ich mich täusche (also dass es nur zufällig übereinstimmt) ist $1/10000 = 0.00001 = 0.01\%$. Die Wahrscheinlichkeit dass es wirklich meines ist ist damit $1.0 - 0.00001 = 0.9999$ oder 99.99%.

Identifikation: ich habe nun einen Haufen von 10000 Schlössern und möchte meines identifizieren. Ich muss 10000 Tests durchführen, die Wahrscheinlichkeit bei jedem einzelnen Test richtig zu liegen beträgt 0.9999. Um jetzt eine korrekte identifikation durchführen zu können, muss ich bei jedem der 10000 tests richtig liegen, die entsprechende Wahrscheinlichkeit dafür ist $0.9999^{10000} = 0.37$. The Wahrscheinlichkeit das falsche Schloss zu nehmen (also eine Fehlidentifikation zu machen) ist $1.0 - 0.37 = 0.63$!!!!

Und dieses Ergebnis obwohl die Wahrscheinlichkeit korrekt zu sein bei einem einzelnen Test 0.9999 ist !!

Biometric Verification vs. Identification II

Die Wahrscheinlichkeit von Fehlidentifikation steigt schnell mit der Grösse der Datenbank (i.e. der Grösse des Schlösser-Haufens).

1000 Schlösser: $1.0 - 0.9999^{1000} = 0.09$. 10000 Schlösser: $1.0 - 0.9999^{10000} = 0.63$.
100000 Schlösser: $1.0 - 0.9999^{100000} = 0.99995$.

Beispiel: FBI Criminal Master File mit Fingerabdrücken von 50.000.000 Personen.
Welche Genauigkeit wird von jedem einzelnen Vergleich benötigt um eine Identifikation mit 99.99% Zuverlässigkeit durchführen zu können ?

$$X^{50.000.000} = 0.9999$$

$X = 0.999999999998$ (gleiche Fehlerwahrscheinlichkeit für jeden Vergleich angenommen). Das bedeutet: einen Fehler pro 500,000,000,000 Fingerabdrücke.
Vergleich: die Weltbevölkerung beträgt ca. 6,500,000,000.

HOUSTON, WE HAVE A PROBLEM !!!

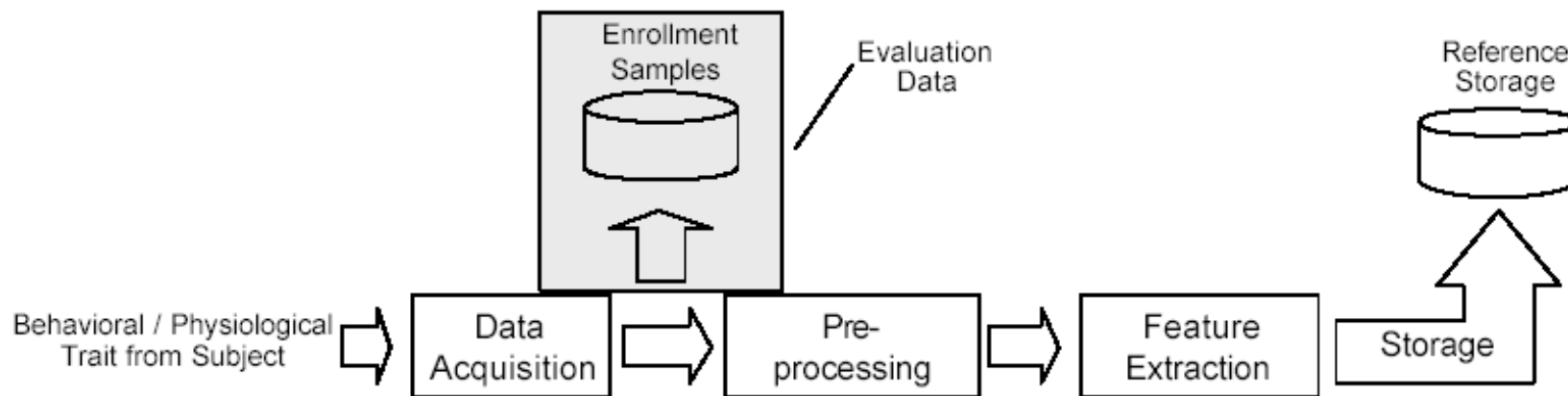
Glücklicherweise ist für viele Applikationen Verifikation ausreichend. Der grundsätzlichen Problematik sollte man sich aber jedenfalls bewusst sein.

Biometric Verification vs. Identification III



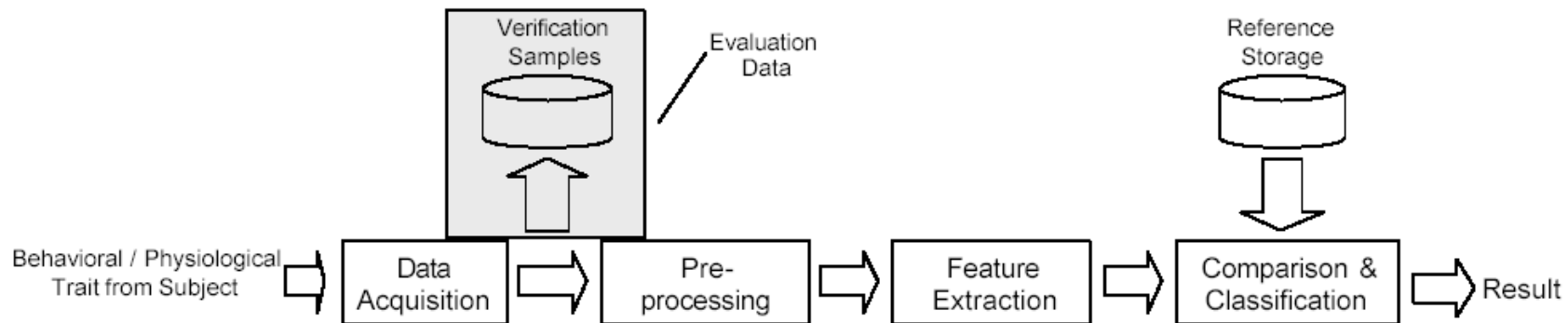
Enrollment vs. Authentication

Alle biometrischen Authentifizierungssysteme operieren in zwei Modi: Zuerst müssen sich alle Benutzer am System registrieren, dieser Vorgang wird als "Enrollment" bezeichnet. Hier werden die Referenzmerkmale für jeden Benutzer im System gespeichert und mit der Identität des Benutzers assoziiert. Die Enrollment samples (= Originaldaten) können eventuell (verschlüsselt und komprimiert) aufbewahrt werden um eine spätere alternative Feature Extraction durchführen zu können. Die Daten werden dann einer Qualitätsprüfung, einer Vorverarbeitung und einer Merkmalsextraktion unterzogen, die biometrischen Merkmale in die Datenbank geschrieben.



Enrollment vs. Authentication

Authentication ist der komplementäre Vorgang, bei dem die Identität des Benutzers verifiziert oder festgestellt wird. Bei der Verifikation wird zusätzlich zum biometrischen Merkmal die behauptete Identität angegeben. Die Verarbeitungsschritte entsprechen bis zur merkmalsextraktion dem Enrollment Prozess, im Anschluss daran werden Vergleiche mit den Referenzmerkmalen in der Datenbank durchgeführt.



Positive vs. Negative Recognition Mode

- Positive Recognition: Hier prüft das System ob die Person diejenige ist die sie (implizit - Identification oder explizit - Verifikation) vorgibt zu sein. Der Zweck von positiver Erkennung ist es zu verhindern, dass mehrere Benutzer die selbe Identität benutzen. Hier soll getestet werden, ob die Verification samples von einer Person stammen, die im System enrolled ist. Es wird also also geprüft, ob diese Person ein Enrollment durchgeführt hat. Beispiele: klassische Zutrittskontrolle z.B. auf Flughäfen.
- Negative Recognition: Hier prüft das System ob die Person diejenige ist die sie leugnet zu sein. Der Zweck von negativer Erkennung ist es zu verhindern, dass eine einzelne Person mehrere Identitäten benutzt. Hier soll getestet werden, ob die Verification samples von einer Person stammen, die nicht im System enrolled ist. Es wird also also geprüft, ob diese Person tatsächlich kein Enrollment durchgeführt hat. Beispiele: “double dipping” – Auszahlung von Sozialhilfe nur wenn die Person nicht in der Datenbank gefunden wird. Hier ist Authentication immer gleich Enrollment !

Während positive Recognition auch mit traditionellen Methoden der Authentifizierung bewerkstelligt werden kann, ist das für negative Recognition nicht der Fall – hier funktionieren nur biometrische Verfahren ! Weiters gibt es Identifikation und Verifizierung nur bei positive Recognition.

Anwendungsszenarien I

- *cooperative vs. non-cooperative*: bezieht sich auf das Verhalten der sich authentifizierenden Person bei der Interaktion mit dem System. Bei einem positive Recognition System liegt es im Interesse des sich Authentifizierenden, sich möglichst kooperativ zu verhalten (z.B. electronic banking). Anders liegt der Fall bei negative Recognition wo es im Interesse des zu Authentifizierenden liegt, nicht zu kooperieren um nicht erkannt zu werden (z.B. Flughafen Applikation um mit Gesichtserkennung Terroristen zu erkennen).
- *overt vs. covert*: wenn sich der Benutzer der Anwendung eines biometrischen Verfahrens bewusst ist, ist es ein overt Verfahren. Gesichtserkennung kann ein covert Verfahren sein (bei versteckten Überwachungskameras), Fingerabdrücke bei Authentifizierung immer overt. Bei forensischen Applikationen ist allerdings auch Fingerabdruck ein mögliches covert Verfahren.

Anwendungsszenarien II

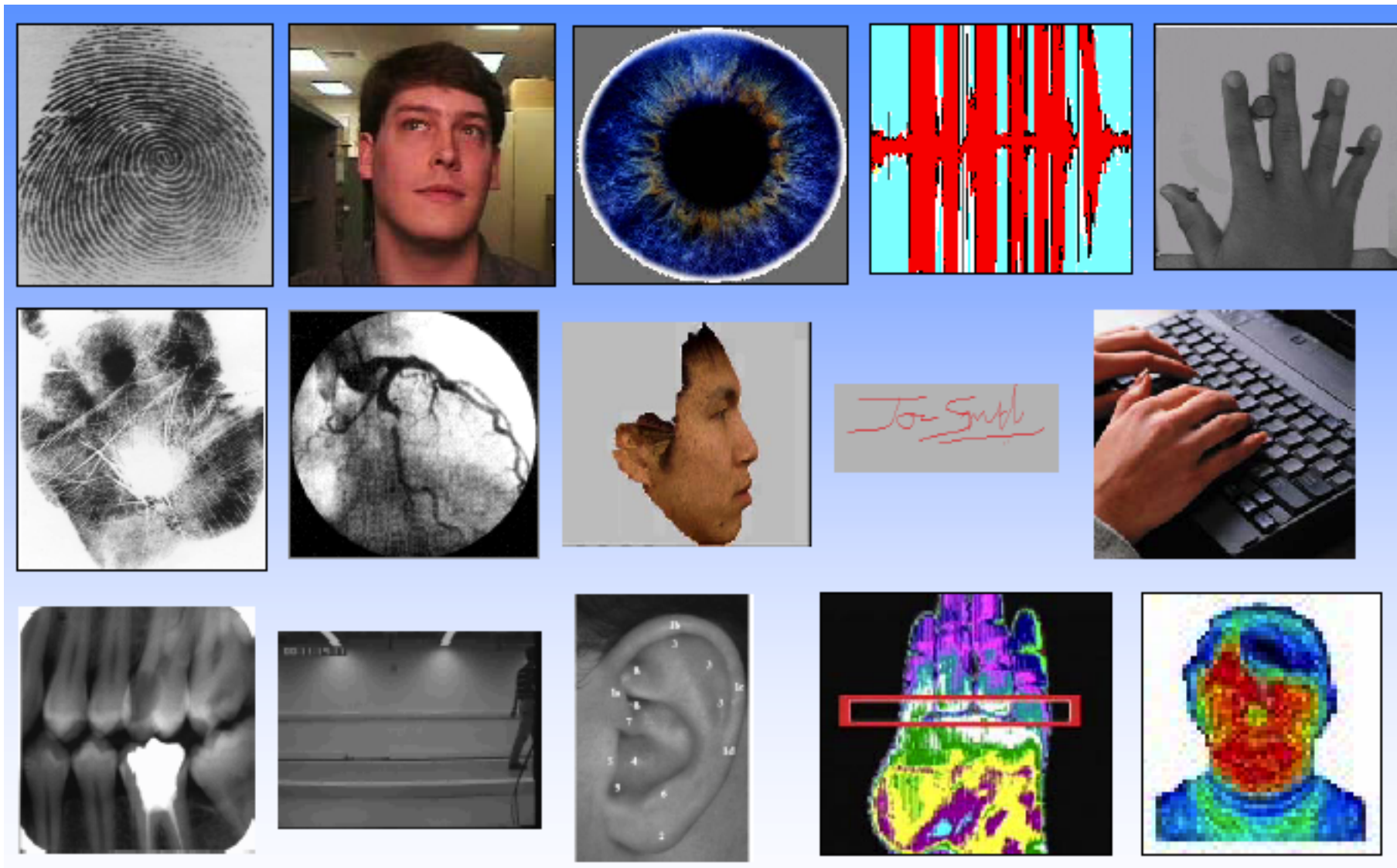
- *habituated vs. non-habituated*: hier geht es darum ob für die Benutzer das Verwenden des biometrischen Systems Routine ist oder nicht. Dies ist für die Erkennungsgenauigkeit wichtig, die mit der Vertrautheit der Benutzer steigt. Zugangskontrolle in einen Arbeitsbereich ist sicher habituated, während eine Führerscheinkontrolle oder Einreisekontrolle wegen der seltenen Ausführung eher non-habituated sein wird.
- *attended vs. non-attended*: bezieht sich auf die Frage ob die biometrische Merkmalsaufnahme beobachtet, geführt oder überwacht wird (z.B. durch einen Angestellten einer Sicherheitsfirma). Ein biometrisches System kann ein attended enrollment (fast alle zumindest im Sinn einer qualitätskontrolle) aber eine unattended authentication haben. Non-cooperative applications müssen de facto attended sein.
- *standard vs. non-standard*: eine standard Umgebung bezieht sich auf eine Anwendung des Systems in einer kontrollierten Umgebung (kontrolliert z.B. bezüglich Temperatur, Druck, Feuchtigkeit, Lichtbedingungen. Häufig sind Indoor Anwendungen standard environments und Outdoor nicht. Klar sind standard environments einfacher zu handhaben, z.B. wechselnde Beleuchtung im Freien bei Gesichtserkennung.

Anwendungsszenarien III

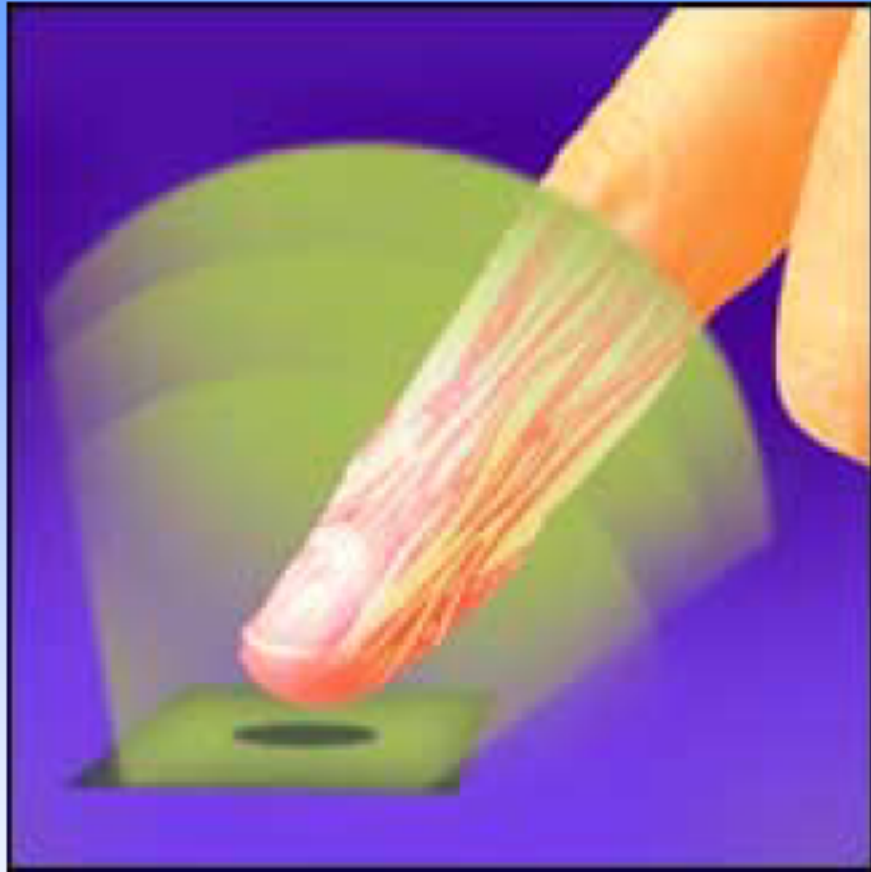
- *public vs. private*: sind die Benutzer des Systems Kunden (public) oder Angestellte (private) der organisation die das biometrische System verwendet ? Die Einstellung bei der Benutzung der Geräte und damit die Leistung hängt vom Verhältnis Endnutzer und Systemmanagement ab. Biometrie bei Führerschein oder Pass ist immer public.
- *open vs. closed*: wird es jemals notwendig sein, dass das System biometrische oder sonstige Daten mit einem anderen System austauscht ? Ein closed System kann proprietäre Dateiformate verwenden, ein offenes muss auf Standards beruhen (wie z.B. dem FBI WSQ Fingerabdruck Kompressionsstandard). Bessere Leistung können aber ev. eigene Systeme bieten (z.B. JPEG2000 part II).

Beispiel: INSPASS (Handgeometrie) zur schnellen Einreise von frequent Travellern in USA (e.g. Kennedy, Newark, LA, Miami, Detroit, ...): positive recognition, cooperative, overt, non-attended, non-habituated, standard, public, closed.

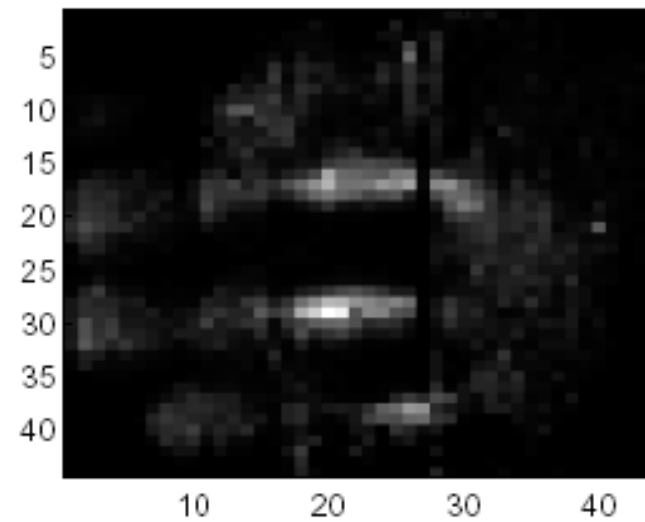
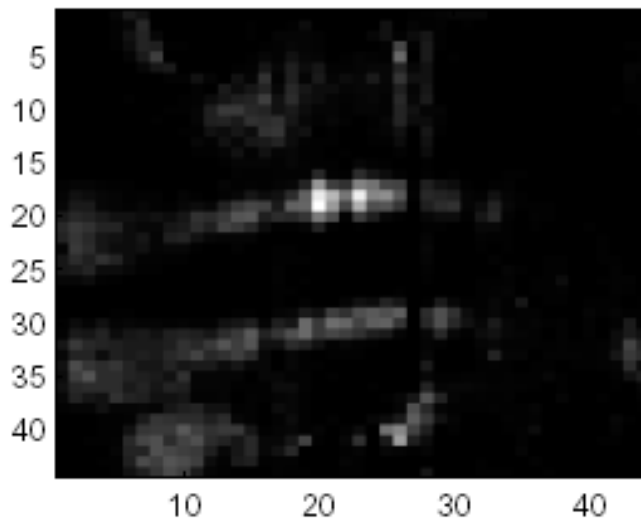
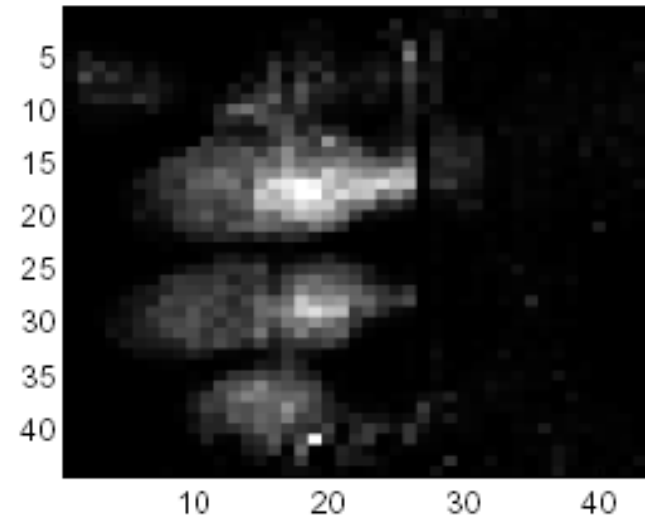
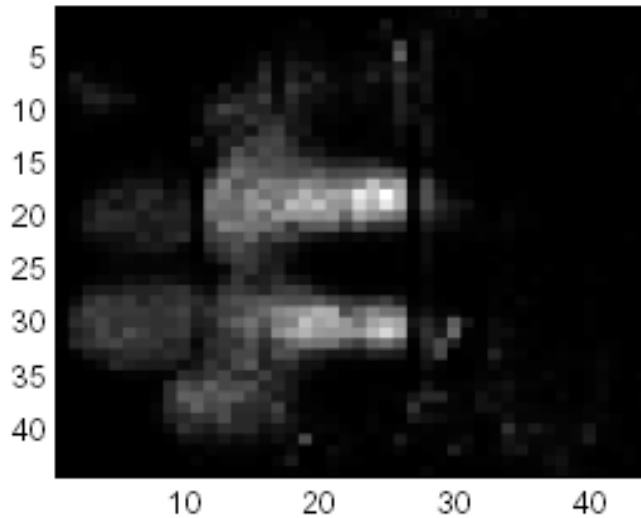
Klassische Biometrische Merkmale (Traits, Features)



Neue Biometrische Merkmale: 3D Fingerblutgefäße



Neue Biometrische Merkmale: Griffmuster



Eigenschaften von biometrischen Merkmalen I

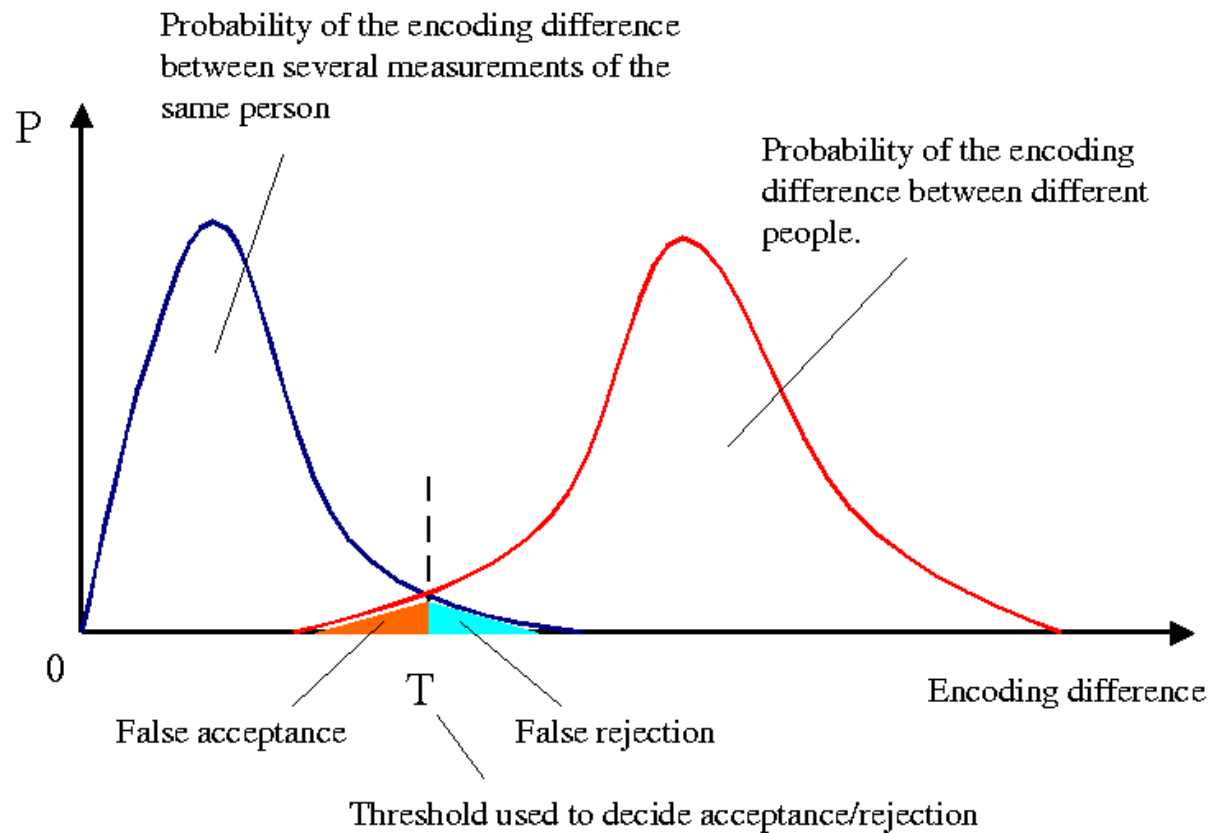
- *Universality*: Jede Person sollte das in Frage kommende biometrische Merkmal besitzen
- *Ascernability - Collectability*: Jedes biometrische Merkmal sollte gut messbar (quantitativ, nicht qualitativ) und zugänglich sein. Dies wird hauptsächlich durch Sensortechnologie festgelegt, aber auch die Vorverarbeitung muss eine bestimmte Mindestqualität garantieren. Hier ist es auch wichtig, dass der Aquisitions- und Verarbeitungsvorgang in vernünftiger Zeit abläuft. DNA Muster sind ein Beispiel die diese Eigenschaft nicht gut erfüllen.
- *Variability - Permanence*: Jedes biometrische Merkmal ist einer natürlichen Veränderlichkeit ausgesetzt (allein schon durch das permanente Absterben und Ersetzen von Zellen), d.h. es gibt Veränderungen von Event zu Event für eine und dieselbe authentische Person. Diese Tatsache wird als Intra-Personal oder Intra-Class (biometrische Verfahren können immer als Klassifikationsprobleme interpretiert werden) variability bezeichnet. Neben dem natürlichen Alterungsprozess sind unterschiedliche Betriebsbedingungen und systematische A/D Wandlungsprobleme weitere Gründe für variability. Klarerweise sollte variability möglichst klein sein um eine ausreichende *permanence* zu erreichen.

Eigenschaften von biometrischen Merkmalen II

- *Distinctiveness*: Ein biometrisches Merkmal muss eine hohe “discriminative power” haben, d.h. es muss sich von Person zu Person deutlich unterscheiden. Das ist gleichbedeutend mit hoher Inter-Personal oder Inter-Class variability.

* Links die Intra-Personal variability, rechts die Inter-Personal variability.

* *Distinctiveness* wird durch die Schnittfläche der beiden Verteilungen bestimmt, die möglichst klein sein sollte. In diesem Bereich kommt es zu falschen Authentifizierungen.



Eigenschaften von biometrischen Merkmalen III

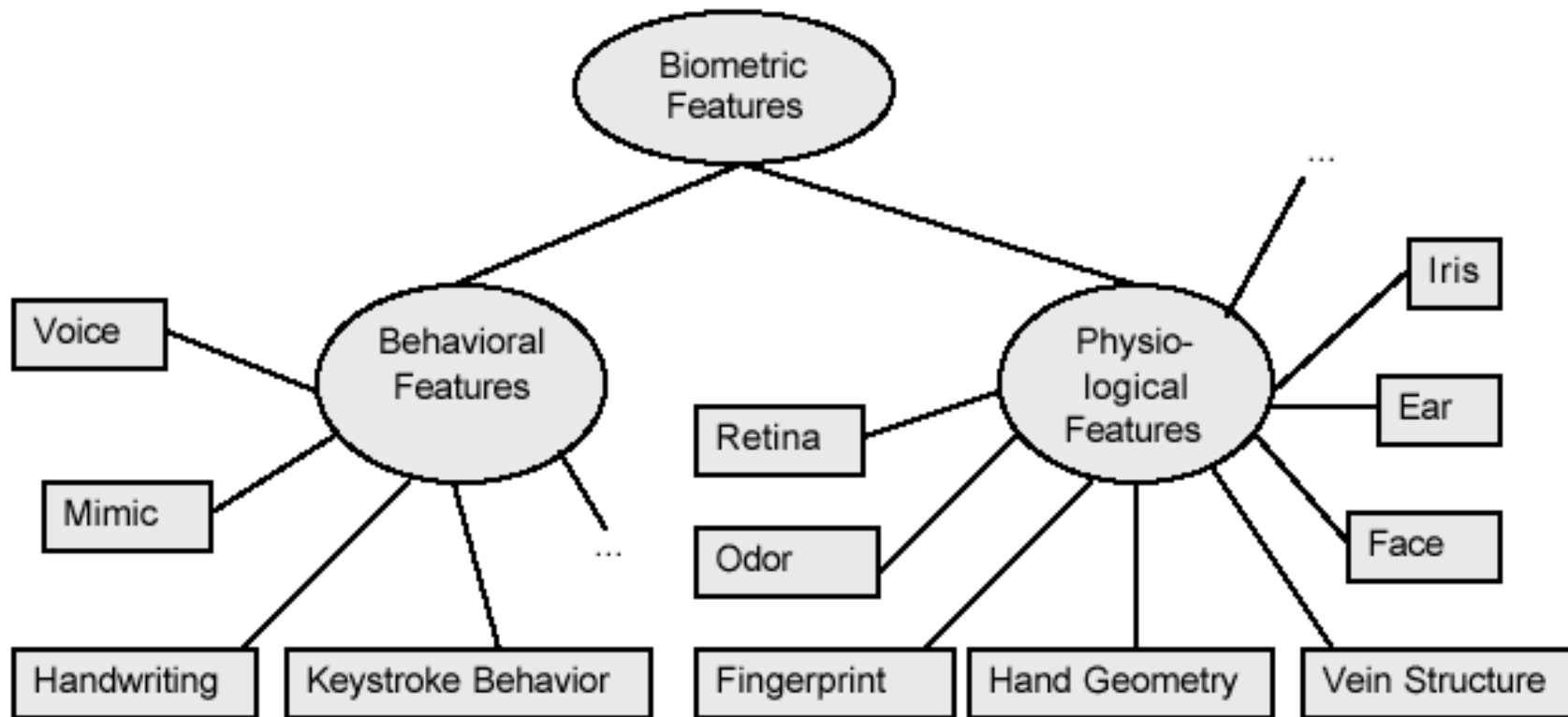
- *Stability*: qualitative Eigenschaft eines Merkmals mit niedriger variability und hoher distinctiveness.
- *Performance*: Geschwindigkeit im Bereich Sensorik, Verarbeitung und Matching, Skalierbarkeit zu grossen Benutzermengen, Erkennungsgenauigkeit, Ressourcenanforderungen um eine bestimmte Genauigkeit zu erreichen,
- *Acceptability*: hier geht es um die Frage ob Personen gewillt sind, sich der Benutzung des Merkmals in der benötigten Häufigkeit zu unterziehen. Entnahme von Hautproben wird wohl eher niedere Werte aufweisen. Retina Scan hat hier eher schlechte Werte durch aufwändigen Scanprozess.
- *Circumvention*: Wie einfach oder schwierig ist es das System zu täuschen oder anzugreifen ?

Aquisition von biometrischen Merkmalen

- Physisches Entfernen von organischem Material: das ist im forensischen Bereich weit verbreitet. Klassisches Beispiel sind Haarteile, Speichelproben, Samenflüssigkeit, Hautteile, etc. zur DNA Analyse. Im Bereich User Authentifizierung wird das eher nicht verwendet, da die Extraktion der DNA Sequenzen biochemisch geschieht und zeitaufwändig ist. Ausserdem kann organisches Material auch verloren werden und damit von anderen Personen illegitim weiterverwendet werden.
- Verhalten der Person: ist durch drei Wesentliche Faktoren bestimmt: **Biologische Beschaffenheit** der Organe die das Verhalten generieren, **gelernte Charakteristika** wie das Verhalten zu generieren ist, und die **Intention** mit der das Verhalten gezeigt wird.
- Physiologische Personenmerkmale: sind individuelle biologische Strukturen die z.B. mit optischen Methoden aufgenommen werden können ohne physische Proben zu nehmen.

Behaviour vs. Physiology: Prinzip

Bei der Aquisition von biometrischen Merkmalen basierend auf physiologischen Eigenschaften kann die sich authentifizierende Person inaktiv bleiben. Diese Merkmale werden daher auch als *passive* Biometrics bezeichnet. Im Gegensatz dazu muss bei der Aquisition von verhaltensbasierten Merkmalen Aktivität gezeigt werden, daher werden solche auch als *active* Biometrics bezeichnet. Dies hat z.B. wichtige Implikationen bzgl. cooperative oder covert environments.



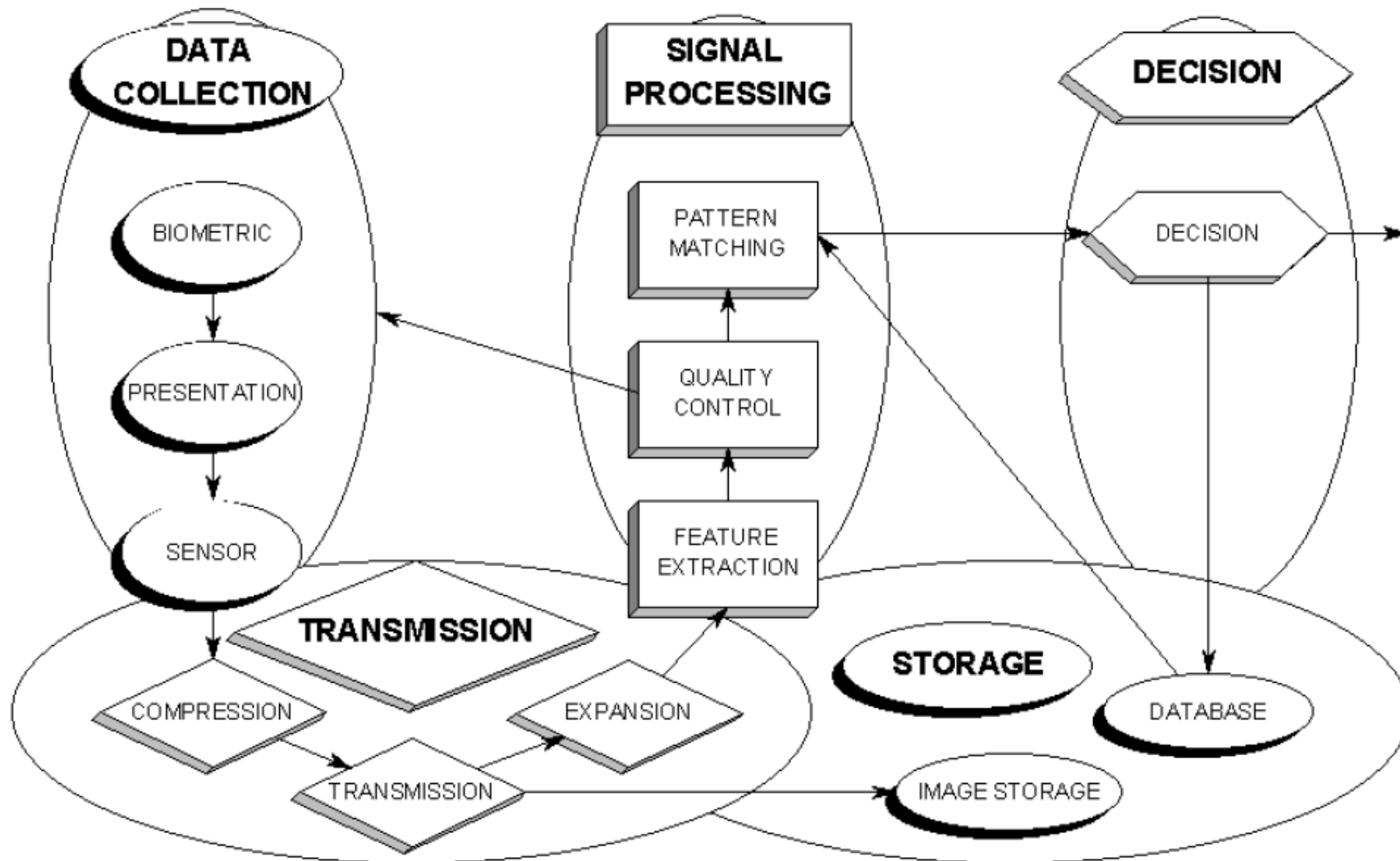
Behaviour vs. Physiology: Aquisition und Liveness

Aquisition: im Fall von active Biometrics gibt es zwei Alternativen der Messung: es wird entweder gesamte Verhaltensprozess aufgezeichnet oder nur das Endergebnis. Im ersten Fall sind A/D Wandlungsprobleme zu berücksichtigen. Das Ergebnis der Messung sind temporal geordnete Messwerte die durch Sampling gewonnen werden. Diese Methoden werden als "On-line" bezeichnet. Im Gegensatz dazu betrachten "off-line" Methoden nur das Endergebnis, nicht den dynamischen Aspekt. **Beispiel**: handschriftliche Signatur kann beides sein. Off-line features können aus on-line features gewonnen werden aber nicht umgekehrt.

Liveness: Da bei physiologischen Merkmalen keinerlei Aktivität notwendig ist es für die Sicherheit eines solchen Systems sehr hilfreich die Lebendigkeit des aufgenommenen Materials zu bestätigen (z.B. Angriffe durch Gummi-Fingerabdrücke, abgeschnittene Finger, Masken, Fotos u.s.w.). Strategien hierbei sind: Randomisierung (mehrere Aquisition Vorgänge), Aufzeichnung früherer Samples (um im Zweifel Entwicklung feststellen zu können), Multibiometrics (Kombination von mehreren Merkmalen oder verschiedenen Sensoren für ein Merkmal), Multi-factor Authentifizierung (Kombination von Biometrics, Besitz und Wissen). In supervised environments kann Liveness detection entfallen.

Auch bei verhaltensbasierten Merkmalen gibt es Täuschungsmöglichkeiten, z.B. durch erzwungenes Verhalten. Dies könnte ev. auch durch weitere Tests untersucht werden.

Biometrische Systeme: Systemmodell



Data Collection

Der Output des Sensors wird beeinflusst durch:

- Das biometrische Merkmal (mit diversen variabilities)
- Die Art in der das Merkmal präsentiert wird und den Eigenschaften der Umgebung
- Der technischen Charakteristik des Sensors (die je nach Umgebungseigenschaften wieder unterschiedlich sein kann)

Die Stability wird durch Änderungen dieser Items negativ beeinflusst. Wird ein offenes System angestrebt, müssen die Sensorcharakteristik und die Art der Präsentation standardisiert werden.

Transmission

Bei manchen biometrischen Systemen geschieht die Sensorik an einem anderen ort als Speicherung und Verarbeitung. In diesem Fall müssen sie aquirierten Daten übertragen werden. Bei hochaufgelösten Bilddaten kann Kompression notwendig sein um Übertragungskapazität zu sparen. Ist dies der Fall müssen die Daten vor der Weiterverarbeitung wieder dekomprimiert (“expansion”) werden. Im allgemeinen werden lossy (i.e. verlustbehaftete) Verfahren verwendet. Ein aktuelles Forschungsgebiet ist die Entwicklung von optimalen Kompressionsverfahren für spezielle biometrische Merkmale, wobei die Optimalität auf möglichst geringe Beeinflussung der Signalverarbeitung ausgerichtet werden muss.

Existierende Standards: Fingerabdrücke (FBI WSQ), Gesichtsbilder (JPEG), Sprachdaten (CELP)

Ein weiterer Punkt im Bereich Transmission ist die Berücksichtigung eventueller Übertragungsfehler.

Signal Processing

Signalverarbeitung bedeutet die Vorbereitung (und folgende Durchführung) der biometrischen Daten auf den Abgleich mit abgelegten Templates in der Datenbank. Folgende Komponenten können identifiziert werden:

- **Feature Extraction:** beinhaltet auch Segmentierung, i.e. die Identifikation des biometrischen Merkmals im übertragenen Signal (z.B. Erkennung der Phasen von Sprachaktivität und Aussonderung von Sprachpausen). Feature Extraction selbst ist dann die Identifizierung von wiederholbaren und unterscheidungsrelevanten Eigenschaften im biometrischen Merkmal. Nicht-wiederholbare Störungen und redundante Datenteile müssen entfernt werden. Hier setzen auch die zentralen Methoden der Bildverarbeitung an, die wir uns im Folgenden genauer ansehen werden. Feature extraction ist eine Form der nicht-reversiblen Kompression, d.h. das originale biometrische Merkmal kann aus den Features allein nicht rekonstruiert werden. In manchen Systemen erfolgt die Übertragung nach der feature extraction um Bandbreite zu sparen.
- **Quality Control:** hier wird nach verschiedenen Kriterien geprüft ob das erhaltene Signal von ausreichender Qualität ist um einen vernünftigen matching Prozess durchführen zu können. Ist dies nicht der Fall, muss der Vorgang der Data Collection wiederholt werden. Das Einfügen dieser Qualitätskontrolle hat in den letzten Jahren biometrische Systeme stark verbessert.

Signal Processing - Pattern Matching

Das feature “sample” ist nun von wesentlich geringerer Grösse als das originale Signal. Es wird nun mit den in der Datenbank gespeicherten “templates” oder “models” verglichen (die vom Enrollment Prozess stammen). Die features im template sind vom gleichen Typ als die im sample (also z.B. in beiden Fällen ein Vektor). Ein “model” hat eine komplexere mathematische Formulierung als ein template und wird z.B. in Sprecher- und Gesichtserkennung verwendet, templates sind typisch für Fingerabdrücke, Iris, und Handgeometriesystemen.

Ziel des matching Prozesses ist ein quantitatives Ergebnis des Vergleichs, das im Anschluss an das Entscheidungssystem geschickt wird. Je nachdem ob es sich um Verifikation oder Identifikation handelt, wird einmaliges oder wiederholtes matching durchgeführt.

Das Ziel ist es für features von einem Individuum kleine Unterschiede zu erhalten und grosse für solche von unterschiedlichen Individuen. Wirkliche 0-Werte sind praktisch nie zu erwarten.

Storage

Zwei Typen von Daten werden potentiell gespeichert:

1. Templates: wenn es sich um ein reines Verifikationssystem handelt, kann auch eine verteilte Datenbankstruktur gewählt werden. In diesem Fall können die templates auf Smartcards, optischen Karten oder Magnetkarten gespeichert werden und es muss keine zentrale Datenbank existieren. Auch in diesem Fall ist eine zumindest zusätzliche zentrale Datenbank von Vorteil, da gefälschte Karten so erkannt und Kartenduplikate ohne wiederholtes enrollment ausgestellt werden können. In Identifikationsanwendungen ist es wichtig dass die Speicherung strukturiert erfolgt (z.B. durch Indizierung und Klassifikation) um exhaustive Search vermeiden zu können (Geschwindigkeit steigt, aber auch die Fehleranfälligkeit).
2. Raw-Daten: da aus den features die Rohdaten nicht mehr rekonstruiert werden können, kann es für manche Systeme von Vorteil sein die Rohdaten ebenfalls zu speichern. Dies muss in besonders geschützter (i.e. verschlüsselter) und komprimierter Form getan werden. In diesem Fall kann auch auf andere features "umgestiegen" werden, wenn z.B. die Lizenzgebühren eines Herstellers zu hoch werden oder ein featuresatz sich als nicht mehr geeignet herausstellt.

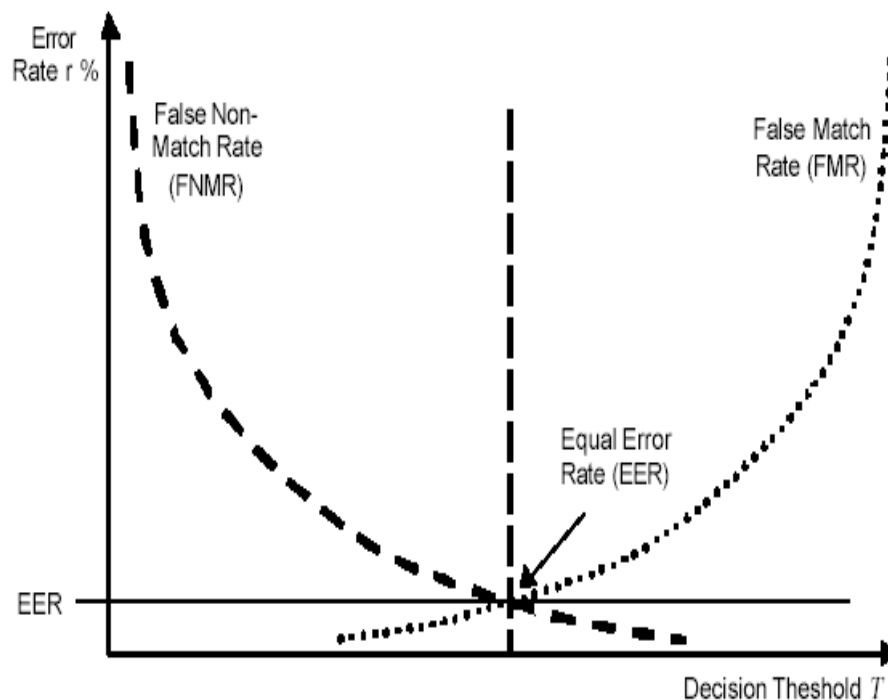
Decision

Das Decision Subsystem setzt die Systempolicy durch Steuerung der Datenbanksuche um, bestimmt Treffer und nicht-Treffer basierend auf dem vom patternMatcher gelieferten Ähnlichkeitswert und gibt die letztgültige Entscheidung aus.

Beispiele für Policies: Ein Benutzer wird nicht authentifiziert, dessen Merkmal nicht aufgenommen werden konnte (klar, aber dies z.B. Anstelle einer wiederholten data collection Prozedur). Es kann fixe Akzeptanz Thresholds gegen oder solche die abhängig sind von diversen Parametern wie z.B. der Person, der Zeit, Umgebungsparametern etc. Weiters können in einem bestimmten Wertebereich mehrere samples verlangt werden um zu einer endgültigen Entscheidung zu gelangen. Ebenfalls können hier Sperrpolicies umgesetzt werden, wie z.B. im Fall von Verifikation sind nur zwei Fehlversuche erlaubt. Auch werden hier Entscheidungen getroffen die Anzahl der zu erwartenden falsch Positiven Authentifizierungen vs. die falsch Negativen betreffend. Dies setzt auch Wissen über die zu erwartende Wahrscheinlichkeit voraus, dass jemand versucht das System zu täuschen.

Bewertung von Biometrischen Systemen I

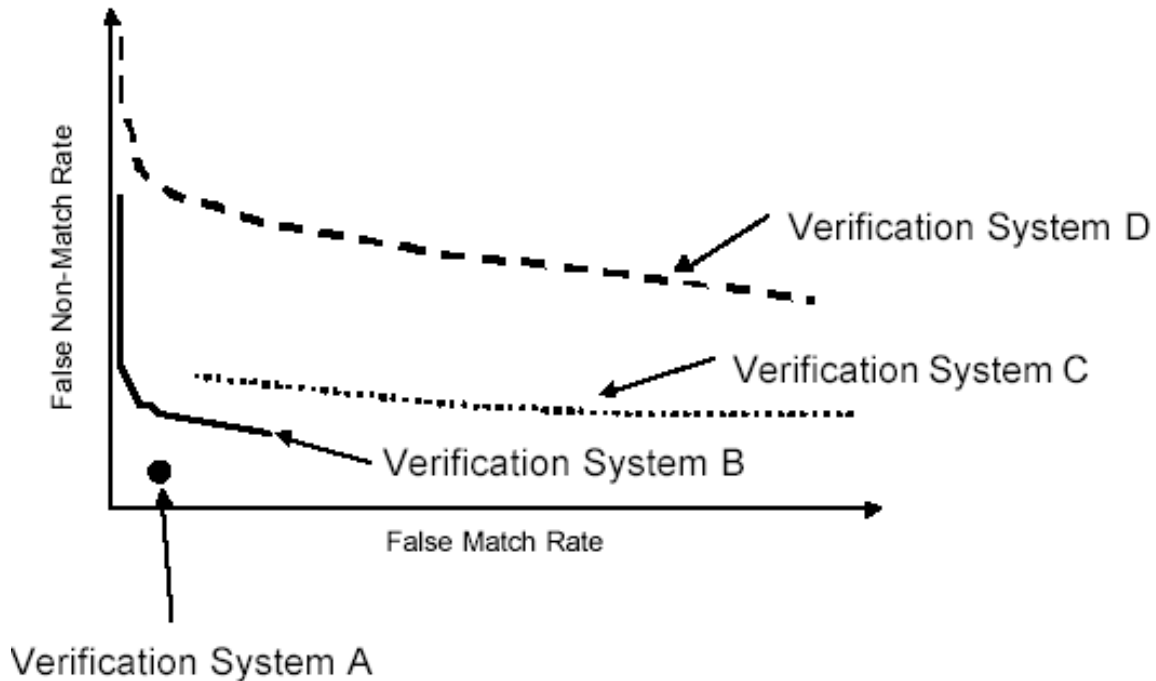
- *False Match Rate (FMR)*: Verhältnis zwischen angezeigten Übereinstimmungen die tatsächlich nicht übereinstimmen und der Gesamtanzahl der durchgeführten Tests (falsch positive Verifikation oder Identifikation, Typ II Fehler)
- *False Non-Match Rate (FNMR)*: Verhältnis zwischen nicht-angezeigten Übereinstimmungen die tatsächlich übereinstimmen und der Gesamtanzahl der durchgeführten Tests (falsch negative Verifikation oder Identifikation, Typ I Fehler)



* FMR und FNMR werden oft als abhängig von einer Entscheidungsschranke T (für Erkennung zugelassener Unterschied zwischen Sample und Template) dargestellt.

* Equal-error-rate ist der Punkt wo FMR und FNMR den gleichen Wert annehmen. Das ist ein oft verwendetes Maß für die Systemgenauigkeit und oft werden (ohne praktische Begründung) biometrische Systeme mit diesen Parametern betrieben.

Bewertung von Biometrischen Systemen II



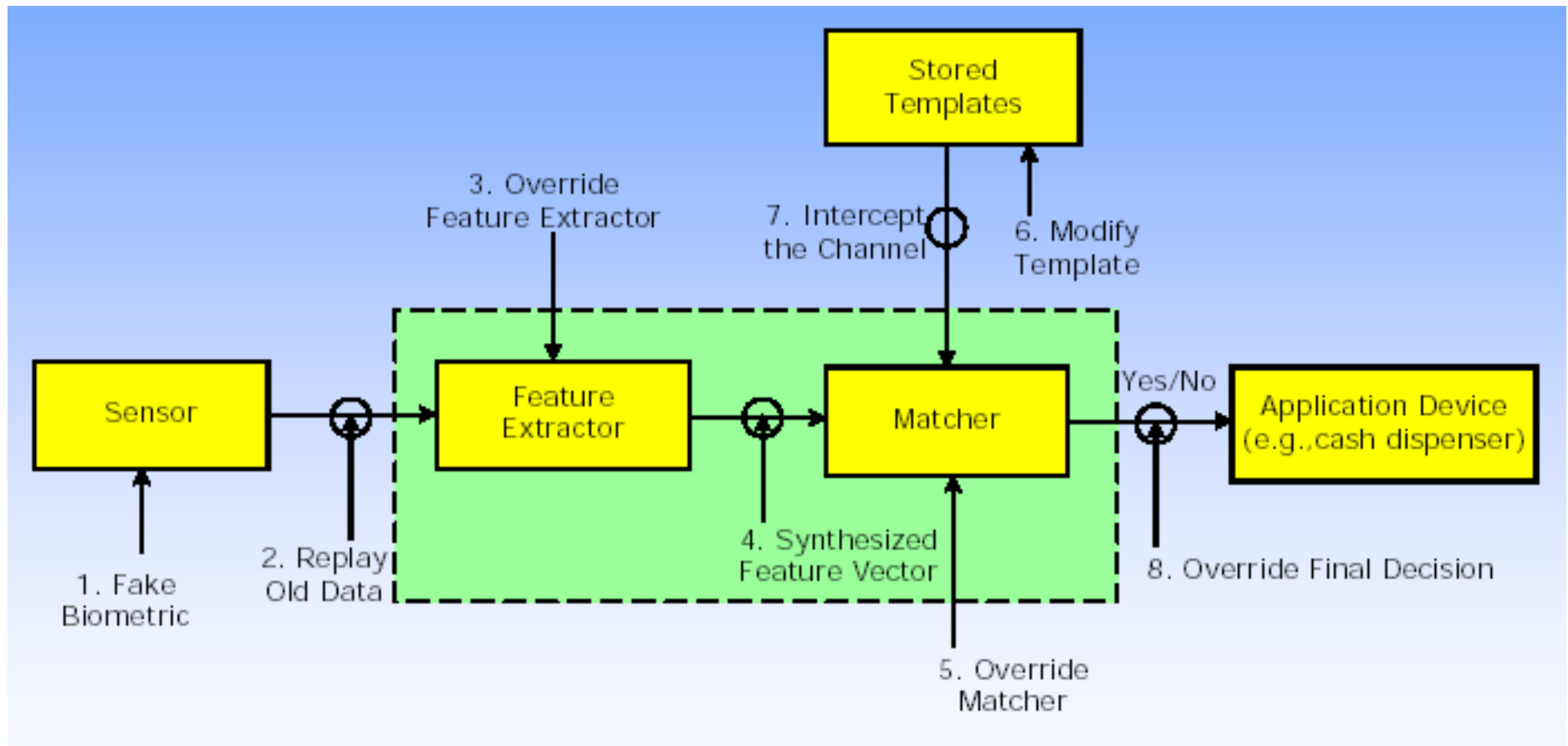
- * *Receiver Operating Characteristic (ROC)*: zeigt FNMR als eine Funktion von FMR.
- * Optimaler Wert natürlich bei (0,0); mittels ROC können gut verschiedene Verfahren verglichen werden bzw. Verfahren mit best. Eigenschaften ausgewählt werden (z.B. Systeme mit hoher FNMR bei niedriger FMR).

- *Binning Error Rate (BER)*: Prozentsatz der Samples die nicht in eine Partition der Daten eingeteilt werden, die den original Templates entsprechen würden (hier geht es offenbar um grossen Datenmengen die nicht einfach sequentiell durchsucht werden, sondern z.B. Klassifikation benutzen; v.a. Identifikationsprobleme).

Bewertung von Biometrischen Systemen III

- *Penetration Coefficient (PC)*: Durchschnittliche Anzahl von Vergleichen für jedes Sample im Verhältnis zur Datenbank Größe (Such Komplexität).
- *Transaction Time (TT)*: Zeitbedarf für eine Authentifikations Transaktion; $T_{collect} + T_{compute}$
- *FAR und FRR*: Akzeptanz kann auf mehreren Matches bzw. non-Matches beruhen, ist daher wesentlich unklarer. Für einfache System gilt ohnehin $FAR = FMR$ und $FRR = FNMR$.
- *FIR und CIR*: im Fall von Identifikation, falsch positive vs. korrekt positive durch Gesamtanzahl.
- *Threshold trade-off*: Ein geringerer Wert der Entscheidungsschranke führt zu geringerer Erkennungswahrscheinlichkeit von nicht-authentischen Personen (FMR sinkt), aber natürlich auch zu höherer Wahrscheinlichkeit einer Ablehnung für authentische Personen (FNMR steigt). Wahl der Schranke ist Kompromiss zwischen Sicherheit und Benutzerfreundlichkeit.

Sicherheit von Biometrischen Systemen



Privacy von Biometrischen Systemen I

Wird ein biometrisches System benutzt, ist es möglich persönliche Informationen über die Benutzer zu erhalten ?

- Sind Eigenschaften des Merkmalsträgers eruierbar ? Retinascan könnte im Zusammenhang mit Retinadiagnose problematisch sein, Stimmerkennung gibt Aufschluss über Geschlecht, Unterschrift über Charaktereigenschaften, Gang über eventuelle Verletzungen des Bewegungsapparats, Herzrhythmus über ev. Herzerkrankungen – diese Dinge scheinen unproblematisch zu sein, mit wenigen Ausnahmen.
- Befürchtungen dass Biometrics benutzt werden können um Personen an personenbezogene Daten zu binden oder Reisetätigkeit nachzuvollziehen (wie es mit Kreditkartenbelegen oder Telefongesprächen getan wird). Im Falle von Telefonbüchern oder Kontodaten gibt es auch “umgekehrte” Daten die die Verbindung von einer Nummer zum Namen herstellen. Im Fall von Biometrischen Daten gibt es praktisch keine öffentlichen Datenbanken und auch nur wenige umfassende im staatlichen Bereich.

Privacy von Biometrischen Systemen II

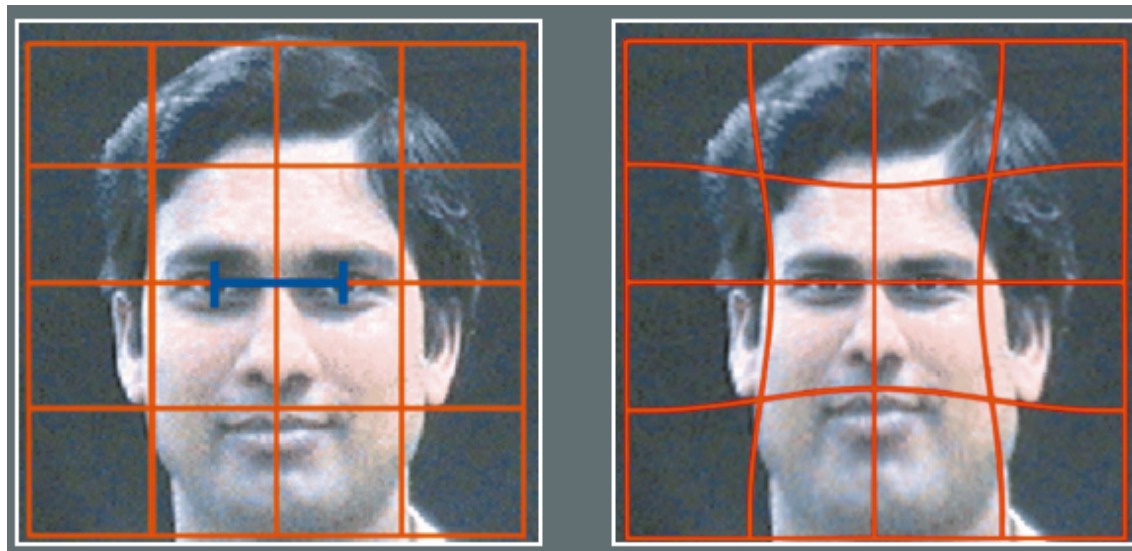
- Die Umkehrbarkeit ist nicht in der gleichen Weise wie z.B. bei Telefonbuchdaten gegeben, da die Eindeutigkeit auf die konkrete biometrische Datenbank beschränkt ist, wenn überhaupt (durch die verschiedenen Typen von benutzten templates). Die Raw Daten sind hier kritischer, was deren (verschlüsselte) Aufbewahrung eher zum Problem macht.
- Biometrische Merkmale sind nicht geheim, manche können direkt aus öffentlichen Daten generiert werden (z.B. Foto auf der Webpage für Gesichtserkennung). Sie ähneln oft vom Konzept her public-keys, die jedoch, wenn komprimiert, nicht ausgetauscht werden können.

Insgesamt scheint jedoch (bis auf den letzten Punkt) das Privacy Problem kleiner zu sein als bei klassischen Authentifizierungsmethoden.

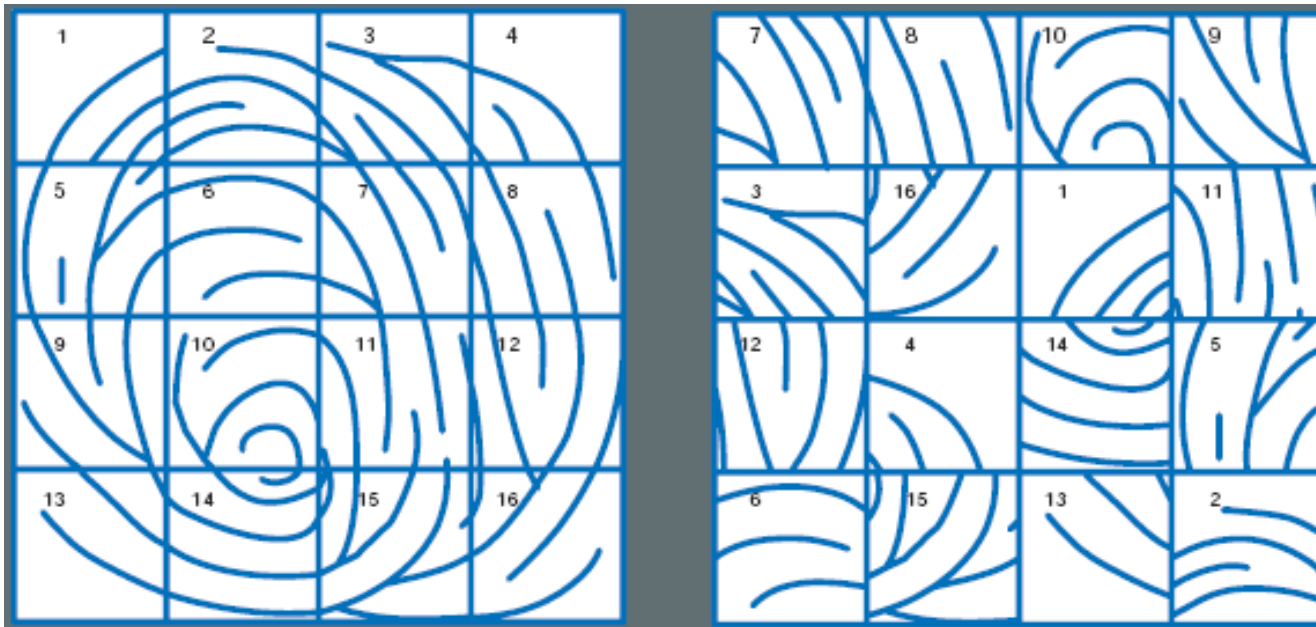
“Cancelable” Biometrics I

Es geht hier um die eben besprochene Problematik dass Biometrische Merkmale wie bei tokenbasierten Systemen nicht widerrufen und ausgetauscht werden können, was insbesondere im Fall der Kompromittierung schlecht ist.

Die Idee um Widerrufbarkeit zu erreichen ist eine wiederholbare Störung des Biometrischen Signals einzuführen [?]. Beim Enrollment und jeder Authentifizierung wird die gleiche Störung angewendet, bei unterschiedlichen biometrischen Installationen aber unterschiedliche (damit ist auch das Verknüpfen von biometrischen Daten erschwert). Im Falle der Kompromittierung wird dann nur die Störung ausgetauscht. Achtung: die Störung muss so beschaffen sein, dass sich die biometrischen Merkmale trotz vorhandener Störung extrahieren lassen.



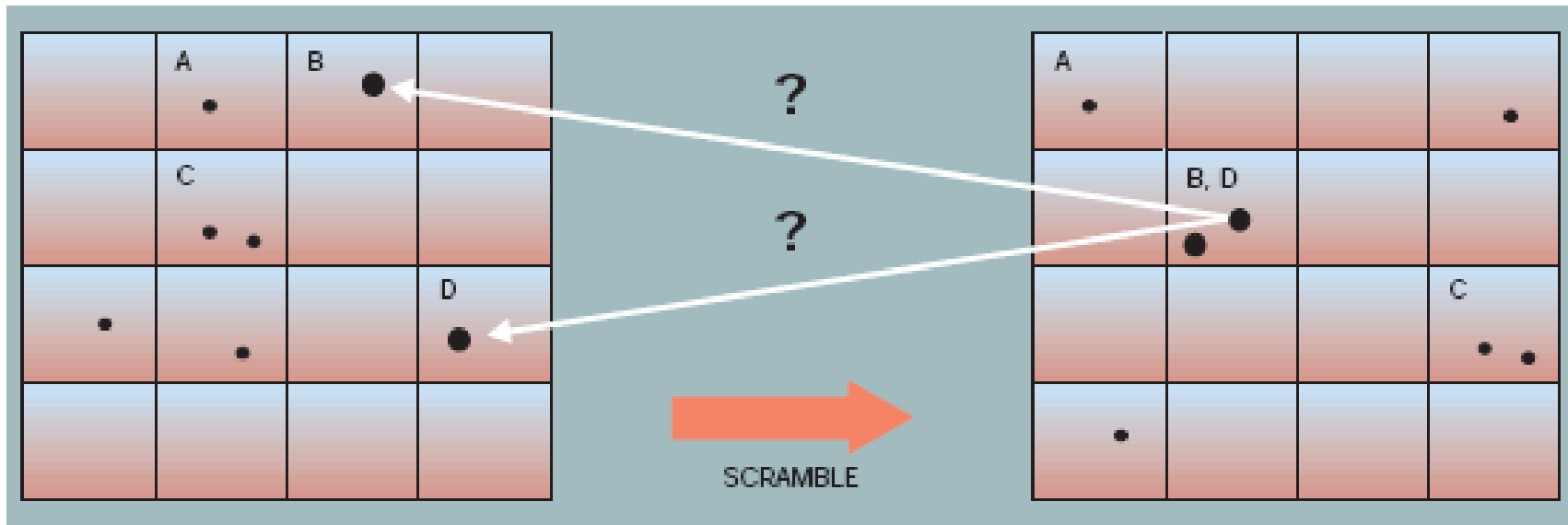
“Cancelable” Biometrics II



Die Störungen sollten durch nicht-invertierbare Transformationen generiert werden, dass das ungestörte Signal auch bei Kenntnis der Transformation und der gestörten Daten nicht rekonstruiert werden kann. Die Transformation kann entweder direkt auf das gewonnene Signal oder auf die features angewendet werden.

Beispiele für Anwendung auf das Signal: Image morphing (für Face recognition) und Blockpermutation (für Minutien-basierte Fingerprint recognition). Nicht-Invertierbarkeit ist nur bei Geheimhaltung der Transformationsparameter gegeben, im Fall der Blockpermutation ist selbst dies mehr als fragwürdig (Angriff durch testen der Grenzähnlichkeit).

“Cancelable” Biometrics III



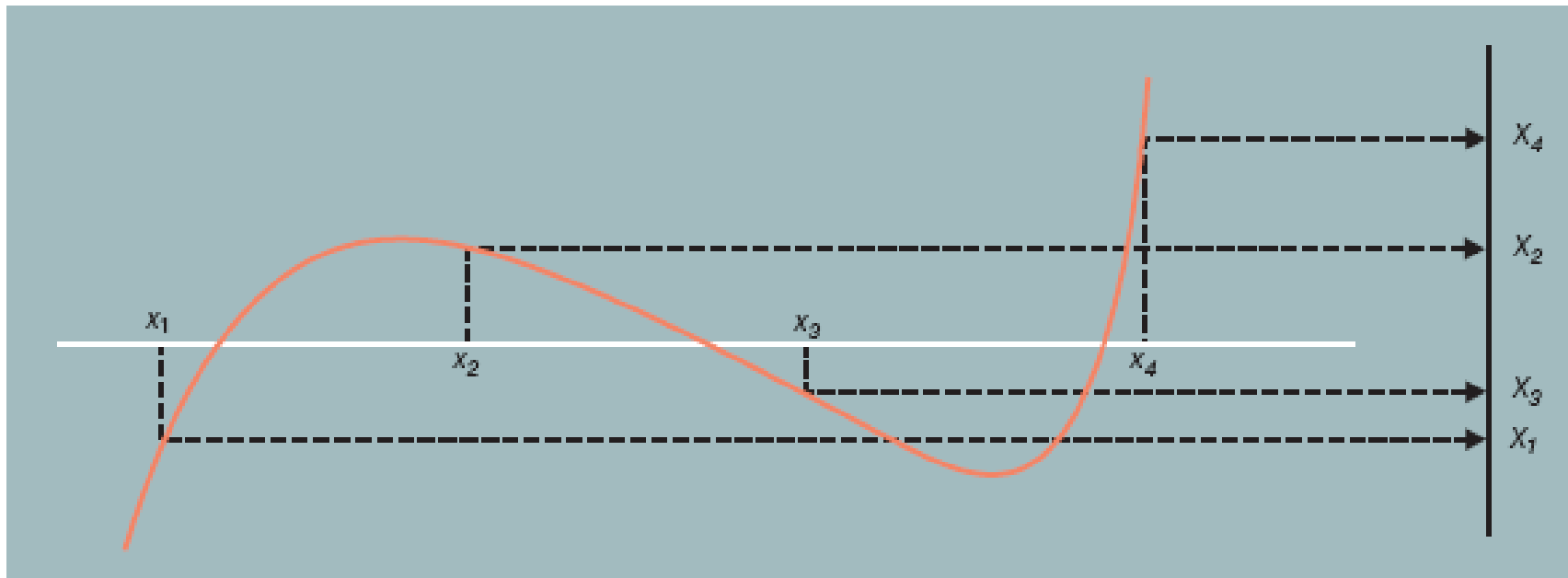
Die Graphik zeigt eine Transformation in der Feature Domain für Fingerprints, wo zusätzlich zur Permutation mehrere Blöcke auf einen abgebildet werden. Wenn es zu keinen Überdeckungen kommt, bleiben die Features erhalten und die Transformation ist nicht invertierbar.

Um eine wiederholbare Transformation zu erhalten, muss das biometrische Signal und die resultierenden Features ausgerichtet (image registration) werden – dies kann z.B. durch “cores” und “deltas” bei Fingerprints geleistet werden (siehe dort!).

“Cancelable” Biometrics IV

Im Folgenden ein Beispiel für eine nicht-invertierbare Transformation in der Feature Domain, anwendbar für Punktemuster (wie bei Fingerprints). Ein Minutiensatz S besteht beispielsweise aus $S = \{(x_i, y_i, \phi_i), i = 1, \dots, M\}$. Eine nicht-invertierbare Funktion der x -Koordinate ist z.B. ein Polynom höherer Ordnung:

$$F(x_i) = \sum_{n=0}^N \alpha_n x_i^n = \prod_{n=0}^N (x_i - \beta_n)$$



Die anderen Koordinaten können mit analogen Transformationen bearbeitet werden.

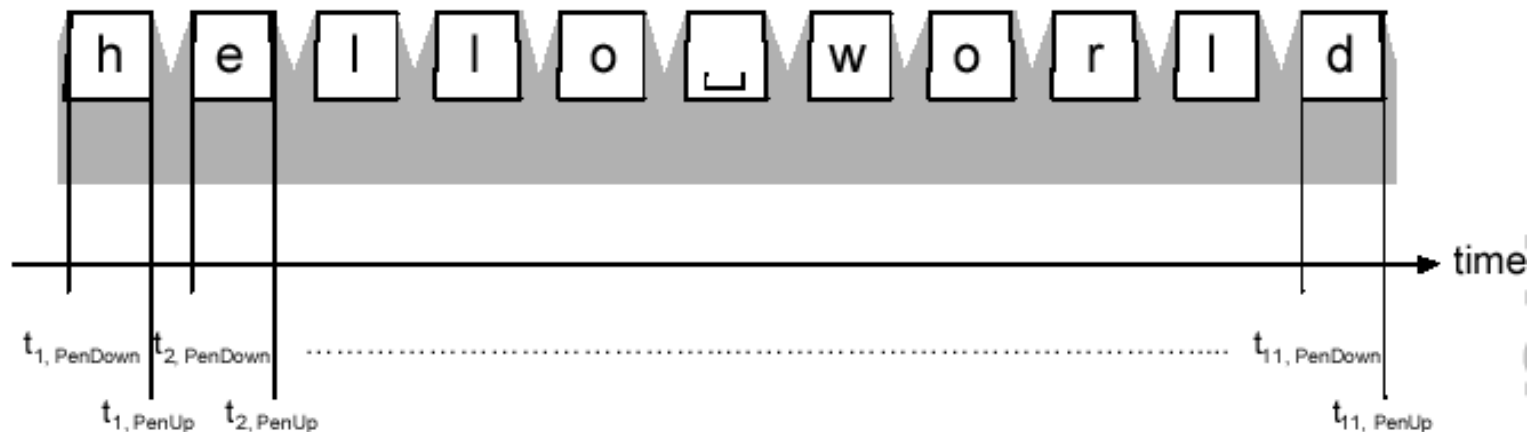
Table of Contents: Nicht-visuelle Biometrie

- Keystroke Dynamics
- On-line Signature
- Biorhythmen
- Sprechererkennung
- Ungewöhnliche Merkmale

Keystroke Dynamics: Grundlagen

Vorteile: Es wird keine besondere Sensorik benötigt, da Keyboards fast in jedem Haushalt, jeder Arbeitsstelle, u.s.w. verfügbar sind (verteilte Infrastruktur, hohe collectability und geringer Preis - ausser man will Pressstärke messen). Weiters wird die Akzeptanz sehr hoch sein, da sehr viele Menschen Umgang mit Tastaturen gewöhnt sind (habituated environment). Die gewonnenen Daten (welche Taste gedrückt, Presszeit, Loslasszeit) können mit wenig Aufwand übertragen, gespeichert und verarbeitet werden.

Nachteile: Intra-personal variability kann sehr gross sein (Person ist müde, verletzt - Sehnenscheidenentzündung, Person steht; insbesondere bei wenig geübten Tippern ist variability extrem hoch !), durch die Verteiltheit ist auch das environment natürlich unattended und kann sehr unterschiedlich sein (z.B. unterschiedliche Keyboard Layouts führen zu verschiedenem Tippverhalten).



Keystroke Dynamics: Login vs Monitoring

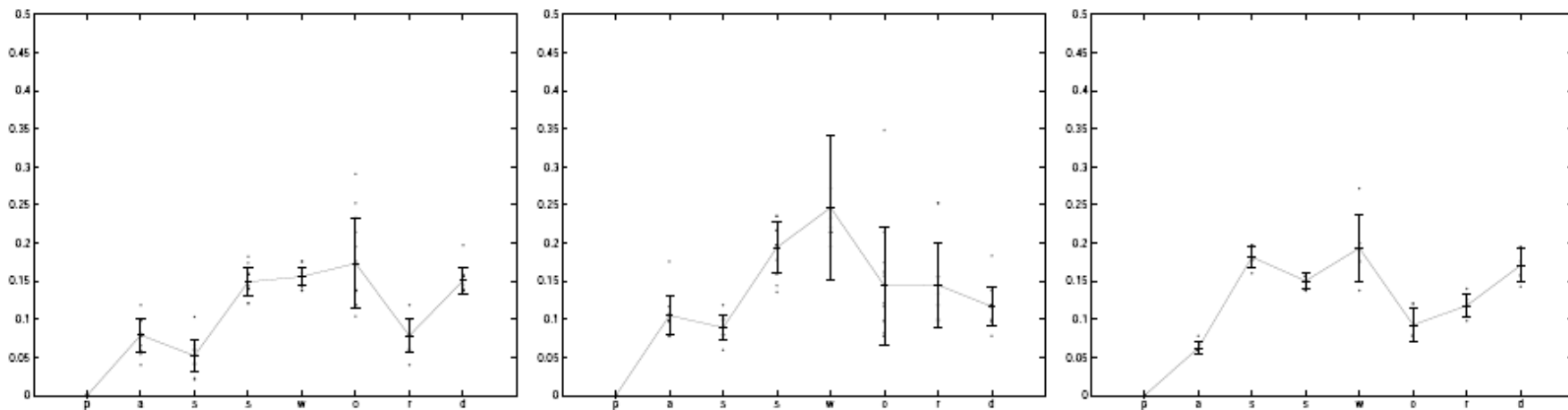
Login: ist das klassische Szenario von “key-stroke enhanced” Login. Ein Benutzer gibt Login und Passwort (oder nur Login) ein, zusätzlich wird das Tippverhalten aufgezeichnet. Login wird nur bei ausreichend ähnlichem Tippverhalten gewährt (Beispiel für Authentifizierung durch Wissen **und** Verhalten). Beim Enrollment müssen Login und Passwort öfter als in klassischen Systemen eingegeben werden. Hier stehen relativ wenige Daten zur Verfügung um das Tippverhalten zu charakterisieren, dafür ist es ein fixer Text.

Monitoring: wird konstant während einer Session eingesetzt. Während die Benutzer am Terminal arbeiten, sammelt ein Hintergrundprozess Daten und bewertet ihr Tippverhalten. Bestehen Zweifel an der vorgegebenen Identität kann automatisches Ausloggen oder eine Warnung an den Administrator die Folge sein. Grössere Datenmenge, freie Texte.

In der Literatur wird das als Gegensatz Verifikation und Identifikation bezeichnet, ist aber eigentlich in beiden Fällen Verifikation !

Keystroke Dynamics: Latency

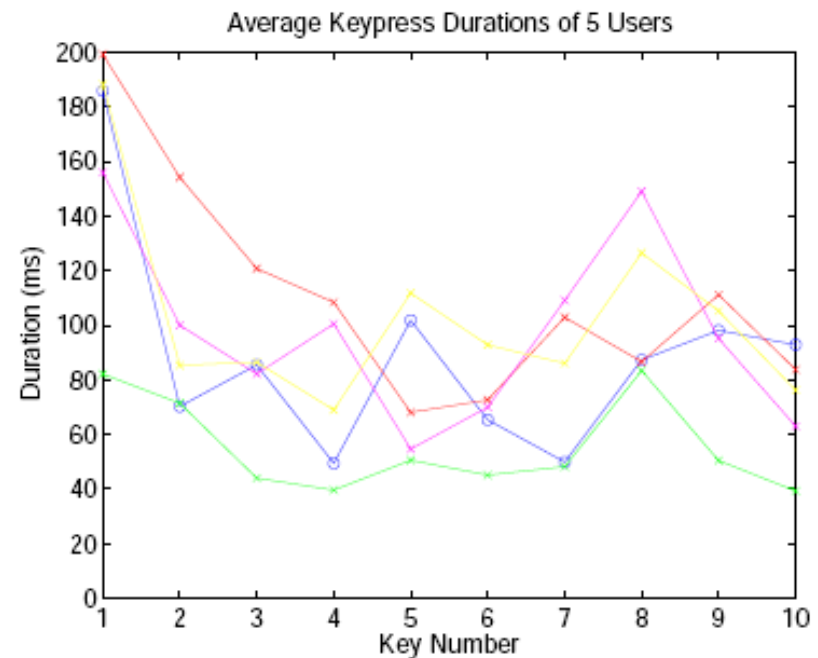
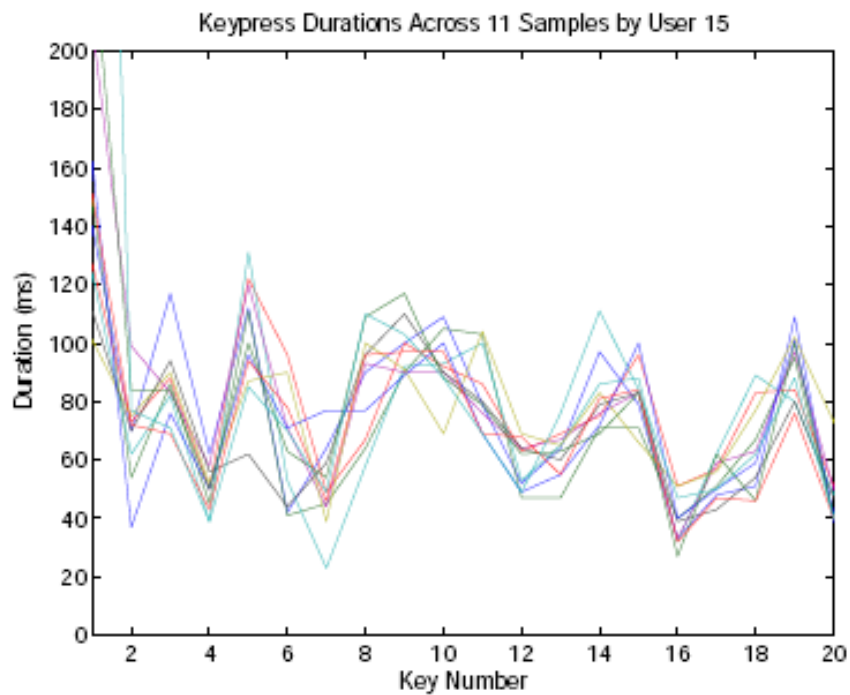
Ist **das** klassische Keystroke Feature. Beim Loginszenario wird als Element i des Featurevektors $t_{i+1, PenDown} - t_{i, PenDown}$ verwendet. Auf diese Vektoren können dann klassische Metriken angewendet werden. Es ist zu beachten dass die $t_{i, PenUp}$ nicht der logischen Anordnung der Buchstaben entsprechen müssen (da man eine Tasten zu verschiedensten Zeitpunkten loslassen kann), auch $t_{i, PenUp} < t_{i+1, PenDown}$ ist nicht gewährleistet.



Beim Monitoring Szenario werden Latenzzeiten von einer Menge von Buchstabenpaaren gesammelt und die entsprechenden Werte zwischen Benutzern verglichen.

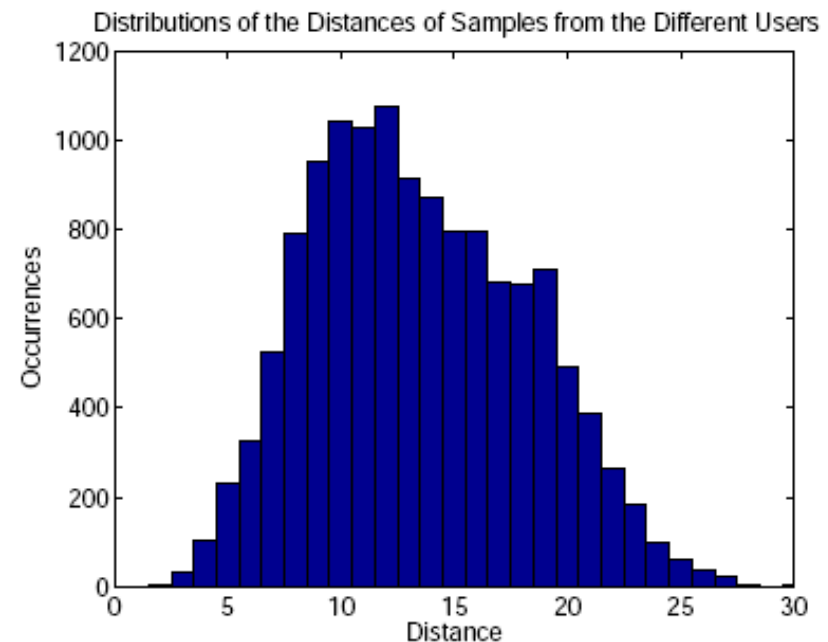
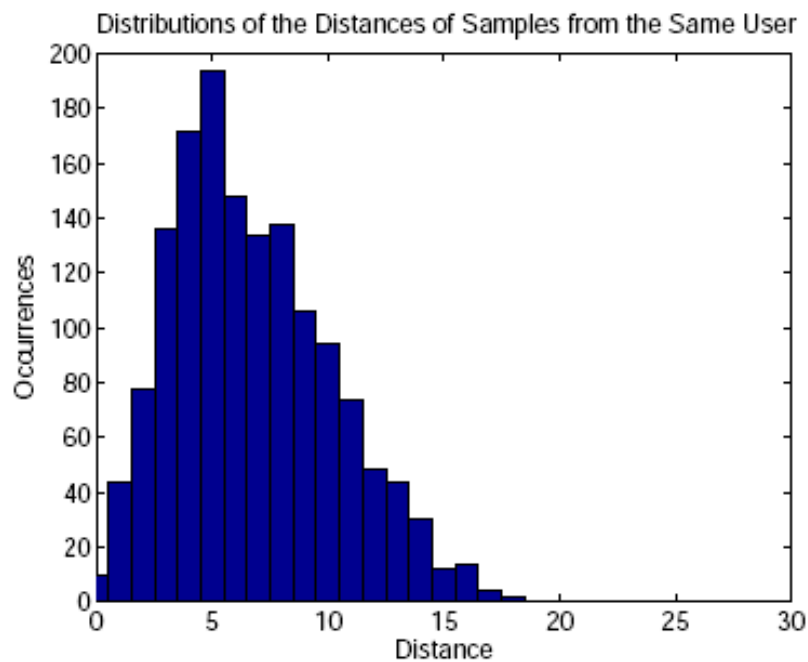
Keystroke Dynamics: Druckdauer

Ein Element i des Featurevektors beim Loginszenario ist $t_{i, PenUp} - t_{i, PenDown}$. Dieses Feature wird auch gern in Kombination mit der Latenz eingesetzt. Analoge Überlegungen bzgl. Login und Monitoring treffen zu.



Keystroke Dynamics: Anordnung der Tastenevents

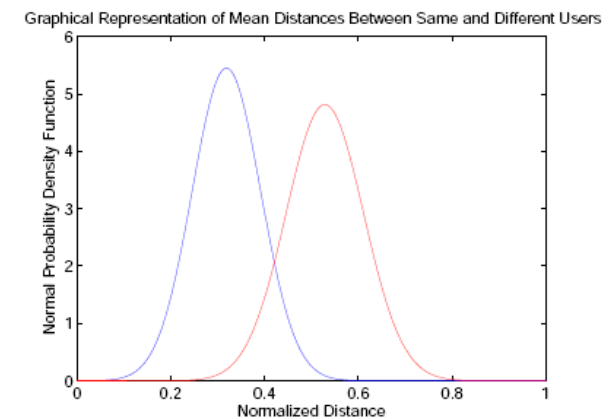
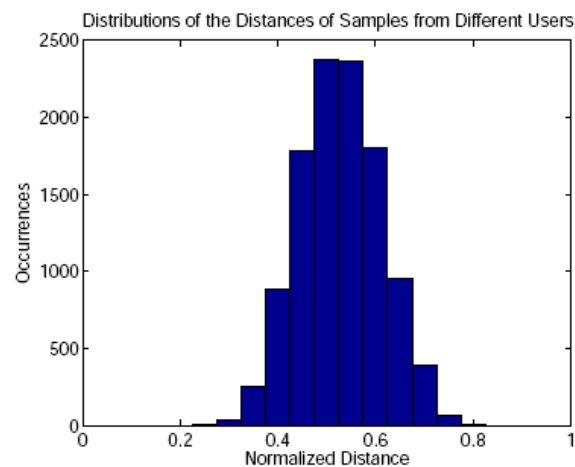
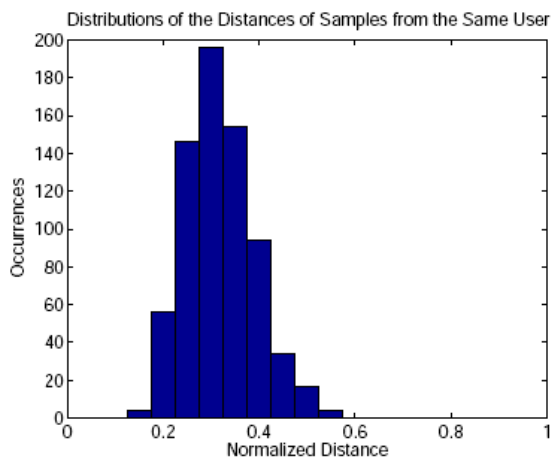
Diese Technik wird nur beim Login Szenario angewendet. Erstellt wird hier die zeitliche Anordnung der $t_{i, PenUp}$ und $t_{i+1, PenDown}$ für alle Tastenevents (die in Wirklichkeit nicht so aussieht wie im einführenden Bild). Werden nun zwei Login Vorgänge verglichen, so werden im wesentlichen die Anzahl der unterschiedlichen Positionen gezählt (geteilt durch zwei, weil die Anzahl der Tasteneventvertauschungen als Mass genommen wird).



Keystroke Dynamics: Relative Geschwindigkeit

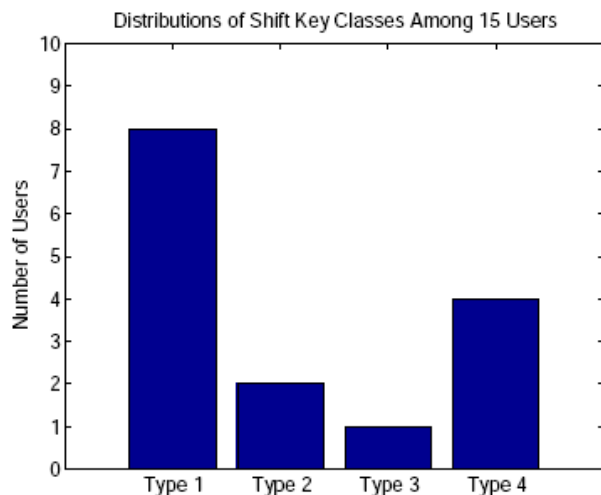
Ist bei der absoluten Tippgeschwindigkeit die Intra-personal variability auch recht hoch, ist die relative Geschwindigkeit zwischen verschiedenen Tastenevents vielleicht eher konstant. Um dies als Mass zu verwenden, werden die Latenz oder Presszeiten sortiert, als Vektor repräsentiert und Tastenpaare die nicht überall vorkommen entfernt (z.B. Taste und Backspace bei Fehlern). Sei $S[i]$ die Position des Messwerts im sortierten Sample, kann der Abstand zwischen S und S' wie folgt bestimmt werden:

$$Abstand = \sum_i |S[i] - S'[i]|$$



Keystroke Dynamics: Shift Benutzung und andere Ideen

Für jede Taste kann entweder die linke oder die rechte Shifttaste verwendet werden. Für habituated user können ihre muster in der Shift-Tastenbenutzung zur Unterscheidung benutzt werden. 4 Klassen von Shift Benutzern hypothetisiert: reine links und rechts Shift Nutzer, Gegenseite Shift Nutzer (benutzen die Taste auf der gegenüberliegenden Seite der Tastatur um die Schreibeffizienz zu erhöhen) und chaotische Shift Nutzer.

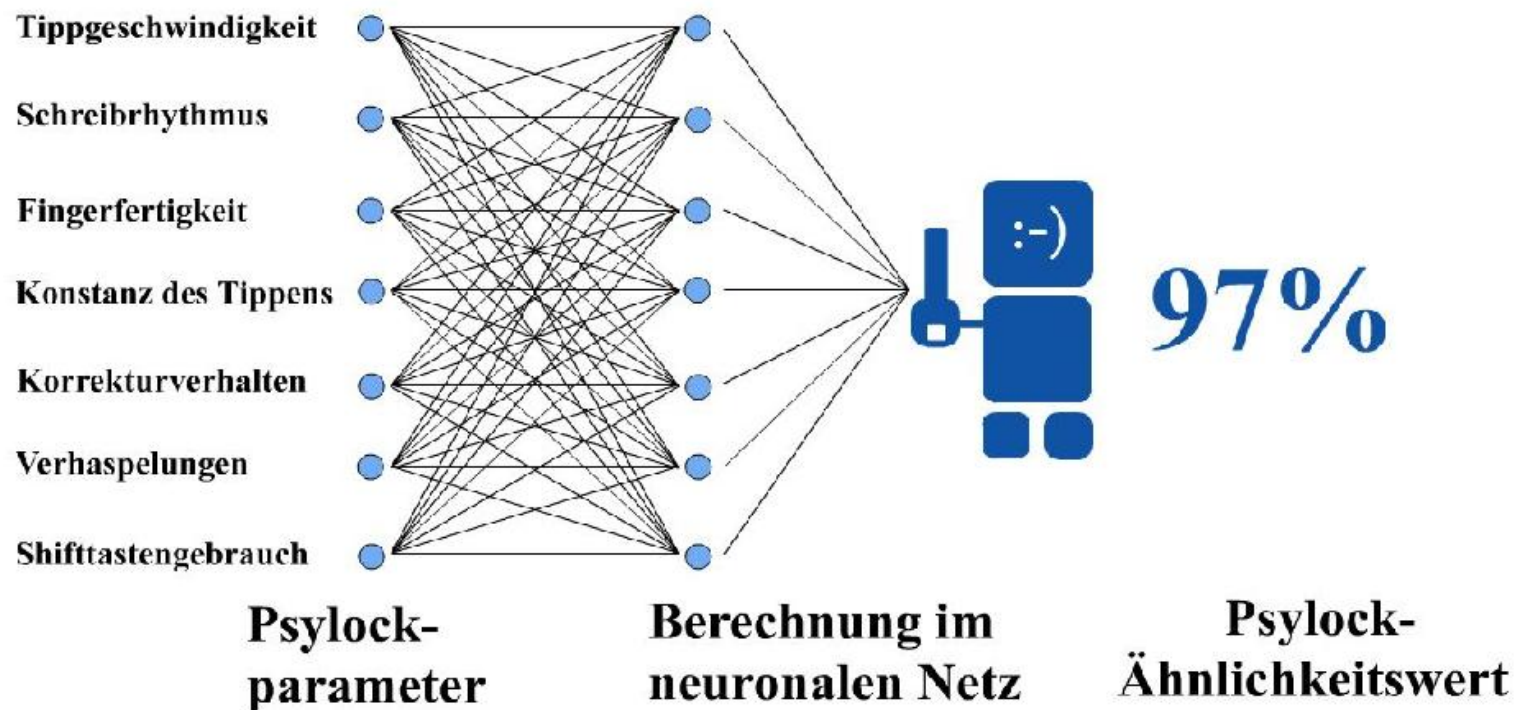


- * Wirklichkeit: reine links und rechts Shift Nutzer existieren.
- * Typ 3: beidseitig und fixiert bei allen Tasten, nicht aber gegenseitig; Typ 4: die letzte Klasse nur bei Mehrheit der Tasten festgelegt.

FAZIT: Shift Benutzung kann nur als zusätzliches Feature verwendet werden um eine Klassifizierung anzuwenden, ähnlich wie Rechts- und Linkshänder (beim Rechtshänder sind die rein rechten Tastenpaare “schneller”).

Keystroke Dynamics: Kommerzielle Systeme

- www.biopassword.com: Straightforward Implementierung von Latency und Pressdauer, geht auch in Windows Netzwerken. Beim Enrollment werden ID und PWD 15x eingetippt. Wurde von Testern gut bewertet.
- www.psylock.de: Aufwändigeres System mit zusätzlichen Features; wird von Byometrics mit Fingerprint und Iris kombiniert. Nette Demo online !



Keystroke Dynamics: Privacy

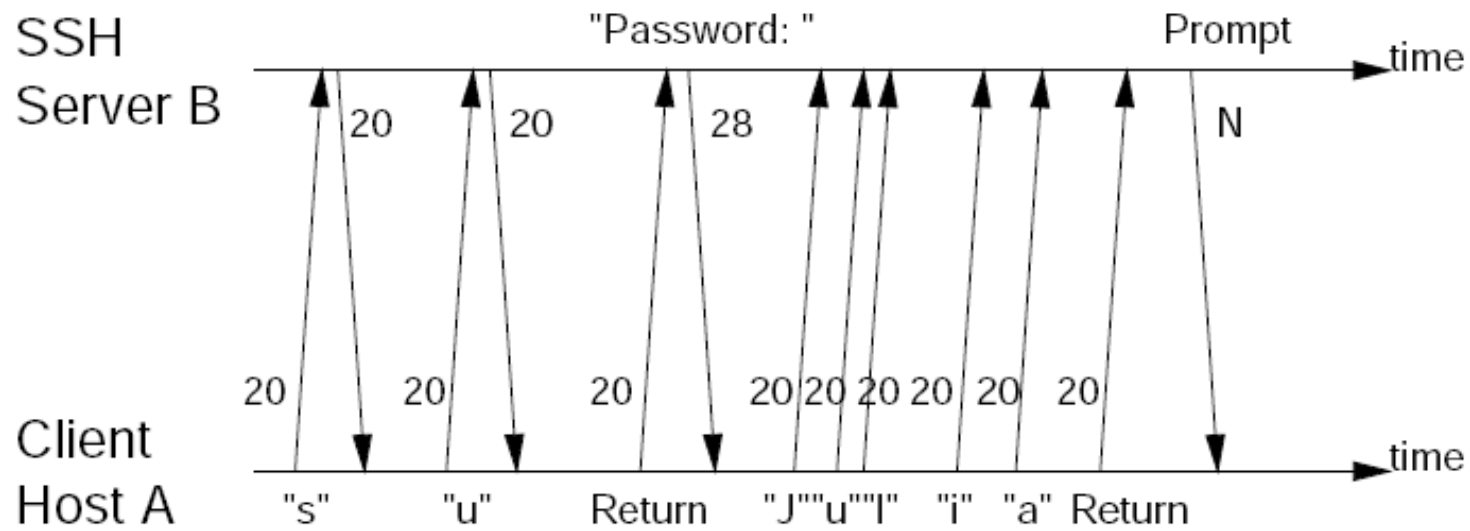
Grundlegendes Problem: ist ein Profil des Tippverhaltens bekannt, kann der Inhalt von getipptem Inhalt rekonstruiert werden basierend auf den Daten des Tippverhaltens. Das kann für Angriffe ebnutzt werden, entweder durch Profil eines bestimmten Benutzers oder durch erstellen von typischen durchschnittlichen Profilen um die Anzahl der noch möglichen Buchstabenkombinationen für eine brute-force Attacke zu verringern (hier natürlich geringere Genauigkeit).

Beispiel SSH: im interaktiven Modus werden die einzelnen Tastenevents in separaten IP-Paketen sofort nach dem Drücken versendet. Das gilt allerdings nicht für den initialen Login (hier ist das gesamte PWD in einem Paket enthalten), sondern beim Benutzen der etablierten SSH Verbindung. Problem ist hier z.B. der Wechsel zum super-user account mit nachfolgender su-Passworteingabe.

Die Frage ist nun noch wie man automatisiert einen solchen Vorgang erkennen kann.

Keystroke Dynamics: Angriff geg. SSH I

Alle normale Tastendrucke die zum SSH-server gehen generieren ein return-Paket (durch die darstellung des Characters am Bildschirm), dies ist für Passwordeingabe nicht der Fall. So kann dieses Muster leicht aufgespürt werden.

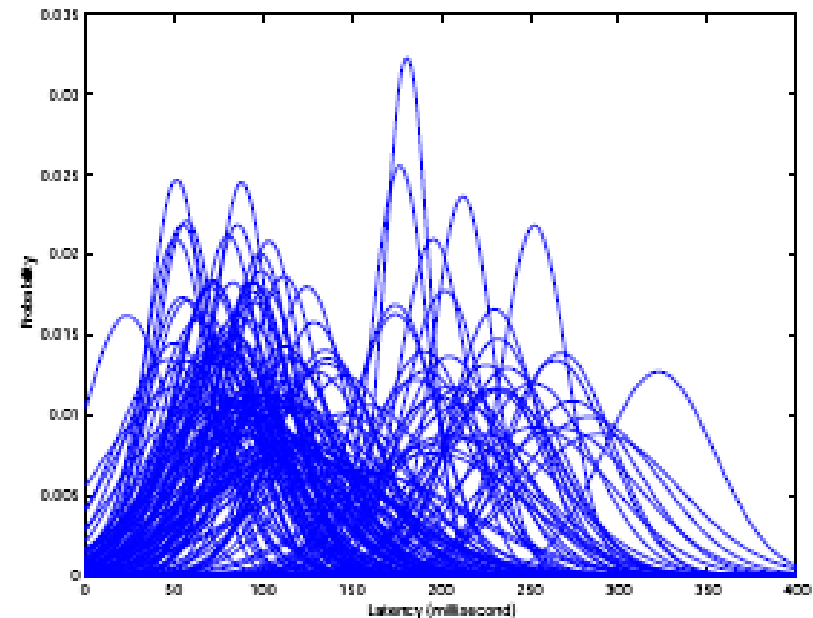
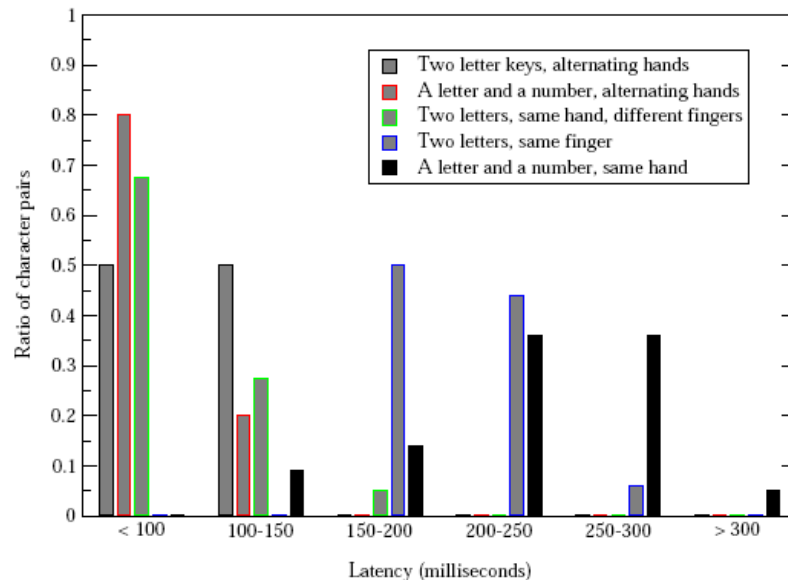


Zur Durchführung eines konkreten Angriffs müssen nun noch Daten über das Tippverhalten vorliegen, wie sie z.B. durch Keystroke Biometrics vorliegen.

Abhilfe wäre z.B. das Senden von Dummy Daten bei der PWD Eingabe (gegen das Aufspüren der Login-Sequenz) oder das Benutzen von Zufälligen Delays beim Senden der Pakete (die aber so lange sein müssten dass es den Nutzer stören k önnte).

Keystroke Dynamics: Angriff geg. SSH II

Jedenfalls zeigt sich am Beispiel Latenz dass Daten vorliegen die entsprechende Angriffe möglich machen:

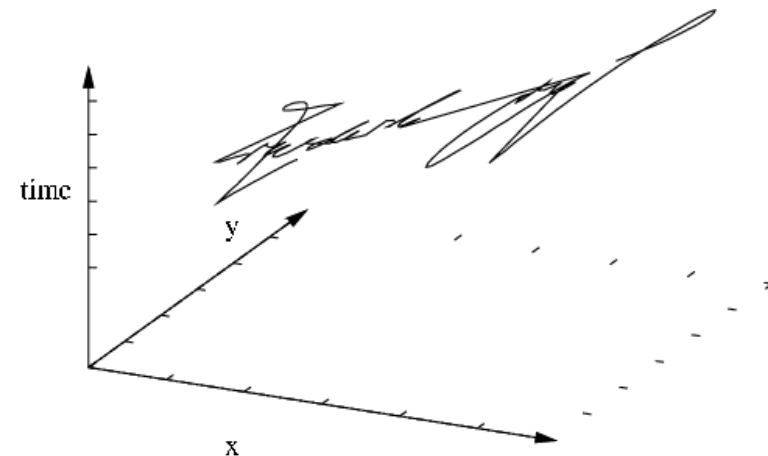
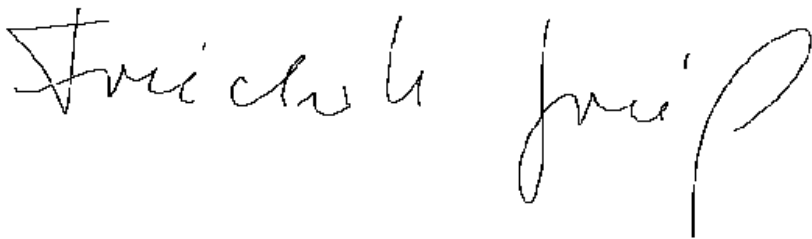


Eine ähnliche Attacke kann gegen PGP durchgeführt werden, da hier Tippverhalten als Zufallsquelle zur Erzeugung von Seed-werten benutzt wird.

On-Line Signaturen

Off-Line Signaturen: Das Resultat des Unterschriftsvorganges ist das biometrische Merkmal. Dies ist im Allgemeinen ein digitales Bild. Hier können Methoden der Bildverarbeitung angewendet werden um Übereinstimmungen festzustellen (siehe im entsprechenden Teil der Unterlagen).

On-Line Signaturen: Das biometrische Merkmal ist die zeitabhängige Dynamik des Unterschriftsvorgangs. Das ist im Allgemeinen eine zeitabhängige Funktion. Hier werden Methoden der Zeitreihenanalyse angewendet um Übereinstimmungen festzustellen.



Pros & Cons Signaturen in der Biometrie

Pros: Unterschrift ist nach wie vor ein sehr weit verbreitetes und akzeptiertes Mittel der Authentifizierung im analogen Bereich (habituated execution). Durch die vermehrte Verfügbarkeit von TabletPCs, PocketPCs, PDAs, Palms etc. ist auch die direkte Verfügbarkeit der notwendigen Sensorik im Bereich On-Line Signaturen gegeben (Scanner für Off-Line Signaturen sind ohnehin verfügbar) – collectability ist gegeben.

Cons: Unterschriften zeichnen sich durch hohe intra-personal Variability aus (im Alterungsprozess häufig vorkommende Krankheiten - Parkinson - sind ein grosses Problem). Dies macht es insbesondere schwierig professionelle Fälschungen zu erkennen, die i.A. leichter anzufertigen sind als bei anderen biometrischen Verfahren (das ist u.a. ein Argument für on-line Signaturen). Auch gilt das Argument der habituated execution nur für analoge Medien (und damit nur für Off-line Signature Verification).

Verwandte Verfahren und Abgrenzungen

- Text abhängig vs. Text unabhängig: es wären auch Text unabhängige biometrische Verfahren vorstellbar, die dann auf einer buchstabenorientierten Erkennung beruhen würden (handwriting recognition). Hier sind die Erfolgsraten viel schlechter, weil direkter Vergleich von Zeitreihen oder visuellen Daten nicht möglich ist. Andererseits ist aber auch das Erstellen von Fälschungen schwieriger (z.B. bei der Authentifizierung muss ein zufällig ausgewählter Text mit schwierigen Buchstaben geschrieben werden).
- Character, Text & Handwriting Recognition: ist einerseits schwieriger, da textunabhängig und weil man wissen will was es heisst, andererseits aber einfacher weil keine Fälschungen zu berücksichtigen sind. Zeitliche Dynamik ist hier unerwünscht und wird durch entsprechendes Resampling korrigiert.
- Signature Recognition ist implizit immer eine Verifikation, da der behauptete Name geschrieben wird. Dies kann aber nur bei vorgeschalteter Texterkennung ausgenutzt werden.
- Varianten sind das Zeichnen von Symbolen, Schreiben von Passwörtern (hier entfällt der Automatismus der Unterschriftsleistung und es ist eine Kombination mit wissensbasierter Authentifizierung)

On-Line vs. Off-Line Signaturen I

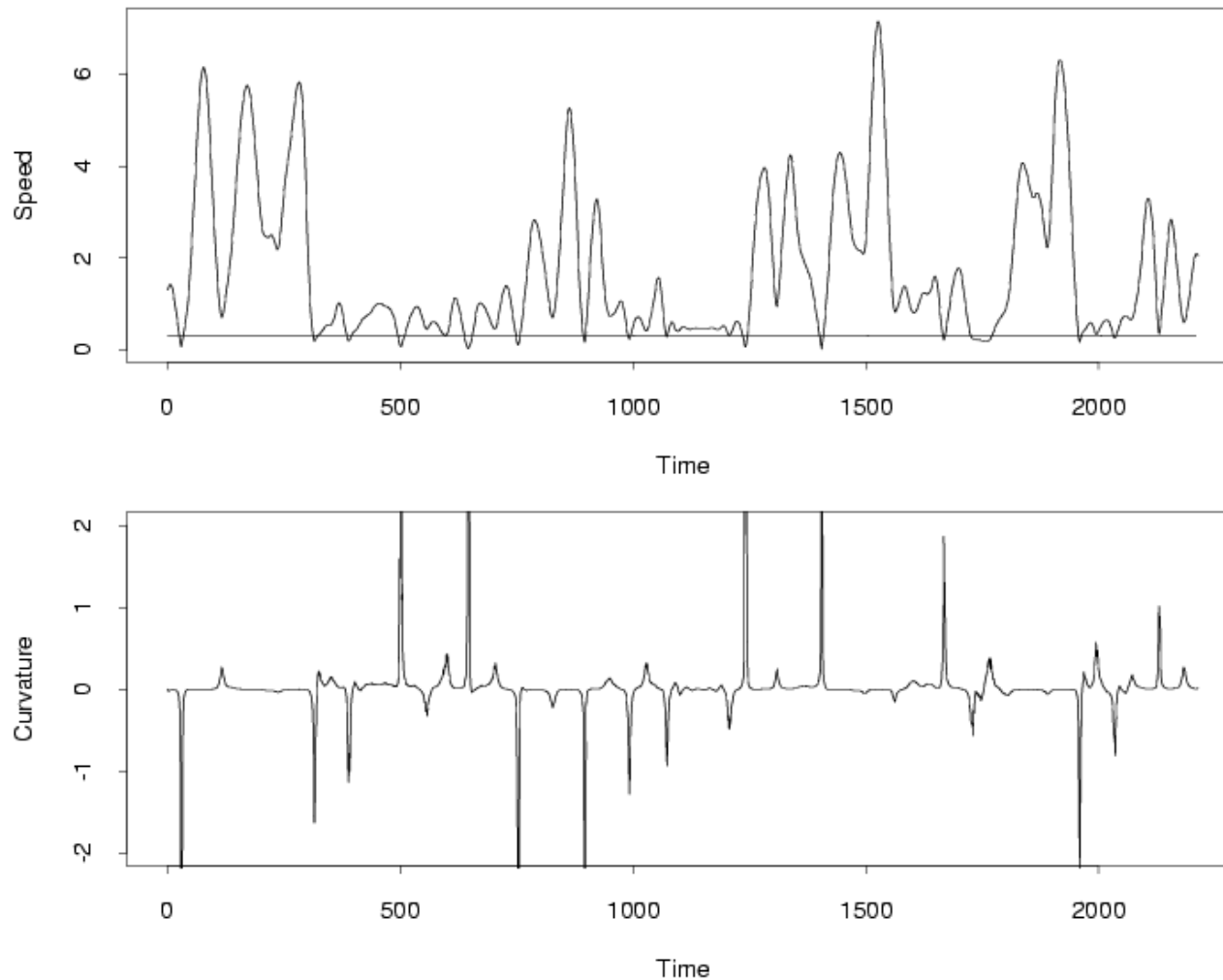
On-line Signaturen halten die Dynamik der Unterschrift fest. Bezogen auf Fälschungen ist es oft der Fall dass je genauer der visuelle Eindruck nachempfunden wird, desto weiter entfernt sich die Dynamik von einer wirklichen Unterschriftsleistung. Ausnahme ist hier eine Fälschung wo der Angreifer das Erstellen der Signatur flüssig erlernt (wesentlich aufwändiger).

Andererseits wird der gute Erfolg von on-line Features insbesondere bei der Fälschungserkennung auch darauf zurückgeführt, dass Fälscher die optische Ähnlichkeit nachzumachen versuchen. Sobald das Ziel einer Fälschung auch dynamische Aspekte umfasst, könnte das Ergebnis anders aussehen. Intra-personal Variability ist höher als bei Off-line Features.

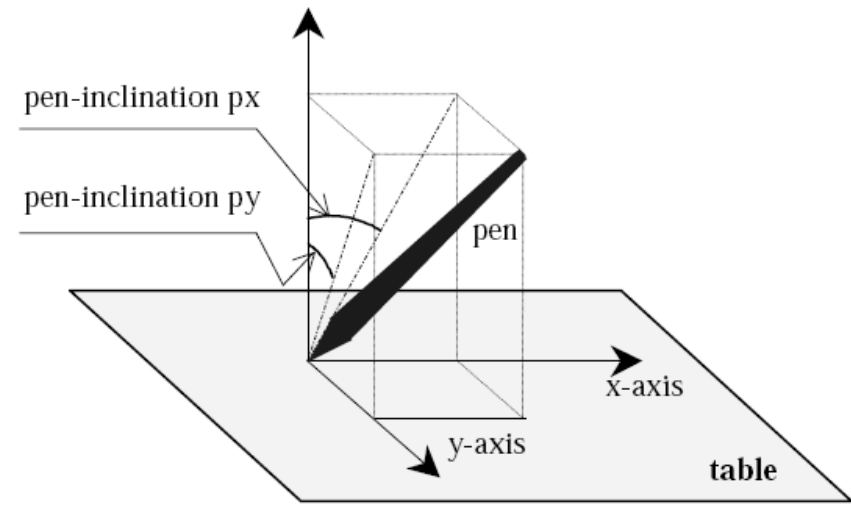
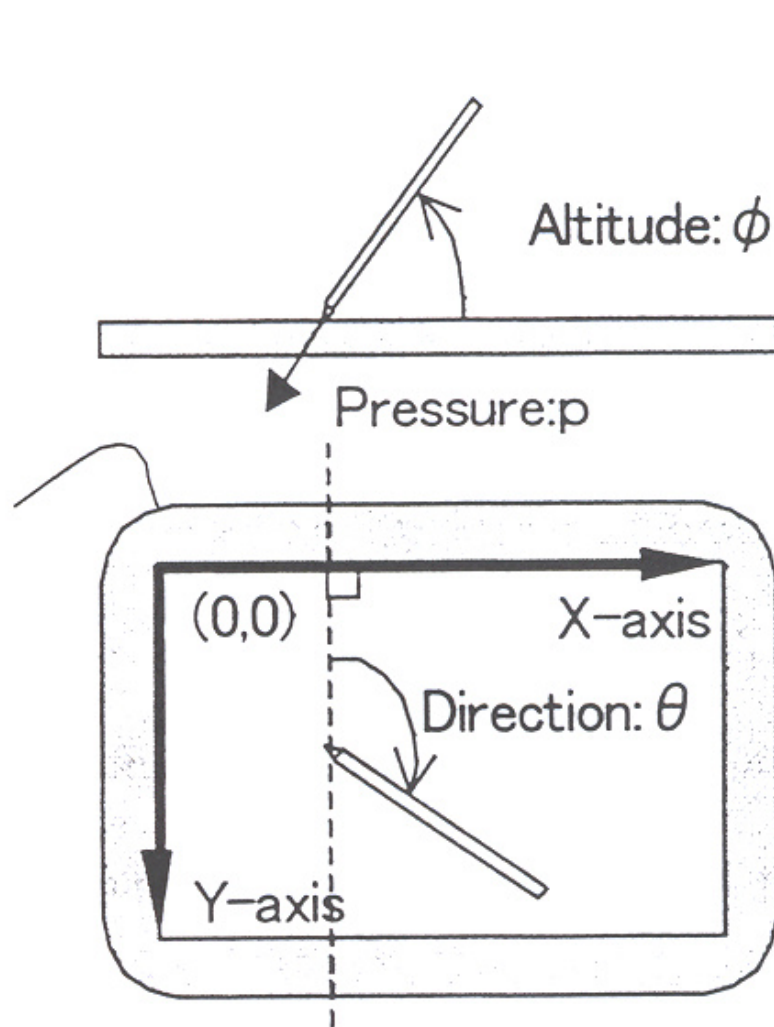
Viele on-line Signatur Verfahren lassen sich ebenfalls mit off-line Features realisieren, da sie nicht wirklich den dynamischen Aspekt berücksichtigen. Werden beispielsweise bei zeitlich geordneten Punkten die x-Koordinate, die y-Koordinate, die Krümmung, u.s.w. (also lokale geometrische Eigenschaften) als Features verwendet, so kann dies auch mit einer Anordnung parametrisiert nach der Bogenlänge (z.B. Pixelcount) des Unterschriftszugs (anstelle der Zeit) realisiert werden (dies entspricht dann einem Re-sampling der Daten). Solche Verfahren sind also nicht intrinsisch on-line und die Features können aus dem Unterschriftsbild generiert werden !

On-Line vs. Off-Line Signaturen II

Oft sind auch dynamische Features (on-line) mit gestalt-basierten (off-line) Features hoch korreliert, wie hier am Beispiel von Geschwindigkeit und Krümmung gut zu sehen ist:



On-Line Signature Feature Recording



Wir haben also folgende Features (zusätzlich zum Unterschriftsbild) potentiell zur Verfügung: $x(t)$, $y(t)$, $p(t)$, $\phi(t)$, und $\Theta(t)$.

Diese Features können durch Sensorik im Tablet oder im Stift aufgezeichnet werden. Beispiel: WACOM ArtPad 2 pro Serial.

Global vs. Local On-Line Features

Globale Features betreffen die Signatur als Ganzes während lokale Features zeitabhängige Eigenschaften an einer Stelle der Signatur sind. In der Literatur werden wesentlich häufiger lokale Features verwendet, auch um Verfahren der Zeitreihenanalyse anwenden zu können.

Beispiele für globale on-line Features:

- Gesamtsignaturzeit, Pen-down Ratio (Pen-Down oder Stroke Zeit durch Gesamtzeit), Anzahl Strokes
- Basierend auf Geschwindigkeit und Beschleunigung: $v_x = \frac{dx}{dt}$, $v_y = \frac{dy}{dt}$, $a_x = \frac{dv_x}{dt}$, $a_y = \frac{dv_y}{dt}$. Die Pfadgeschwindigkeit v etwa ist definiert als $v = (v_x^2 + v_y^2)^{1/2}$. Hier werden dann mittlere Geschwindigkeiten und Beschleunigungen, entsprechende Varianzen, auch Histogramme über die Verteilung derselben und Korrelation zwischen v_x und v_y verwendet.
- Histogramme über $\phi(t)$, $\Theta(t)$ und $p(t)$.
- Durch die Korrelation mit manchen gestalt-bezogenen Features (s.o.) muss darauf geachtet werden, dass im Fall der Verwendung von beiden Typen diese Korrelation durch gute Wahl vermieden wird um echte Vorteile zu haben !

Globale On-Line Features

Ranking	Feature Description	Ranking	Feature Description
1	signature total duration T_s	2	$N(\text{pen-ups})$
3	$N(\text{sign changes of } dx/dt \text{ and } dy/dt)$	4	average jerk \bar{j} [3]
5	standard deviation of a_y	6	standard deviation of v_y
7	(standard deviation of y)/ Δ_y	8	$N(\text{local maxima in } x)$
9	standard deviation of a_x	10	standard deviation of v_x
11	j_{rms}	12	$N(\text{local maxima in } y)$
13	$t(\text{2nd pen-down})/T_s$	14	(average velocity \bar{v})/ $v_{x,\text{max}}$
15	$\frac{A_{\text{min}}=(y_{\text{max}}-y_{\text{min}})(x_{\text{max}}-x_{\text{min}})}{(\Delta_x=\sum_{i=1}^{\text{pen-downs}}(x_{\text{max}} i-x_{\text{min}}))\Delta_y}$	16	$(x_{\text{last pen-up}} - x_{\text{max}})/\Delta_x$
17	$(x_{\text{1st pen-down}} - x_{\text{min}})/\Delta_x$	18	$(y_{\text{last pen-up}} - y_{\text{min}})/\Delta_y$
19	$(y_{\text{1st pen-down}} - y_{\text{min}})/\Delta_y$	20	$(T_w \bar{v})/(y_{\text{max}} - y_{\text{min}})$
21	$(T_w \bar{v})/(x_{\text{max}} - x_{\text{min}})$	22	(pen-down duration T_w)/ T_s
23	$\bar{v}/v_{y,\text{max}}$	24	$(y_{\text{last pen-up}} - y_{\text{max}})/\Delta_y$
25	$\frac{T((dy/dt)/(dx/dt)>0)}{T((dy/dt)/(dx/dt)<0)}$	26	\bar{v}/v_{max}
27	$(y_{\text{1st pen-down}} - y_{\text{max}})/\Delta_y$	28	$(x_{\text{last pen-up}} - x_{\text{min}})/\Delta_x$
29	(velocity rms v)/ v_{max}	30	$\frac{(x_{\text{max}}-x_{\text{min}})\Delta_y}{(y_{\text{max}}-y_{\text{min}})\Delta_x}$
31	(velocity correlation $v_{x,y}$)/ v_{max}^2 [4]	32	$T(v_y > 0 \text{pen-up})/T_w$
33	$N(v_x = 0)$	34	direction histogram s_1 [4]
35	$(y_{\text{2nd local max}} - y_{\text{1st pen-down}})/\Delta_y$	36	$(x_{\text{max}} - x_{\text{min}})/x_{\text{acquisition range}}$
37	$(x_{\text{1st pen-down}} - x_{\text{max}})/\Delta_x$	38	$T(\text{curvature} > \text{Threshold}_{\text{curv}})/T_w$
39	(integrated abs. centr. acc. a_{Ic})/ a_{max} [4]	40	$T(v_x > 0)/T_w$
41	$T(v_x < 0 \text{pen-up})/T_w$	42	$T(v_x > 0 \text{pen-up})/T_w$
43	$(x_{\text{3rd local max}} - x_{\text{1st pen-down}})/\Delta_x$	44	$N(v_y = 0)$
45	(acceleration rms a)/ a_{max}	46	(standard deviation of x)/ Δ_x
47	$\frac{T((dx/dt)(dy/dt)>0)}{T((dx/dt)(dy/dt)<0)}$	48	(tangential acceleration rms a_t)/ a_{max}
49	$(x_{\text{2nd local max}} - x_{\text{1st pen-down}})/\Delta_x$	50	$T(v_y < 0 \text{pen-up})/T_w$

Lokale On-Line Features

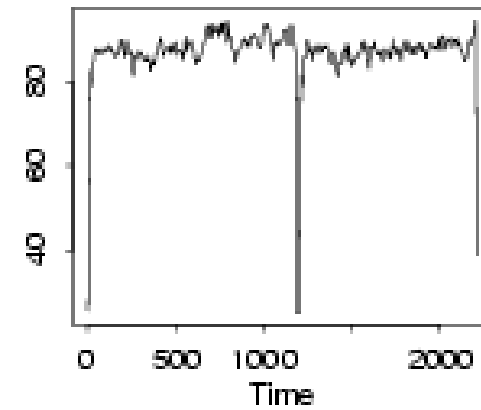
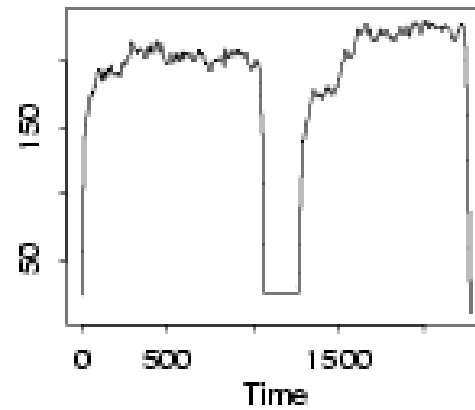
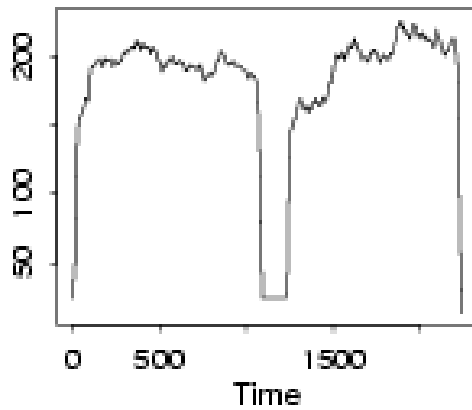
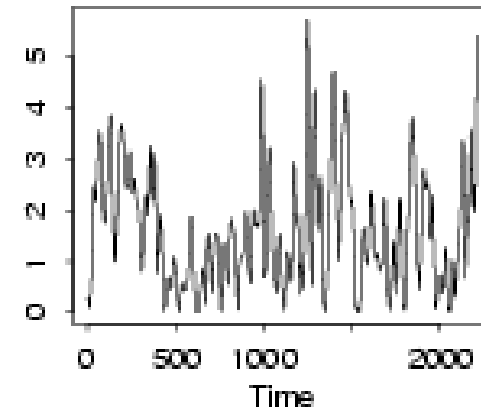
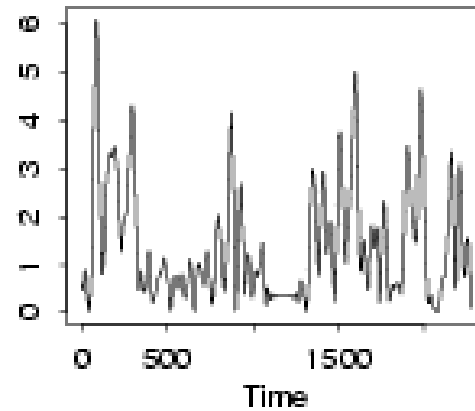
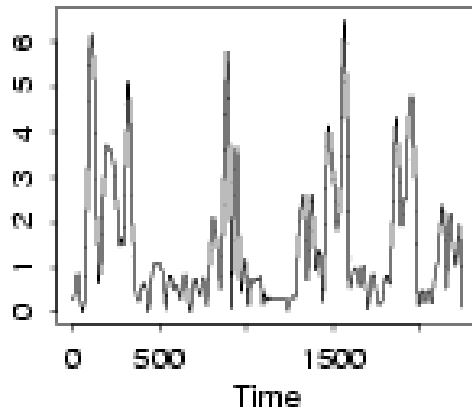
Hier werden klassischerweise zeitabhängige Funktionen betrachtet: $x(t)$, $y(t)$, $p(t)$, $\phi(t)$, $\Theta(t)$, $v_x(t)$, $v_y(t)$, $a_x(t)$, $a_y(t)$.

Weitere verwendete Features sind:

- Pfad Tangentenwinkel $\Phi(t) = \tan^{-1}\left(\frac{v_y(t)}{v_x(t)}\right)$ oder der Winkel $\alpha(t)$ zwischen der Verbindungsgerade von zwei Punkten im Signaturbild zur Zeit t und $t + 1$ und der x-Achse. Entspricht von der Idee her der Krümmung, die hier aber zeitabhängig betrachtet wird.
- Stiftbewegung $V(t)$ als dreidimensionaler Vektor mit $V_x(t) = \sin\Theta(t)\cos\phi(t)$, $V_y(t) = -\cos\Theta(t)\cos\phi(t)$, und $V_z(t) = \sin\phi(t)$.

Um einen sinnvollen Vergleich dieser Zeitreihen zu ermöglichen, müssen einige Schritte im Bereich Vorverarbeitung gesetzt werden: **Resampling, Glättung, optimales Alignment – dynamic time warp DTW.**

Beispiel: Lokale On-Line Features Geschwindigkeit & Druck



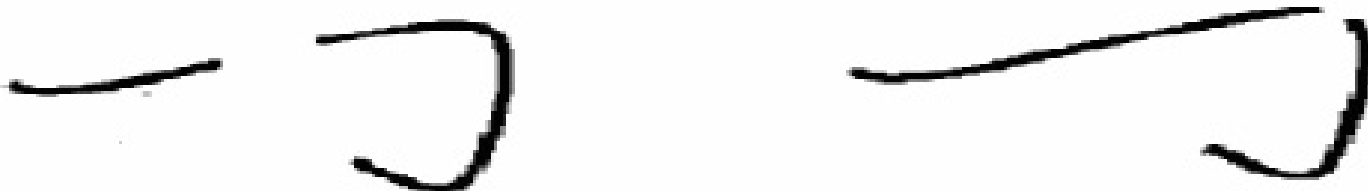
Lokale On-Line Features: Preprocessing

Klassische Schritte beinhalten Positionsnormierung und Grössennormierung. Da diese auch für Off-Line features verwendet werden, zu diesem Thema dort einige Bemerkungen.

Für die Beurteilung von Gestaltmerkmalen (Off-Line Features) müssen dynamische Daten einem uniform Resampling unterzogen werden um temporale Artefakte zu entfernen, On-Line Merkmale keinesfalls !

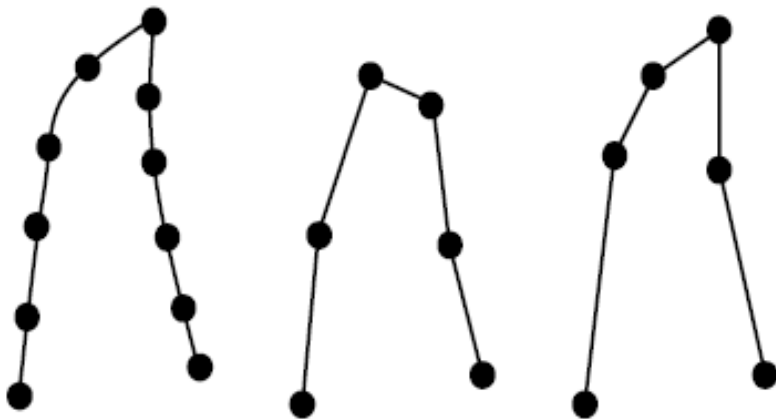
Glättung kann Digitalisierungsrauschen verringern oder entfernen, zu starke Glättung kann aber kontraproduktiv sein (siehe Jitter !).

Strokes: für manche Verfahren werden Signaturkurven immer zu einer geschlossenen Kurve verbunden, in manchen Verfahren nur unter gewissen Bedingungen, um virtual Pen-Ups (durch zu geringen Druck) von echten zu unterscheiden (z.B.: AB und CD sind die Endpunkte von zwei Strokes, ist die Richtung von Vektor BC zwischen den Richtungen von AB und CD werden die Strokes verbunden).



Lokale On-Line Features: Resampling & Jitter

Resampling kann notwendig sein, weil die Abtastrate zu hoch ist (hoher Rechenaufwand) oder für einen Vergleich eine identische Anzahl von Abtastwerten notwendig ist. Hier muss darauf geachtet werden, dass keine wahrnehmungsrelevanten oder sonst wichtigen Punkte ("critical points") entfernt werden. Dies kann abgefangen werden durch $\left| \frac{y(t_i) - y(t_{i-1})}{x(t_i) - x(t_{i-1})} \right| \leq T$, ansonsten wird anders gesampled.

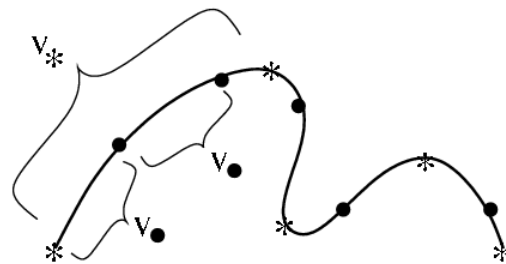


Während eine Gättung möglichst Rauschen vermindern sollte, gibt es auch sog. "Jitter" der auf eine versuchte Fälschung hindeutet. Jitter entsteht dadurch, dass der Fälscher in kleinem Massstab der Signaturkurve zu folgen versucht und immer wieder kleine Richtungskorrekturen anbringen muss. Signaturen sollten auf Jitterverdacht untersucht werden um mögliche Fälschungen früh zu identifizieren.



Significant Points und Segmentierung

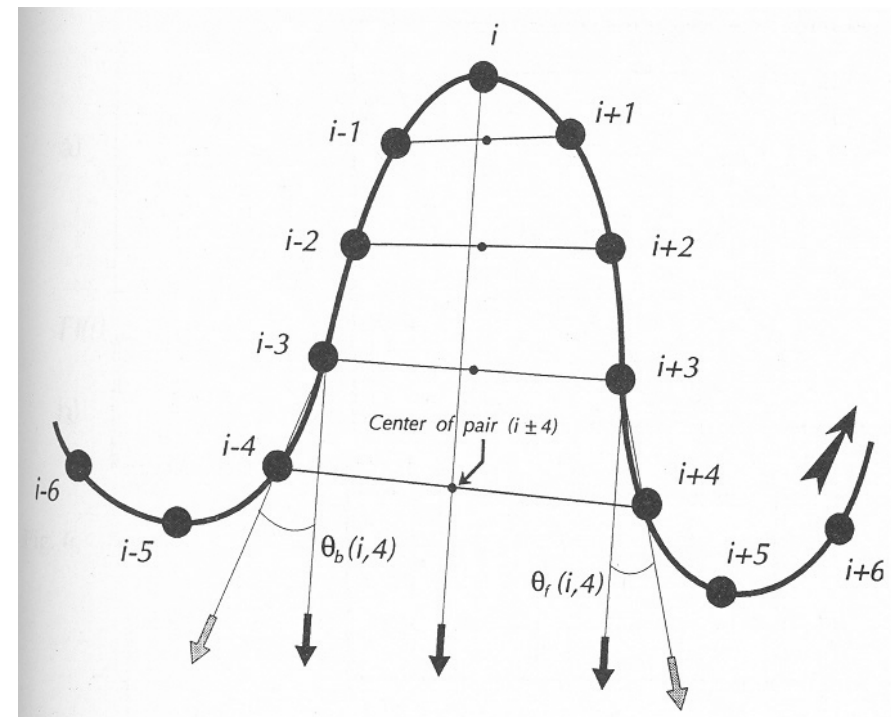
Unterschriften weisen eine Menge von besonders bedeutsamen Punkten auf, deren Eigenschaften wesentlich sind für das Erscheinungsbild und die Dynamik der Unterschrift, sog. signifikante Punkte (besondere Punkte im Unterschriftsbild) oder Extrempunkte (lokale Maxima in einer zeitabhängigen Funktion, z.B. zeitabhängige Krümmung). Solche Punkte sind auch oft Scheitelpunkte und können für Verifikation benutzt werden: durch bestimmte Eigenschaften des Unterschriftszugs zwischen diesen Punkten (siehe Beispiel: Geschwindigkeit zwischen solchen Punkten) oder die Grösse/Lage/Beschaffenheit der Umgebung dieser Punkte. Bestimmung dieser Punkte liefert implizit immer eine Segmentierung einer Unterschrift (einfachste Segmentierung: Stroke-basiert, siehe vorne). Ein Feature pro Segment liefert einen kurzen und potentiell sehr aussagekräftigen Featurevektor.



- * = critical point v_* = speed between two critical points
• = sampling point v_\bullet = speed between two sampling points

Finden von Signifikanten Punkten I

Für jeden Punkt im Signaturbild i wird bestimmt, ob Nachbarpunkte $i \pm n$ (Punktpaare) in seinem Einflussbereich liegen oder nicht. Dafür werden die Winkel $\theta_f(i, n)$ und $\theta_b(i, n)$ bestimmt wie folgt: eine Gerade g wird zwischen den Punkten $i - n$ und $i + n$ gezogen und die Verbindungsgerade h zwischen dem Halbierungspunkt von g und dem Punkt i erstellt. h wird dann parallelverschoben und zwei Gerade durch die Punkte $i - n + 1$ und $i + n - 1$ gezogen. $\theta_b(i, n)$ ist der Winkel zwischen der zu h parallelen Gerade durch den Punkt $i - n + 1$ und der Verbindungsgerade der Punkte $i - n$ und $i - n + 1$, $\theta_f(i, n)$ ist der Winkel zwischen der zu h parallelen Gerade durch den Punkt $i + n - 1$ und der Verbindungsgerade der Punkte $i + n$ und $i + n - 1$.

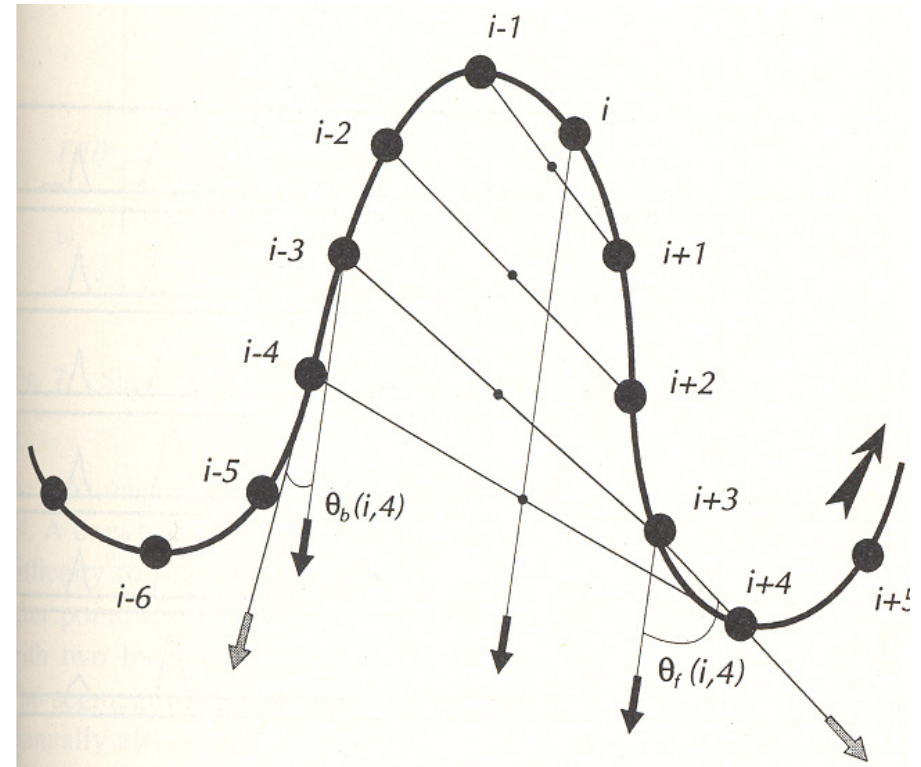


Finden von Signifikanten Punkten II

Je mehr sich diese beiden Winkel 90 Grad annähern, desto weniger signifikant ist i bezüglich seiner Nachbarn. Diese Eigenschaft wird ausgenutzt um zu bestimmen ob ein Punktpaar $i \pm n$ im Einflussbereich von i liegt: Beide Winkel müssen kleiner als eine Schranke sein, die sich zwischen 0 und 90 Grad bewegt. Die Bedeutung des Beitrags des Punktpaares wird nun wie folgt bestimmt:

$$IMP(i, n) = \cos(\theta_b(i, n)) * \cos(\theta_f(i, n))$$

Sind die Winkel sehr klein macht das die Bedeutung von i sehr gross, die Multiplikation zeigt, dass beide Winkel klein sein müssen um einen ausgeprägten Scheitelpunkt zu erhalten.



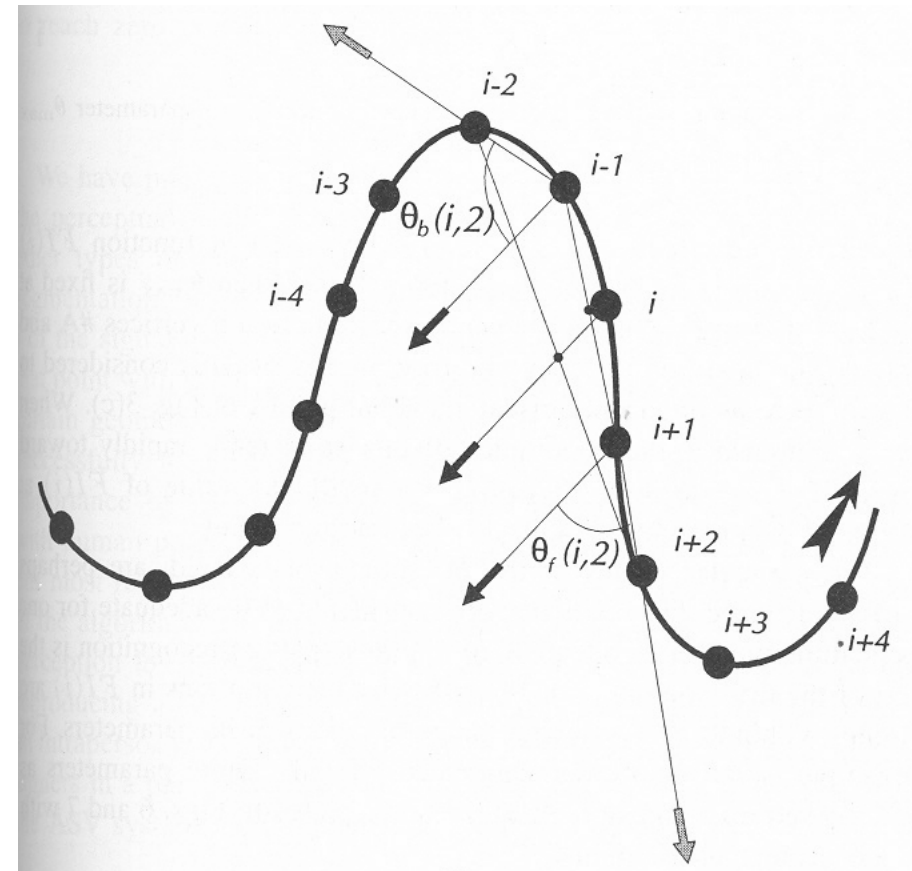
Finden von Signifikanten Punkten III

Um nun die Beiträge aller Punkte im Einflussbereich von i zu bestimmen, wird $IMP(i, n)$ über alle N Punktpaare in der Nachbarschaft von i summiert, deren Winkel $\theta_f(i, n)$ und $\theta_b(i, n)$ kleiner als obengenannte Schranke sind:

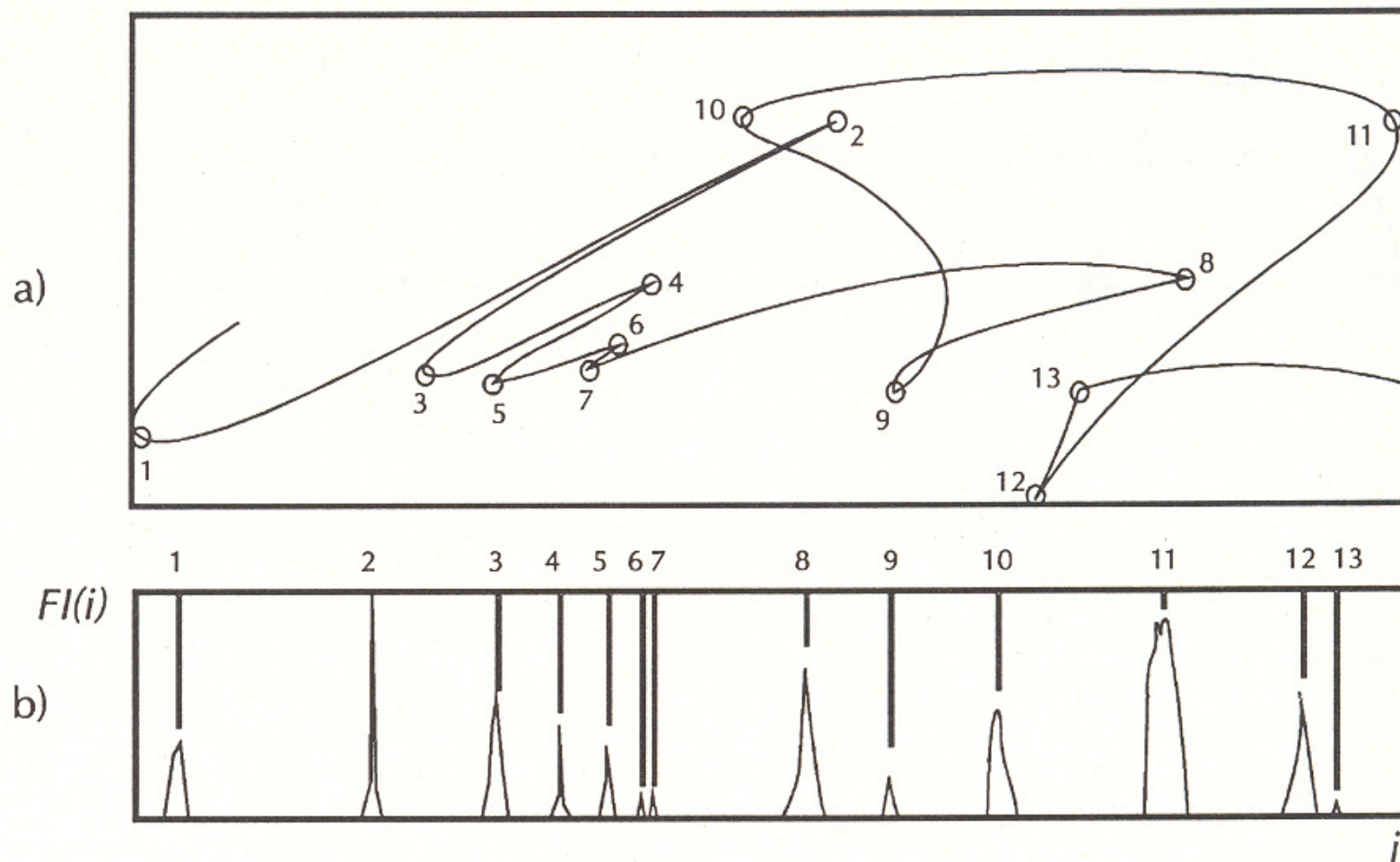
$$FI(i) = \sum_{i=1}^N IMP(i, n)$$

Für die Bestimmung der signifikanten Punkte wird $FI(i)$ für alle Punkte i berechnet, im Anschluss sind die lokalen Maxima dieser Funktion die gesuchten signifikanten Punkte.

Spezielle Beachtung muss breiten Scheiteln gewidmet werden, da die direkten Nachbarn von i in diesem Fall die Signifikanz stark beeinträchtigen (das muss in der Berechnung berücksichtigt werden).



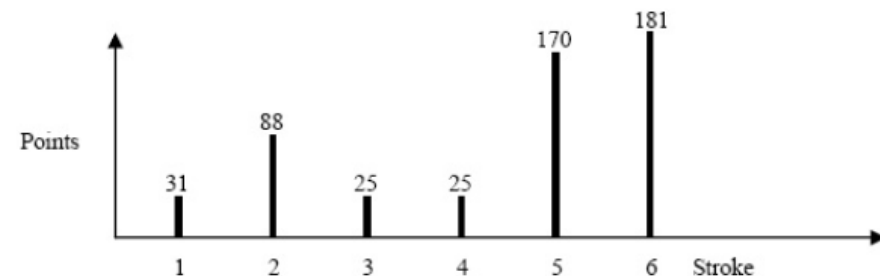
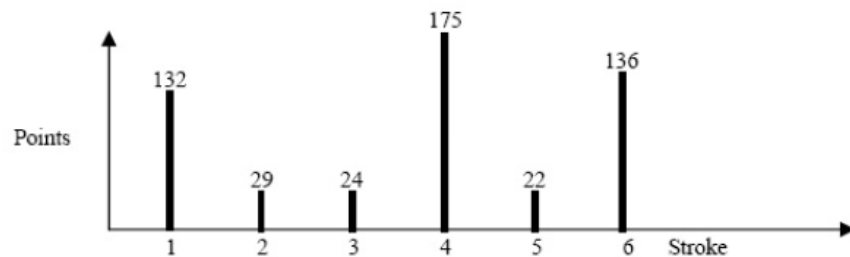
Finden von Signifikanten Punkten: Beispiel



Matching von Segmenten

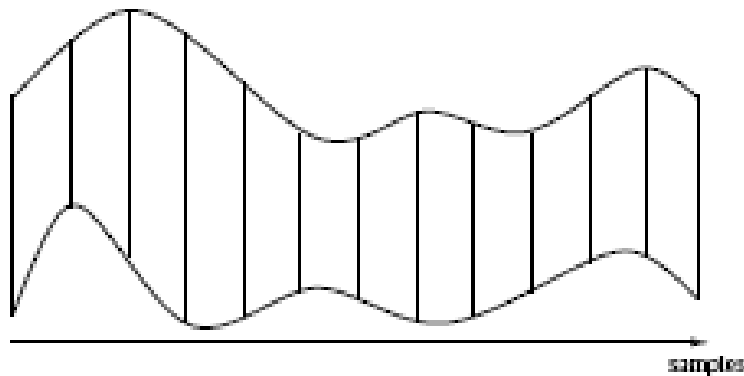
Eine Schwäche von segmentierungsbasierten Verifikationsverfahren ist, dass bei falscher Segmentierung extrem schlechte Matchingwerte zu erwarten sind. Eine mögliche Strategie ist es, Segmente bei schlechten Matchwerten zu rekombinieren und erneut auf Ähnlichkeit zu testen. Begonnen wird hier typischerweise mit dem längsten Segment.

Im Beispiel entspricht stroke 1 der ersten Signatur strokes 1,2 der zweiten Signatur und strokes 5 und 6 in Signatur 1 entsprechen stroke 6 in Signatur 2.

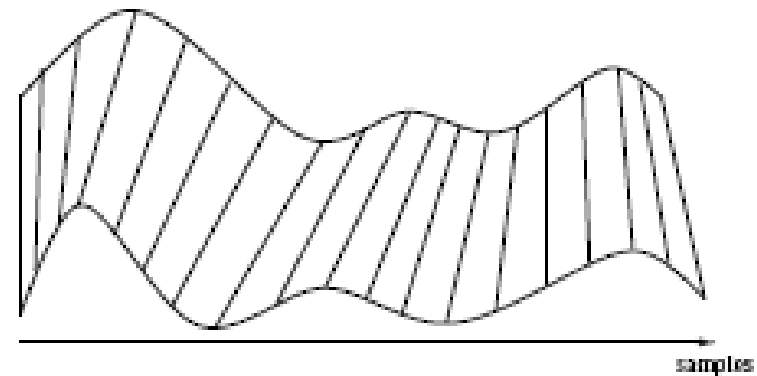


Dynamic Time Warp (DTW)

Wie auch im vorigen Beispiel ersichtlich, bestehen einzelne Segmente aber auch gesamte Unterschriften praktisch nie aus gleich vielen Samplewerten. Dies kommt von der Intra-personal Variability insbesondere bei dynamischen Merkmalen. Die Daten sind aber meistens nicht gleichmässig verzerrt sondern nicht-linear und lokal unterschiedlich. Dies führt häufig zu unverhältnismässig schlechten Matching Werten und einer entsprechend hohen FNMR.



(a) naive alignment after resampling,



(b) alignment with DTW.

DTW passt zwei zu matchende Unterschriften durch nichtlineares warping bestmöglich aneinander an.

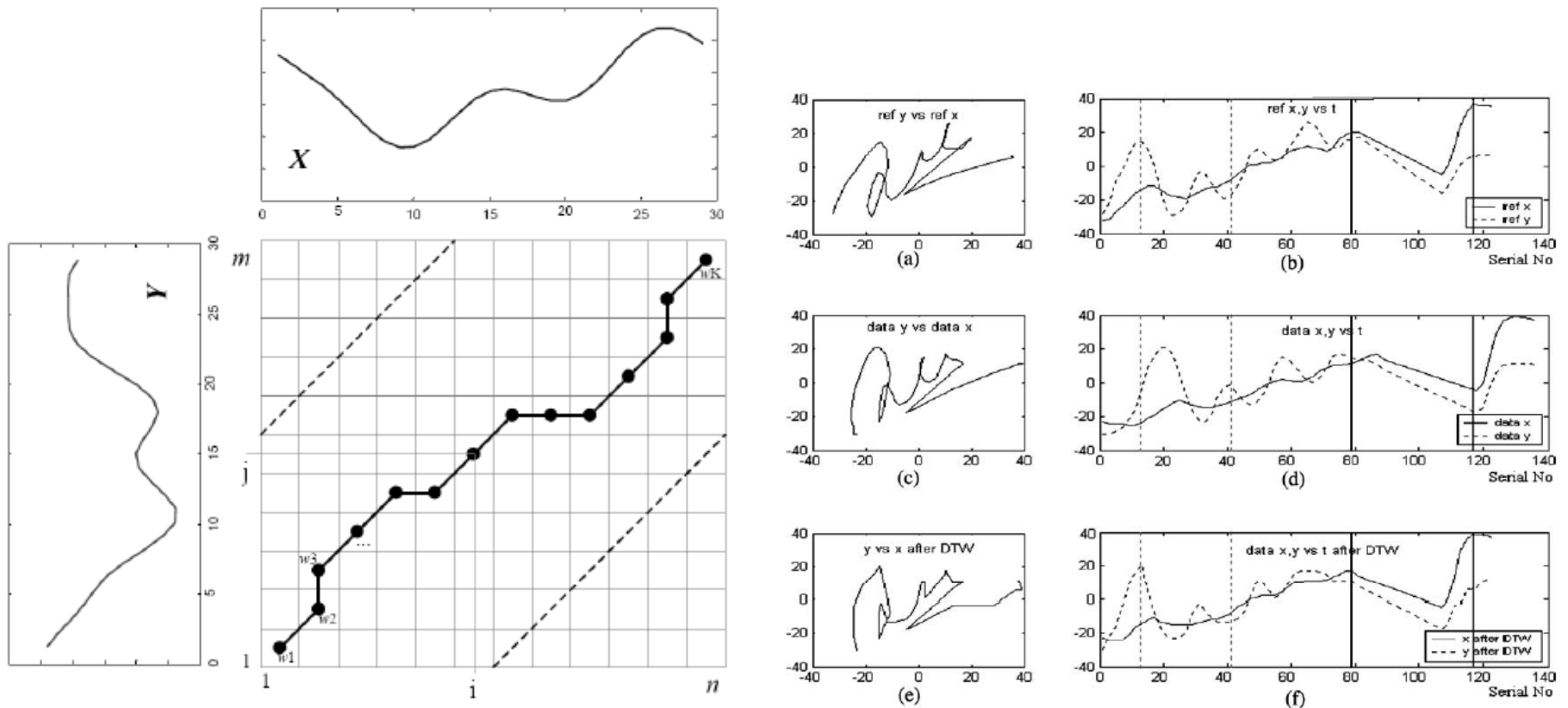
DTW: Grundprinzip

Gegeben sind zwei Zeitreihen $X = x_1, x_2, \dots, x_n$ und $Y = y_1, y_2, \dots, y_m$. Um diese Zeitreihen auszurichten wird eine $n \times m$ Matrix d konstruiert mit $d(i, j) = (x_i - y_j)^2$ (typische aber nicht notwendige Wahl). Ein Warping Pfad $W = w_1, w_2, \dots, w_k$ ist eine zusammenhängende Menge von Matrixelementen die eine Abbildung zwischen X und Y definiert. Das l -te Element von W ist definiert als $w_l = (i, j)_l$. Es gibt verschiedene Einschränkungen bezüglich des Verlaufs von W die je nach Anwendung angewendet werden:

- $w_1 = (1, 1)$ und $w_k = (m, n)$: Anfangs- und Endpunkte der Zeitreihen werden angeglichen. Der Pfad muss an gegenüberliegenden Ecken der Matrix d enden.
- Stetigkeit: sei $w_k = (a, b)$. Dann gilt für $w_{k+1} = (c, d)$ die Eigenschaft: $c - a \leq 1$ und $d - b \leq 1$. Der Pfad bewegt sich nur zu anliegenden oder diagonal benachbarten Zellen.
- Monotonie: sei $w_k = (a, b)$. Dann gilt für $w_{k+1} = (c, d)$ die Eigenschaft: $c - a \geq 0$ und $d - b \geq 0$.

Es gibt exponentiell viele Pfade W , das Interesse gilt allerdings demjenigen der den Abstand zwischen X und Y minimiert.

DTW: Grundprinzip Visualisierung

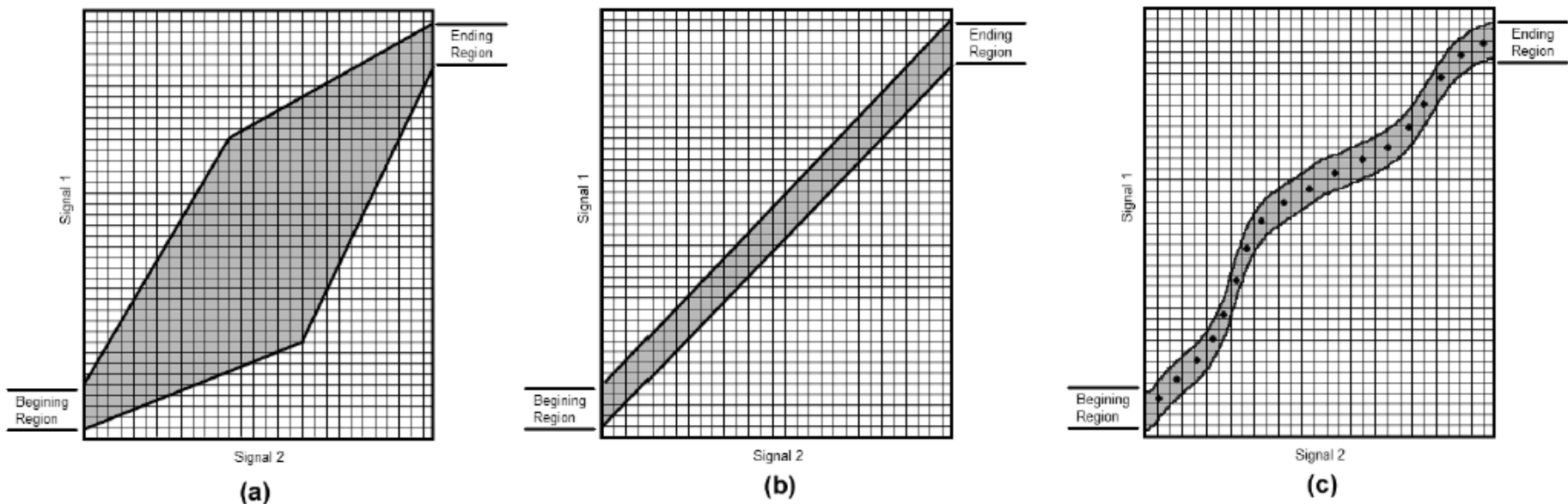


Beispiel rechnen: $X = \{2, 4, 8, 13, 9, 5, 8, 12, 15, 18\}$ und $Y = \{2, 3, 5, 9, 12, 8, 4, 9, 16\}$

DTW: Dynamic Programming

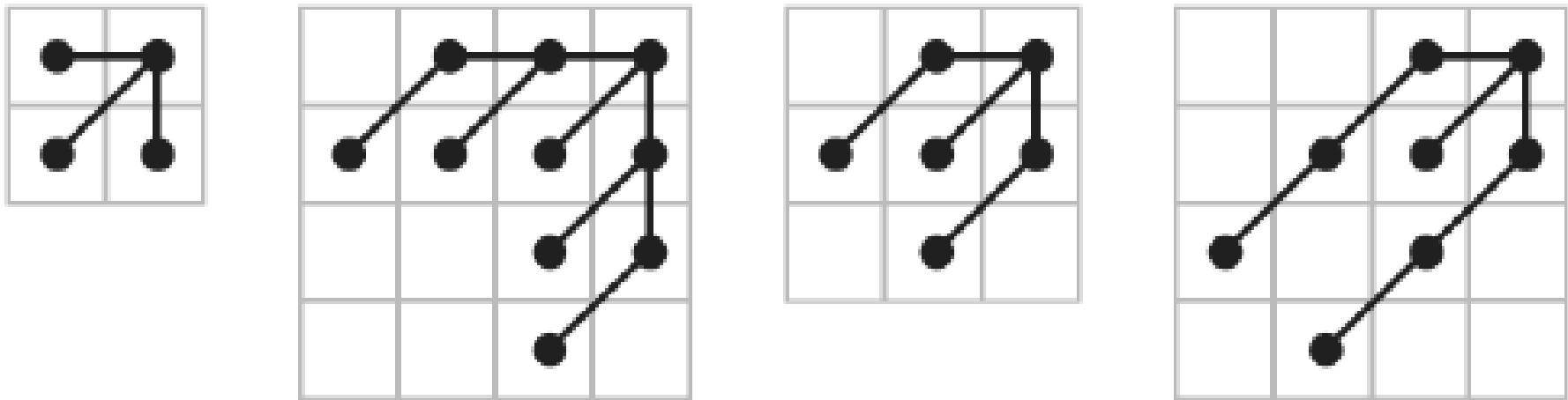
Um den minimalen Warping Pfad zu ermitteln wird eine Matrix $D(i, j)$ wie folgt erstellt:
 $D(i, j) = d(i, j) + \min(D(i-1, j), D(i, j-1), D(i-1, j-1))$. Initialisiert wird das Verfahren mit $D(1, 1) = d(1, 1)$. Um den optimalen Pfad finden zu können muss in jeder Zelle mitgespeichert werden was der bisher minimale Kostenpfad war. Der optimale Pfad wird dann mittels Backtracking gefunden.

Um den Zeitaufwand für die Berechnung von D zu beschränken und zu starke Verzerrung zu vermeiden, wurden mehrere Strategien vorgeschlagen: Beschränkung des Pfades auf bestimmte Regionen oder Beschränkung der Weite eines einzelnen Schrittes.



DTW: Erweiterungen

In der Konstruktion von $D(i,j)$ gibt es noch viele weitere Möglichkeiten als die eine gezeigte Variante (sie entspricht der linken Graphik). Der Einsatz solcher spezifischen Muster ist natürlich anwendungsabhängig.

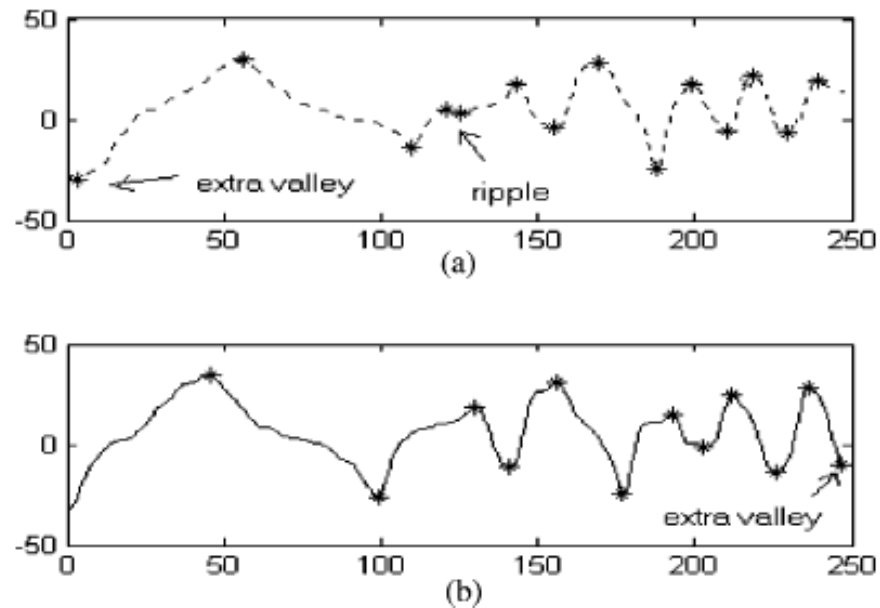


Auch andere Kostenfunktionen können verwendet werden, z.B. können besonders flache oder steile Pfadteile extra bestraft werden oder auch besonders kurze Pfade belohnt werden.

DTW hat aber ein grosses Problem in Zusammenhang mit Signaturen: auch Fälschungen werden ausgerichtet und können dadurch ein gutes Ergebnis erzielen. Dies wird im Folgenden behandelt.

Extreme Point (EP) DTW

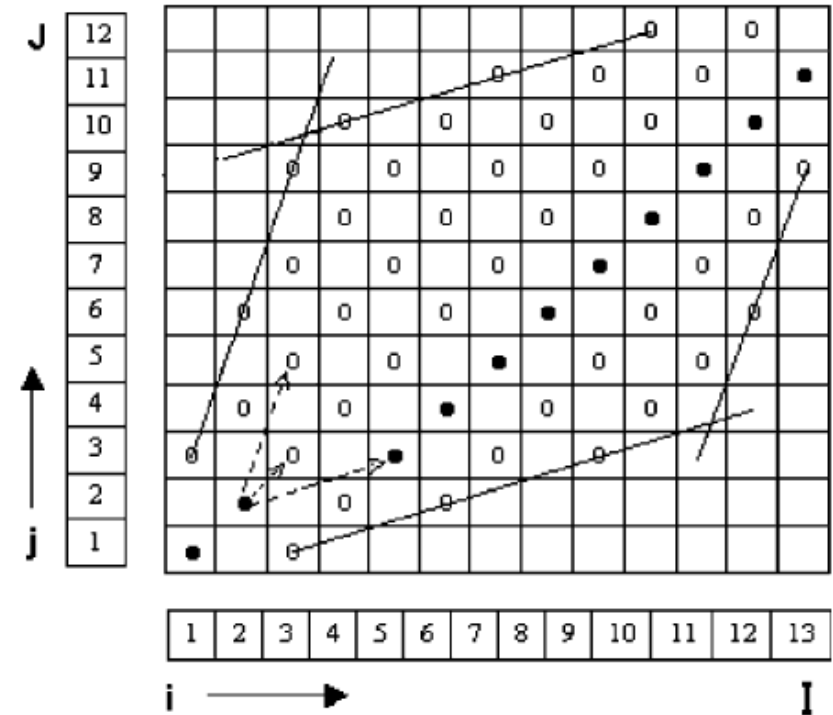
Zielsetzung ist neben der besseren Sensitivität gegenüber Fälschungen auch eine Rechenzeitverkürzung. Grundidee ist es, nur lokale Extremwerte von Zeitreihen einem DTW zu unterziehen. Voraussetzung ist dass “zu lokale” Extrema ausgeschlossen werden (ein Mindestmass an Berg- oder Taleigenschaft wird verlangt). Die EPs müssen für korrekte Ausrichtung immer Paare von Maxima oder Minima sein. Probleme gibt es dabei am Anfang und Ende und durch sog. “ripples”.



Das Verfahren läuft wie klassisches DTW wobei die Matrix nur EPs enthält und spezielle lokale Pfadverläufe definiert werden, die das Überspringen von ripples erlaubt (was in der Kostenberechnung extra bestraft wird).

EPDTW Beispiel

Die in der Matrix eingetragenen Kreise sind alle (innerhalb der Pfadgrenzen) möglichen Maxima-Maxima oder Minima-Minima Paare, von diesen werden die globalen Kosten im Sinn von $D(i, j)$ berechnet. Die schwarzen Kreise entsprechen dem minimalen Kostenpfad (der durch das ripple-Paar unterhalb der Hauptdiagonale verläuft).

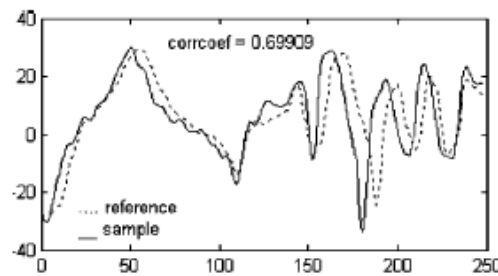


EPDTW Segment Warping & Visualisierung

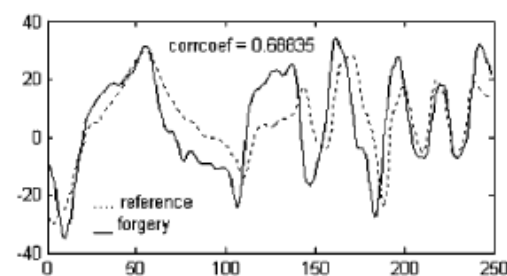
Nach der korrekten Ausrichtung der Extremwerte werden die dazwischenliegenden Segmente linear gestreckt/gestaucht: seien (X_n, Y_n) und (X_{n+1}, Y_{n+1}) die EPs der Referenzzeitreihe auf die die EPs des Samples (x_n, y_n) und (x_{n+1}, y_{n+1}) gewarped wurden. Durch DTW gilt: $x'_n = X_n$ und $x'_{n+1} = X_{n+1}$. Für einen Wert zwischen den EPs gilt:

$$x'_j = X_n + (x_j - x_n) \frac{X_{n+1} - X_n}{x_{n+1} - x_n}$$

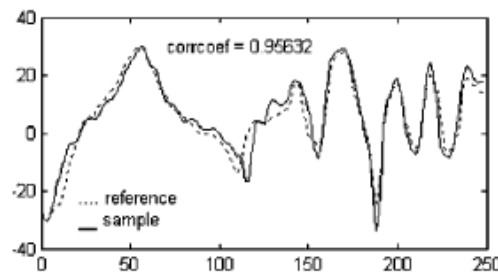
Im Gegensatz zum klassischen DTW wird hier lokale Kurvenform weniger zerstört.



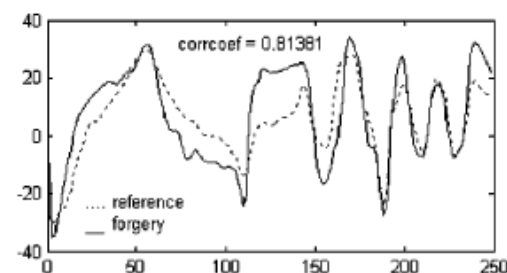
(a)



(b)

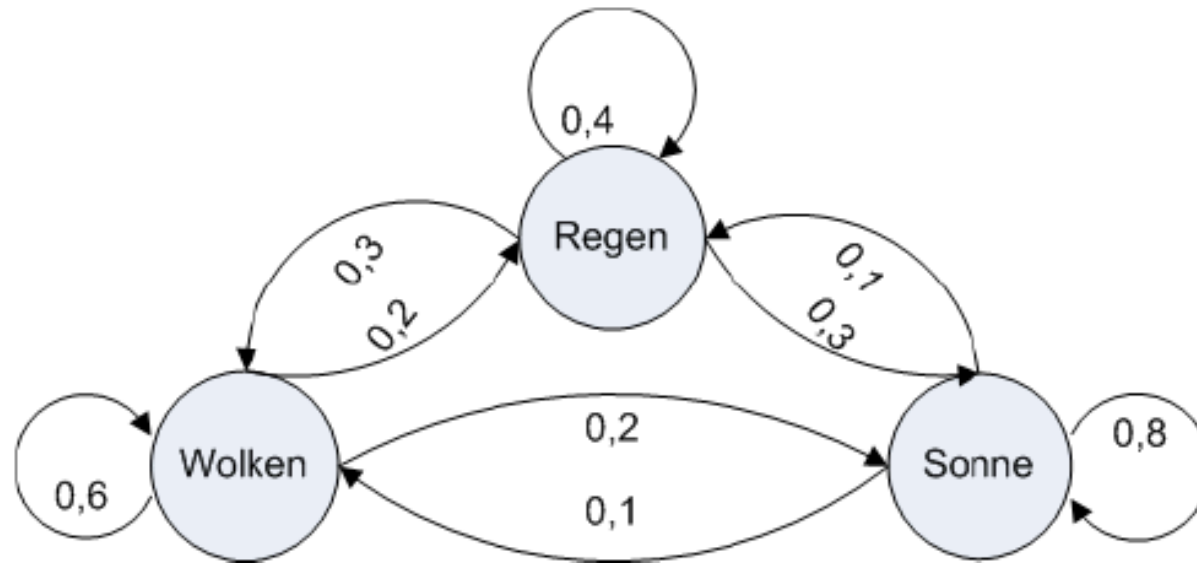


(c)



(d)

Discrete Markov Models I



Bei der Beobachtung eines Systems will man oft wissen, in welchem Zustand es sich zum Zeitpunkt der Beobachtung befindet. Zum Beispiel kann das Wetter das System sein und die Frage ist wie das Wetter morgen wahrscheinlich sein wird. Im Beispiel gibt es drei Zustände, die mit bestimmter Wahrscheinlichkeit aufeinander folgen.

Definition: ein diskretes Markov Modell ist ein System von n Zuständen $\omega(i)$. Zu diskreten Zeitpunkten t geht das System mit Übergangswahrscheinlichkeiten $a_{ij} = P(\omega_{t+1}(j) | \omega_t(i))$ vom Zustand $\omega(i)$ in den Zustand $\omega(j)$ über. Diese Übergangswahrscheinlichkeiten werden in einer Zustandsübergangsmatrix $A = (a_{ij})$ zusammengefasst. Es gilt $\sum_{j=1}^n a_{ij} = 1$ für alle i .

Discrete Markov Models II

Mit diesem System kann man die Wahrscheinlichkeit mit der ein bestimmter Zustand vorliegt ausrechnen. Dazu muss allerdings bekannt sein, wie diese Wahrscheinlichkeiten zum Zeitpunkt der Systeminitialisierung waren. Das wird durch die "Anfangsverteilung" Π beschrieben. $\Pi = (\pi_1, \dots, \pi_n)$ mit $\pi_i = P(\omega_1 = \omega_1(i))$. Ein Markov Modell wird durch das Tupel (A, Π) vollständig beschrieben. Wesentlich: die Gedächtnislosigkeit des Systems, d.h. ein Zustand hängt nur vom vorherigen Zustand ab.

Beispiel: Wie gross ist die Wahrscheinlichkeit für SSSRRSWS wenn am ersten Tag die Sonne scheint ? $P(SSSRRSWS|Modell) = ??$

$$P(S)P(S|S)P(S|S)P(R|S)P(R|R)P(S|R)P(W|S)P(S|W) = \pi_s a_{ss} a_{ss} a_{sr} a_{rr} a_{rs} a_{sw} a_{ws} = 1$$

0.8 0.8 0.1 0.4 0.3 0.1 0.2 = irgendwas.

Hidden Markov Models I

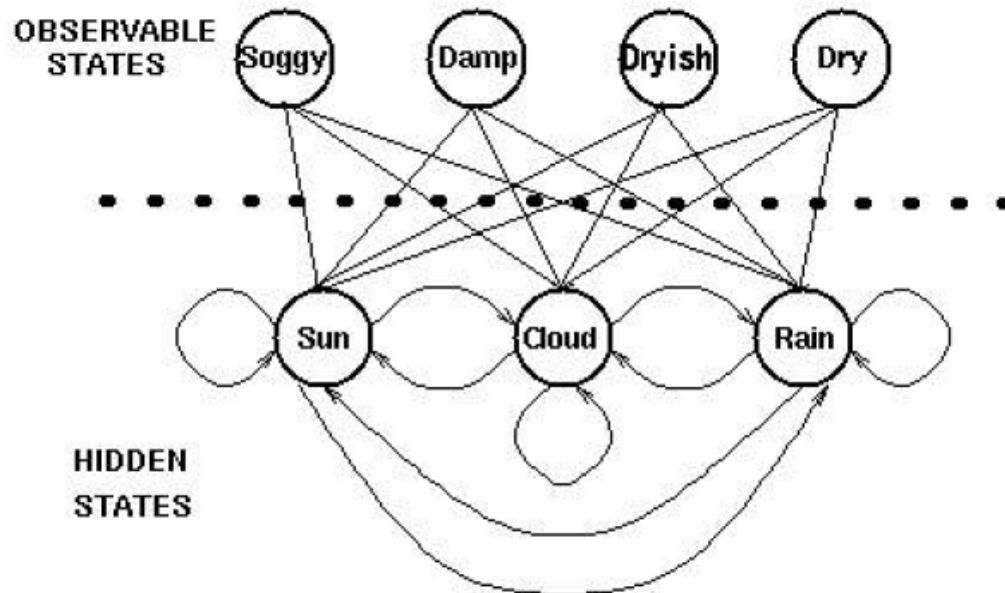
Oft ist es jedoch nicht möglich ein System direkt zu beobachten sondern nur die Auswirkungen auf die Umwelt sind beobachtbar. Beispielsweise könnte man versuchen das Wetter zu bestimmen indem wir die Feuchtigkeit eines Holzstücks das im Freien liegt feststellen und daraus auf das Wetter zurückzuschliessen. Das Wetter selbst ist nicht beobachtbar ("hidden states" - wenn man sich z.B. im Haus befindet und jemand das Holz herein bringt) sondern nur seine Auswirkungen auf das Holz.

Definition: Ein Hidden Markov Model besteht aus n Zuständen $\omega(i)$ die nicht direkt beobachtet werden können. Jeder dieser Zustände emittiert zum Zeitpunkt t einen von $1 \leq k \leq m$ sichtbaren Zuständen (Symbol) $v_t(k)$, das gesamte System generiert beim Durchlaufen von T verborgenen Zuständen die Sequenz $V^T = \{v_1(k), \dots, v_T(k)\}$.

Für die Übergangswahrscheinlichkeiten gilt wie vorher $a_{ij} = P(\omega_{t+1}(j) | \omega_t(i))$. Die Wahrscheinlichkeit für die Emission eines bestimmten Symbols $v_t(k)$ zum Zeitpunkt t wenn sich das System im Zustand $\omega_t(j)$ befindet ist $b_{jk} = b_j(v_t(k)) = P(v_t(k) | \omega_t(j))$. Diese Wahrscheinlichkeiten sind durch die Gedächtnislosigkeit nicht von t abhängig und werden in der Matrix $B = b_{jk}$ zusammengefasst. Es gilt $\sum_{k=1}^m b_{jk} = 1$ für alle j .

Hidden Markov Models II

Die hidden states sind im Beispiel Sonne, Wolken und Regen sowie die Beobachtungen staubtrocken, trocken, feucht und nass mit den entsprechenden Zustandsübergangs- A und Emissionswahrscheinlichkeitsmatrizen B .



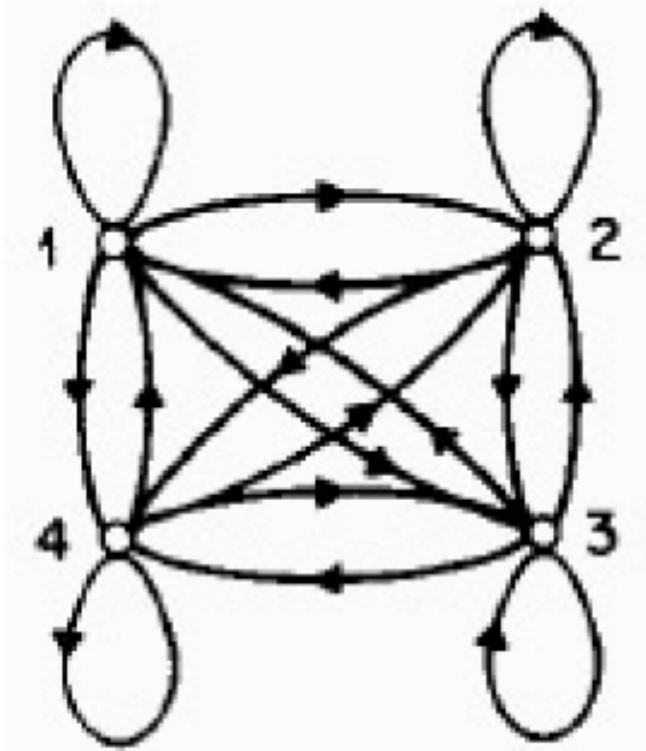
$$A = \begin{pmatrix} 0,5 & 0,25 & 0,25 \\ 0,375 & 0,125 & 0,375 \\ 0,125 & 0,625 & 0,375 \end{pmatrix}$$

$$B = \begin{pmatrix} 0,6 & 0,2 & 0,15 & 0,05 \\ 0,25 & 0,25 & 0,25 & 0,25 \\ 0,05 & 0,1 & 0,35 & 0,5 \end{pmatrix}$$

Dies bedeutet dann beispielsweise dass im Zustand Sonne das Holz mit $P = 0.6$ staubtrocken ist.

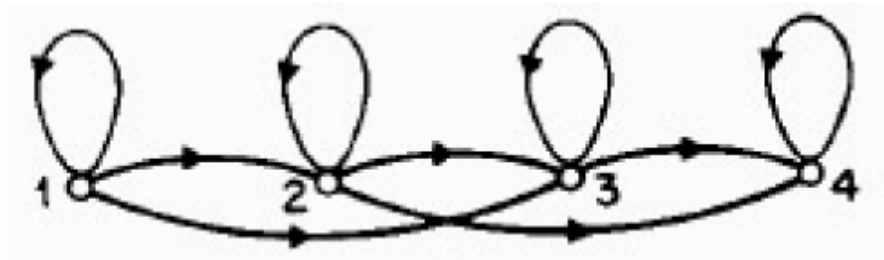
Ein HMM ist durch das Tripel (A, B, Π) charakterisiert.

Hidden Markov Models: Topologien



Ergodische HMM: $a_{ij} \neq 0, \forall(i, j)$

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}$$



Links-Rechts HMM: Bandmatrix $\neq 0$

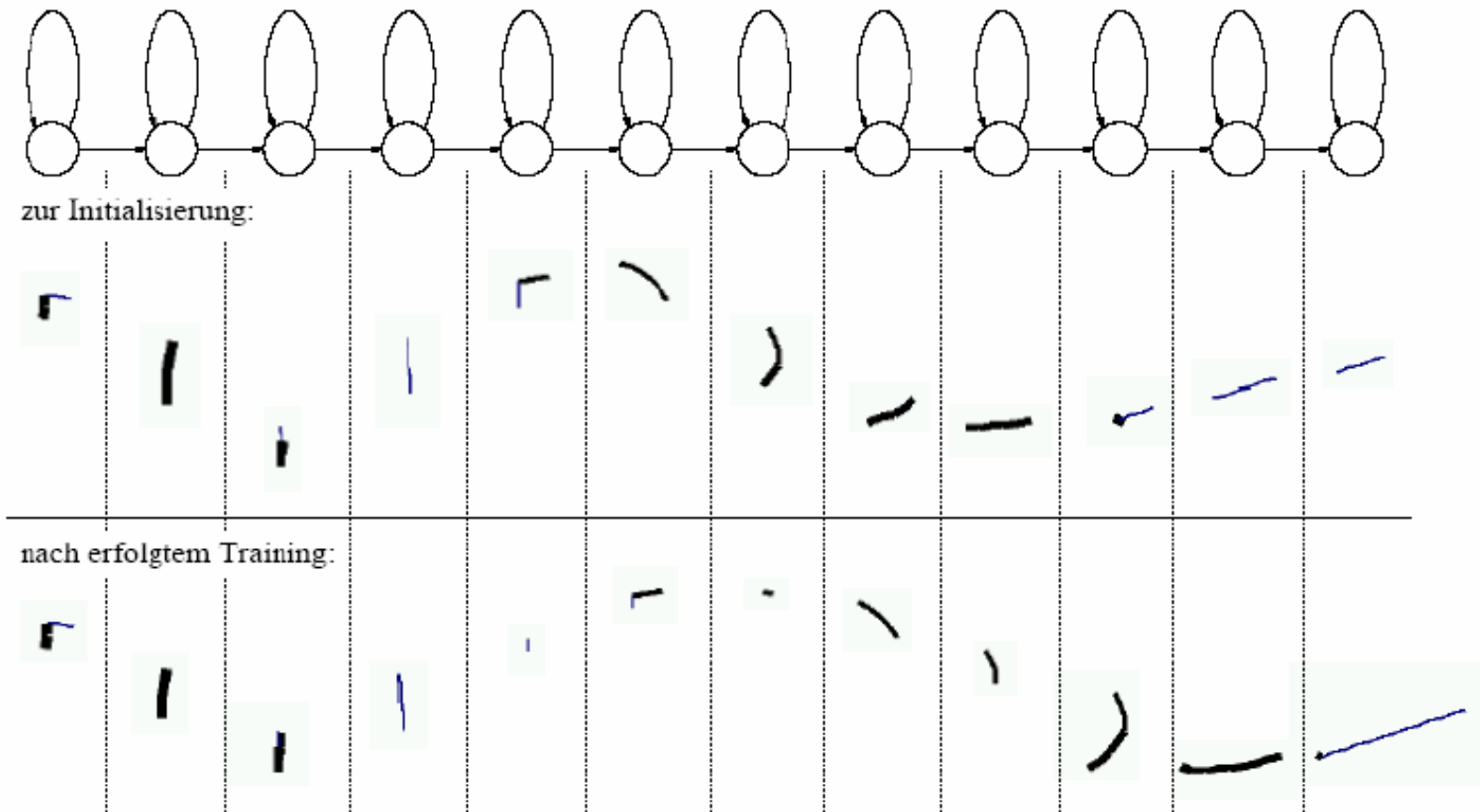
$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & 0 \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & 0 & a_{44} \end{pmatrix}$$

Grundlegende Fragestellungen bei HMM

1. Evaluation: gegeben sei ein HMM $= (A, B, \Pi)$ und eine Beobachtungssequenz V^T : berechne die Wahrscheinlichkeit dass V^T vom geg. HMM erzeugt wurde.
2. Dekodierung: gegeben sei ein HMM $= (A, B, \Pi)$ und eine Beobachtungssequenz V^T : berechne die Folge von verborgenen Zuständen die V^T mit der höchsten Wahrscheinlichkeit erzeugt hat.
3. Lernen: für ein HMM ist die Anzahl der sichtbaren und unsichtbaren Zustände bekannt und eine oder mehrere Beobachtungssequenzen (Trainingssequenzen) verfügbar: berechne die optimalen Parameter A und B.

Die Beobachtungssequenzen entsprechen den berechneten Features der benutzen biometrischen Merkmale und die hidden states entsprechen dem tatsächlichen Merkmal. Einzelne states entsprechen Segmenten der Signatur. Für eine konkrete Signatur wird ein HMM trainiert. Lernen/Training findet beim Enrollment statt. Evaluation bewertet die Features gegenüber einem gelernten HMM (bei Verifikation !), Dekodierung bestimmt ob das biometrische Merkmal bei den beobachteten Features wirklich das wahrscheinlichste ist.

Signaturen und HMM: Beispiel



Hidden Markov Models: Evaluation

Um zu berechnen mit welcher P ein HMM eine beobachtete Sequenz V^T generiert hat könnte man für alle möglichen Sequenzen die Wahrscheinlichkeit des Durchlaufens vom ersten zum letzten Zustands und gleichzeitiger Emission der sichtbaren Sequenz berechnen. Diese Wahrscheinlichkeiten addiert ergeben die Gesamtwahrscheinlichkeit $P(V^T)$, T ist die Länge der Beobachtungssequenz und n die Anzahl der hidden states.

$$P(V^T) = \sum P(V^T | \omega^T) P(\omega^T)$$

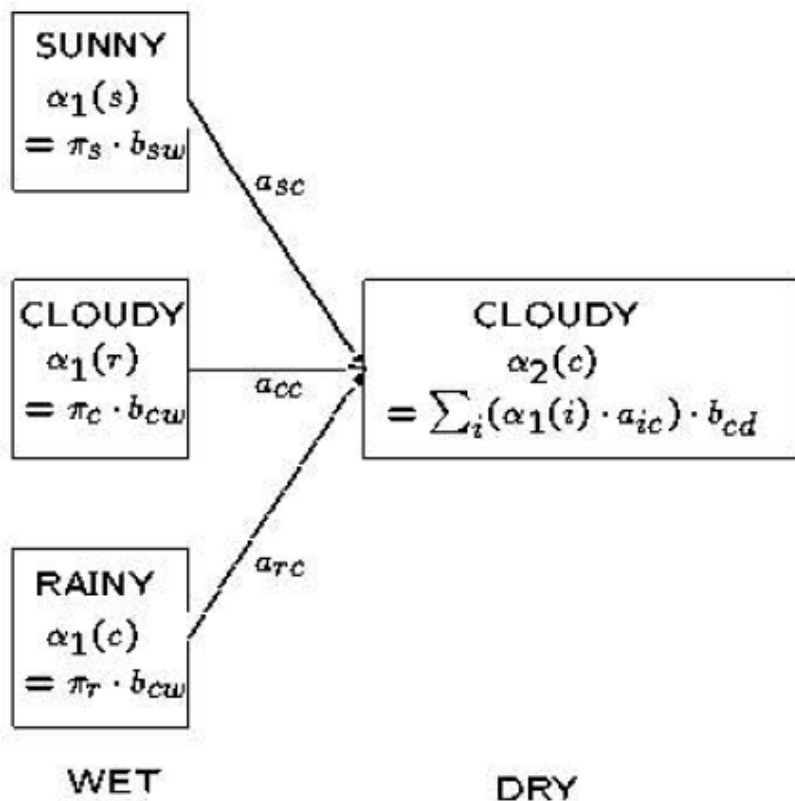
summiert über alle möglichen Sequenzen $\omega^T = \omega_1, \dots, \omega_T$.

$$P(V^T) = \sum_n \prod_{t=1}^T P(v_t | \omega_t) P(\omega_t | \omega_{t-1})$$

Zu beachten ist dass es n^T Zustandssequenzen der Länge T gibt. Es ist für praktische Anwendungen **viel** zu aufwändig alle auszuwerten.

Hidden Markov Models: Forward Algorithmus

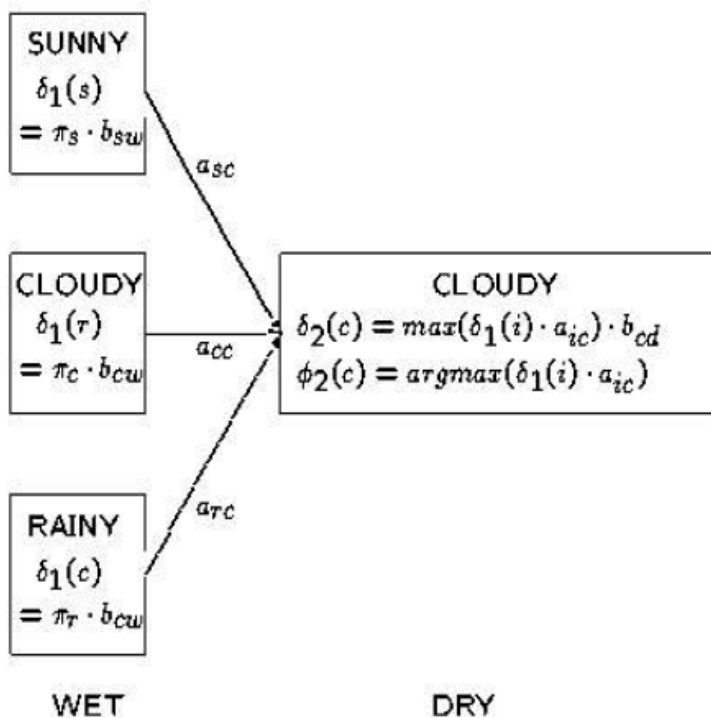
Dieser Algorithmus löst das Evaluations Problem rekursiv. Definiert wird die Wahrscheinlichkeit $\alpha_t(j)$ dass sich das System gerade im Zustand j befindet und bereits die ersten t Elemente von V^T produziert hat. Zur Initialisierung $t = 1$ gilt $\alpha_1(j) = \pi_j b_j(v_1(k))$ (P für Emission des ersten Zeichens im Zustand j).



Allgemein ist $\alpha_t(j) = \sum_{i=1}^n [\alpha_{t-1}(i) a_{ij}] b_j(v_t(k))$. Es wird für alle Zustände j und alle Zeiten t das $\alpha_t(j)$ berechnet (das sind nT , jedes α braucht n Operationen also insgesamt n^2T Operationen). Die Gesamtwahrscheinlichkeit ergibt sich aus $P(V^T) = \sum_{i=1}^n \alpha_T(i)$.

Hidden Markov Models: Viterbi Algorithmus

Auch für das Problem der Dekodierung könnten die Wahrscheinlichkeiten über alle Sequenzen bestimmt werden und dann die mit der grössten Wahrscheinlichkeit herausgesucht werden. Wie bereits gesehen ist das zu aufwändig. Die Lösung ist ein Verwenden des Forward Algorithmus im Sinne Dynamischer Programmierung mit Maximierung der Wahrscheinlichkeit, d.h. das Auffinden des zu jedem Zeitpunkt wahrscheinlichsten Zustands. Definiert wird die Viterbi Variable $\delta_t(i)$ als grösste Wahrscheinlichkeit V^t generiert zu haben wenn man sich im Zustand i zum Zeitpunkt t befindet.



Zur Initialisierung $t = 1$ gilt $\delta_1(j) = \pi_j b_j(v_1(k))$. Allgemein ist $\delta_t(j) = \max_{1 \leq i \leq n} [\delta_{t-1}(i) a_{ij}] b_j(v_t(k))$. Das heisst man sucht Schritt für Schritt den Pfad mit dem grössten P heraus, der bis zum Zustand j geht. Da aber die Abfolge der Zustände gesucht ist muss man den letzten Zustand in einem Pfad $\psi(j)$ speichern: $\psi_t(j) = \operatorname{argmax}_{1 \leq i \leq n} [\delta_{t-1}(i) a_{ij}]$. Am Ende wird der Maximalwert entlang des Pfades nach $t = 1$ zurückverfolgt.

Hidden Markov Models: Lernen/Trainieren

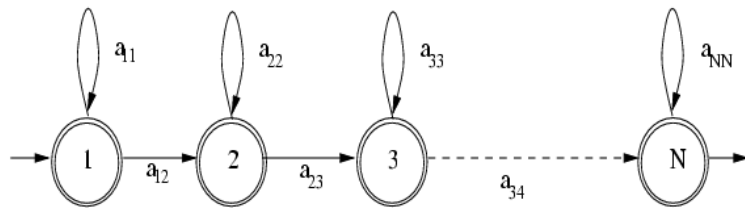
Ein klassisches Verfahren zur Berechnung der Parameter eines HMM ist die “Baum-Welch Rekursion”. Ein geeignetes Startmodell (A, B, Π) wird festgelegt (zufällig oder durch Vorwissen). Für dieses werden die Erzeugungswahrscheinlichkeiten der Trainingsdaten mit dem Forward Algorithmus berechnet. Iterativ werden unter Verwendung der Vorwärts- und Rückwärtswahrscheinlichkeiten (letztere analog aus dem Backward Algorithmus) A und B so angepasst, dass die Erzeugungswahrscheinlichkeiten ansteigen.

Probleme mit HMM:

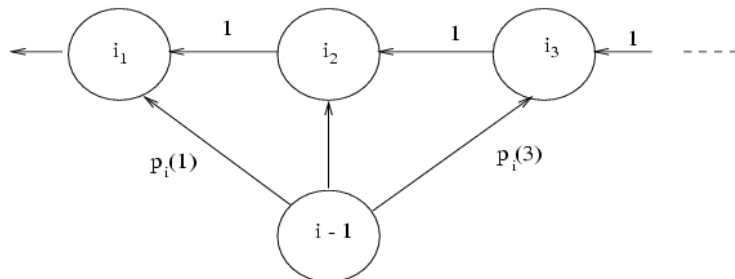
- Brauchen relativ viele Trainingsdaten (aufwändiges Enrollment)
- Anzahl der Zustände nicht mit Baum-Welch Rekursion ermittelbar
- Markov Annahme der Gedächtnislosigkeit oft zweifelhaft

Trotzdem viele erfolgreiche Anwendungen im Bereich Biometrie, v.a. bei Signatur Verifikation und Sprechererkennung (bei textabhängigen Verfahren).

Signaturen und HMM: Zeitschritte



Im klassischen Modell (mit konstanten Zeitschritten) wird ein längeres “Verharren” in einem State i durch a_{ii} modelliert, d.h. der Zustand bleibt während einiger Zeitschritte gleich. Eine elegantere Lösung stellen sog. HMM mit variablen Zeitschritten dar, die auch als HSMM (S für Semi) bezeichnet werden (können alle durch Schemata mit $a_{ii} = 0$ dargestellt werden).



Ein HSMM kann durch ein HMM mit einer grösseren Anzahl von States approximiert werden. Im einfachsten Fall wird ein HSMM State i in mehrere sub-States i_1, i_2, \dots mit konstanten Zeitschritten eingeteilt mit Übergangswahrscheinlichkeiten $a_{i-1, i_k} = p_{i-1}(k)$ und $a_{i_j, i_{j+1}} = 1$.

Typischerweise werden für die Modellierung von Signaturen eine eher kleine Anzahl von States verwendet, z.B. werden in der Literatur 6 States als gut angegeben. Jedoch muss die Anzahl für HMMs nicht notwendigerweise spezifiziert sondern kann im Training optimiert werden. Links-Rechts Topologie mit optionalen State skips ist klassisch.

Biorhythmen: Herzschlag/EKG

Grundsätzlich ist die Idee naheliegend Herztätigkeit als biometrisches Merkmal zu verwenden, da diese relativ leicht sensorisch zugänglich ist und durch die genetische Einmaligkeit des Organs und des Reizleitungsapparats eine hohe Inter-personal Variability zu erwarten ist. Darüberhinaus gibt es aus dem medizinischen Bereich grundsätzlich sehr viel Wissen über diesen Bereich.

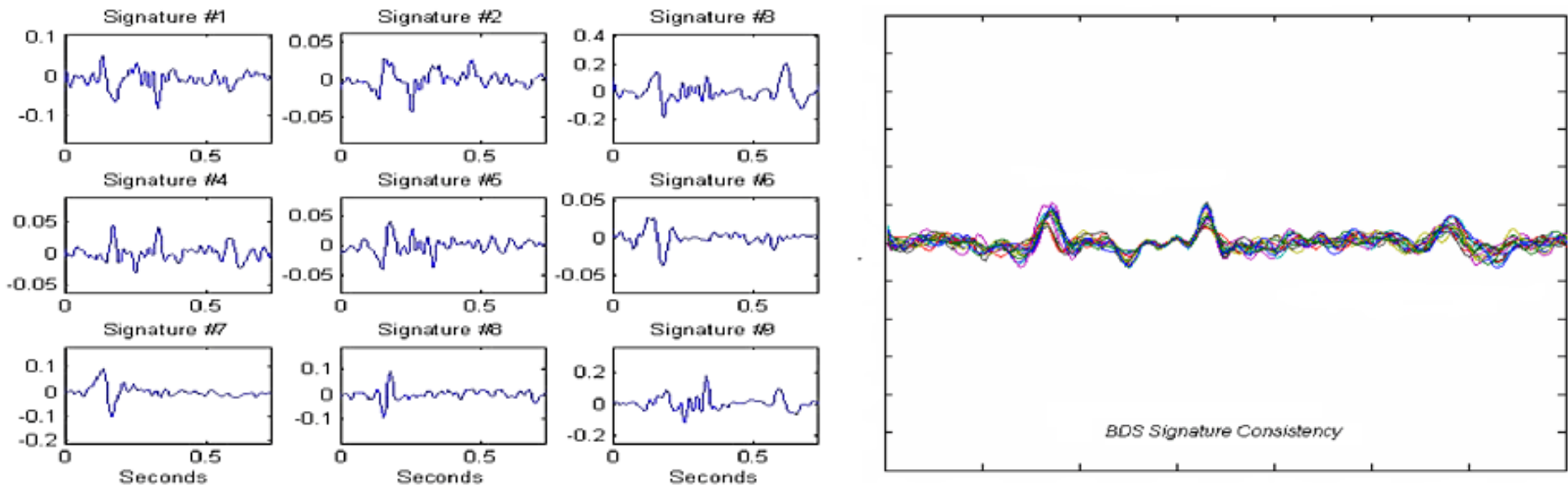
Problematisch erscheint jedoch, dass Messungen über die reine Pulsabnahme hinausgehend aufwändiger sind (EKG) und dass es eine grosse Menge an pathologischen und nicht pathologischen Veränderungen der Herztätigkeit gibt die die Intra-personal Variability beim falschen Merkmal stark erhöhen können (die in der Bevölkerung extrem häufig vorkommen) – wesentlich ist also **was** gemessen wird.

Beispiele: kurzfristige Herzrhythmusstörungen (Tachykardien, Extrasystolen (jeder hat im Durchschnitt 60 pro Tag), Kammer- und Vorhofflimmern und -flattern, etc.), langfristige Veränderungen der Herztätigkeit durch Erkrankungen (z.B. Herzinfarkt, Endocarditis, etc.), stressbedingte Effekte, u.s.w.

Welche EKG Features dieser Anforderung entsprechen könnten, ist unklar !

Biorhythmen: Herzschlag – IDESIA

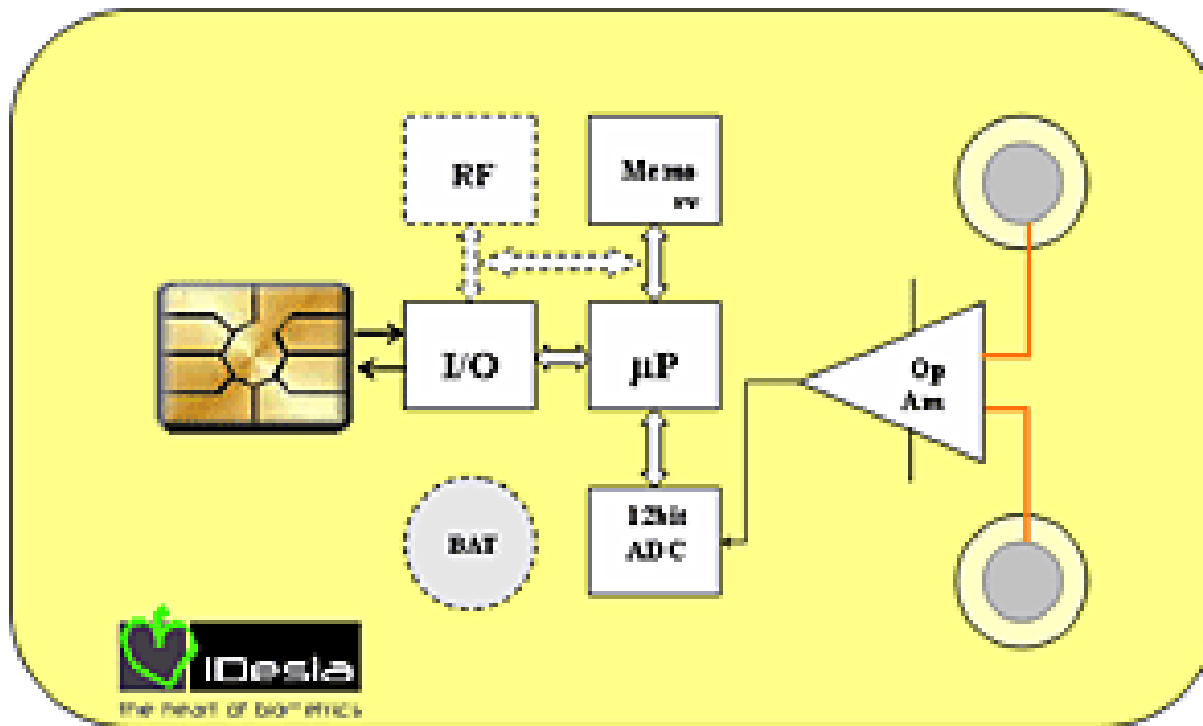
Die Firma IDESIA <http://www.idesia-biometrics.com/> bietet eine fertige Lösung im Bereich Herzschlag an, wobei nicht genauer spezifiziert wird was genau gemessen wird. Die vorliegenden statistischen Auswertungen beruhen auf 160 Personen und können nicht verifiziert werden (man kauft quasi die Katze im Sack). Im Bild die Kurven von 9 unterschiedlichen Testpersonen und die Überlagerung von 16 Messungen einer Testperson. Insbesondere Intra-personal variability ist völlig unklar, es werden hervorragende FMR, FNMR und EER behauptet (nona).



Biorhythmen: Herzschlag – IDESIA Produkte

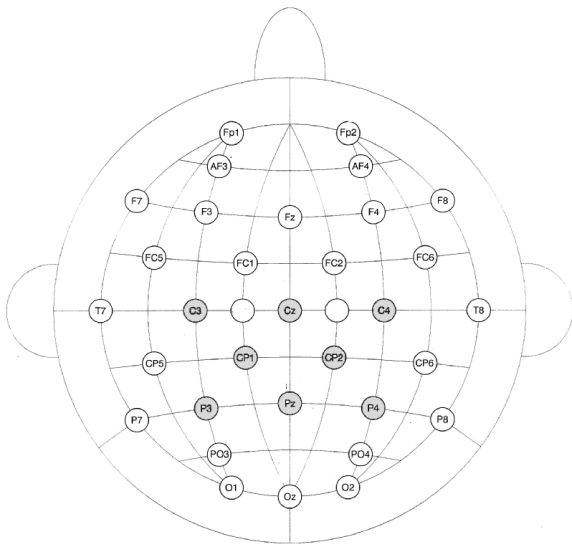


Die Produktpalette umfasst fertige Sensoren zum Anstecken an einen PC (mit je einem Finger einer Hand wird das Signal aufgenommen, Dauer 3-4 Sekunden), Sensor Bausätze zur Integration in ein anderes Gerät, entsprechende Smartcards und Sensoren zur Kombination mit Fingerprintererkennung ("Fusion").



Biorhythmen: EEG

Verglichen zum Herzschlag oder auch dem EKG ist das EEG ein wesentlich komplizierteres Signal, was durch die wesentlich komplexere Natur der Gehirnströme verglichen zur Herztätigkeit begründet liegt. Ein Problem ist die Sensorik: es ist schwierig, Signale mit guter Qualität zu erhalten, da im klassischen Szenario die Sensoren durch Schädeldecke, Haut, Haare von den relativ schwachen Signalemitotoren getrennt sind. Interessanterweise wurde in zwei Studien die Elektrode P4 als vielversprechend identifiziert (da diese bekannterweise den α -Rhythmus deutlich aufnimmt). Der erste Ansatz EEG Daten für biometrische Anwendungen einzusetzen ist den EKG-Verfahren nachempfunden. Die Idee ist, ein individuelles grundlegendes Gehirnwellenmuster zu verwenden. Problematisch ist bei diesem Ansatz die starke EEG Variabilität durch unbewusste und bewusste Vorgänge.



Von Testpersonen in Ruhe wurden sog. EEG Epochen von 8 Elektroden (F7+8, C3+4, FC1+2, P3+4) mit 8.5 sec. Dauer aufgenommen – solche mit Muskel- oder Herztätigkeit-basierten oder anderen Störungen wurden von Spezialisten ausgeschlossen wodurch ca. jeweils 8 Epochen pro Person zur Verfügung standen. Mit diesem Setup konnten 40+ Testpersonen mit guter Genauigkeit unterschieden werden.

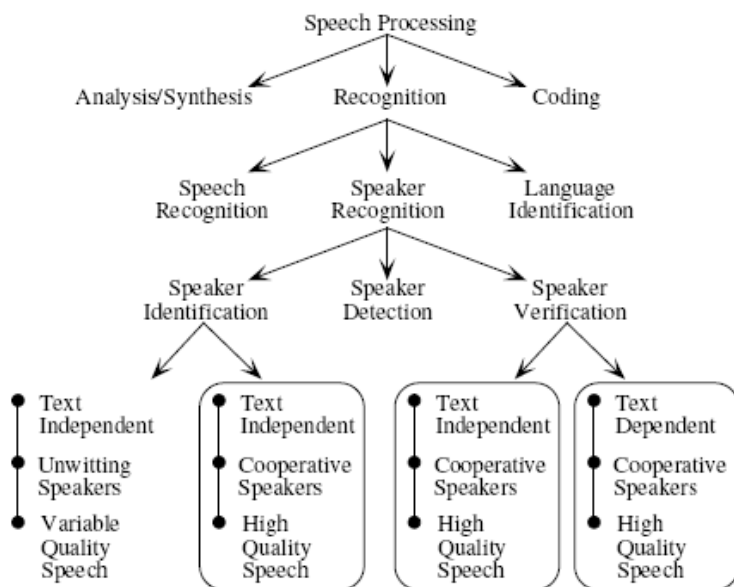
Biorhythmen: EEG/BCI Ansatz

Ein zweiter Ansatz nähert sich der Thematik durch den Zugang der Brain-Computer-Interface (BCI) Technologien. Im Bereich BCI kann man durch vorgestellte Aktivität erlernen, EEG Signale so zu steuern, dass binäre Information problemlos ausgelesen werden kann. Dadurch können vollständig gelähmte Patienten im klinischen Bereich ausschliesslich durch mentale Aktivität einfache Menüsteuerungen vornehmen oder auch kommunizieren. Dieser Ansatz wird auf das Biometrie Szenario übertragen. Im Gegensatz zum vorherigen Szenario wird eine spezifische mentale Aktivität ausgeführt – dies könnte als Entwicklungsendpunkt für eine Kombination aus wissens- und biometriebasiertem Authentifikationsverfahren verwendet werden: ein geheimes Passwort wird intensiv vorgestellt und das entsprechende EEG Signal zur Authentifikation benutzt. Das ist Zukunftsmusik.

Die Realität: Vorstellung von rechter & linker Handaktivität und Wörtern mit identischen Anfangsbuchstaben. Session mit 4 Minuten Dauer, jeweils 15 Sekunden ein mentaler Task, zufälliges Wechseln, 8 Elektroden wie im vorigen Bild. Laplacian Filterung der Rohdaten (Glättung plus 2. Ableitung), 3 0.5 Sekunden lange Segmente mit 50% Überlappung werden einer FFT unterzogen und die Ergebnisse gemittelt. Mit 3 Testpersonen werden Ergebnisse vergleichbar mit denen von Gesichts- und Sprechererkennung erzielt. Völlig unklar bleibt die mögliche Ausweitung auf grössere Testszenarien.

Sprachverarbeitung und Sprache als Biometrisches Merkmal

Sprache ist ein natürliches Signal, sprechen zu müssen wird nicht als besondere Anforderung empfunden, sozusagen habituated. Darüberhinaus gibt es mit diversen Telephonnetzen und VoIP eine existierende, weitestverbreitete und verteilte Infrastruktur. Wesentliches Kriterium ist ob es sich um textunabhängige (freies Sprechen) oder textabhängige (user-spezifisches Passwort oder vom System vorgegebenes Kennwort) Systeme handelt.



Textunabhängige Verfahren sind wesentlich schwieriger zu realisieren, können sich aber auch auf inhaltliche Eigenheiten des Sprechers stützen (bestimmte Worte besonders häufig, Floskeln, Sprechpausen, u.s.w.). Solche Verfahren setzen nach der Featureextraktion beispielsweise anstelle von HMMs Gaussian Mixture Models (GMMs) ein – auf diese Thematik wird nicht näher eingegangen. Je mehr Spracherkennung verbessert wird und je mehr Sprach- und Sprechererkennung zusammenwachsen, desto unwichtiger wird die Unterscheidung zwischen textunabhängigen und textabhängigen Systemen.

Textabhängige Sprecher Erkennung: Feature Extraction

Diese Technologie ist für biometrische Anwendungen gegenwärtig am geeignetsten. Im Bereich der Feature Extraction werden drei Stufen unterschieden:

1. Erkennung von Sprachaktivität: Trennung von Signalsegmenten mit und ohne Sprachaktivität.
2. Feature Extraction: Es ist inzwischen gut belegt, dass ein sprachliches Signal wesentlich von den physiologischen Gegebenheiten des Sprechers geprägt wird. Das Signal wird in sog. temporale "Frames" zerlegt (mit braven Fensterfunktionen: sog. Hamming windows) die ca. 20 ms lang sind und sich jeweils um 10 ms überlappen. Zwei unterschiedliche Strategien werden anschliessend angewendet: LPC & MFCC
3. Kanal Kompensierung: Verschiedene Eingabegeräte und Übertragungsmedien verursachen unterschiedliche Effekte im aufgenommenen Audiosignal. Da eine Verifikation vorzugsweise unabhängig von diesen Umgebungsparametern funktionieren soll, müssen die Features dementsprechend nachbehandelt werden.

Als Ergebnis der Feature Extraktion liegen in jedem Fall Zeitreihen von Features vor, die miteinander verglichen werden müssen. Konzeptuell ist also textabhängige Sprechererkennung der on-line Signatur Erkennung sehr ähnlich, entsprechend werden die selben Verfahren eingesetzt (z.B. HMMs, DTW,

Feature Extraction: LPC I

LPC (Linear Prediction Coefficients) Analysis basiert auf einem linearen Modell der Spracherzeugung. Der physiologische Spracherzeugungapparat wird als Kombination von vier Modulen betrachtet: die Stimmritzen (die im Fall von stimmhaften Lauten einen Strom von Impulsen erzeugen und im Fall von stimmlosen Lauten weisses Rauschen), Rachen- und Mundraum, Nasenraum und die Lippen. Jede dieser Komponenten kann durch einen bestimmten Filter repräsentiert werden: ein Tiefpassfilter für die Stimmritzen, ein AR (autoregressive) Filter für den Vokaltrakt, ein ARMA (autoregressive moving average) Filter für den Nasenraum und ein MA Filter für die Lippen. Der gesamte Spracherzeugungapparat kann durch einen ARMA Filter dargestellt werden.

Die Charakterisierung eines Sprachsignals wird erreicht durch die Bestimmung der Koeffizienten eines globalen Filtermodells das das Sprachsignal beschreibt. Meistens wird für die konkrete Umsetzung das ARMA Filtermodell zu einem AR Filter vereinfacht. Das Prinzip der LPC Analysis beruht also auf einer Schätzung der Parameter eines AR Filters der einen Ausschnitt (i.e. Frame) des Sprachsignals gut beschreibt. In jedem Frame werden optimale Parameter berechnet die zur Berechnung eines Featurevektors verwendet werden.

Bemerkung: diese Idee ist eng verwandt mit DPCM Kompression im Bild und Audiodbereich und im Bereich Sprachkodierung im CELP-coder (z.B. GSM Kompression) realisiert.

Feature Extraction: LPC II

Sei gegeben ein Sprachsignalsegment $S = (s(0), s(1), s(2), \dots, s(N))$. Seien weiters $s(n)$ die beobachtete Ausgabe zum Zeitpunkt n , p die Ordnung der Vorhersage (des Prädiktors), a_k die sog. Prädiktorkoeffizienten (PCs), u_n die gegenwärtige Eingabe und G der sog. Gain-Faktor. u_n entspricht dem Anregungssignal und G ist die Luftmenge die zu unterschiedlicher Lautstärke führt.

$$s(n) = \sum_{k=1}^p a_k s(n-k) + Gu(n)$$

In tatsächlichen Spracherkennungsapplikationen sind u_n und G natürlich unbekannt und werden daher ignoriert. Die LP Approximation $\hat{s}(n)$ wird ermittelt durch

$$\hat{s}(n) = \sum_{k=1}^p a_k s(n-k) \text{ mit } e_n = s(n) - \hat{s}(n)$$

und beinhaltet nur mehr gegenwärtige und vergangene Samplewerte.

Feature Extraction: LPC III

Die a_k sollen so bestimmt werden, dass der MSE E über den gesamten Frame minimal wird:

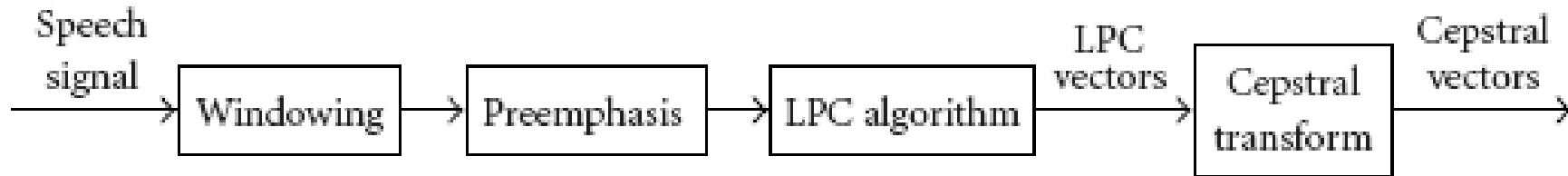
$$E = \sum_{i=1}^N e_i^2 = \sum_{i=1}^N [s(n) - \sum_{k=1}^p a_k s(n-k)]^2$$

Das Minimierungskriterium $\frac{\delta E}{\delta a_k} = 0 \quad \forall k = 1, 2, \dots, p$ führt zu folgender Bedingung:

$$\sum_{k=1}^p a_k \sum_n s(n-k)s(n-i) = - \sum_n s(n)s(n-i) \quad \forall i = 1, 2, \dots, p$$

Dies führt zur Autokorrelationsmethode zur Ermittlung der a_k , ebenso wird Kovarianz und ähnliche Ausdrücke verwendet.

Feature Extraction: LPC IV



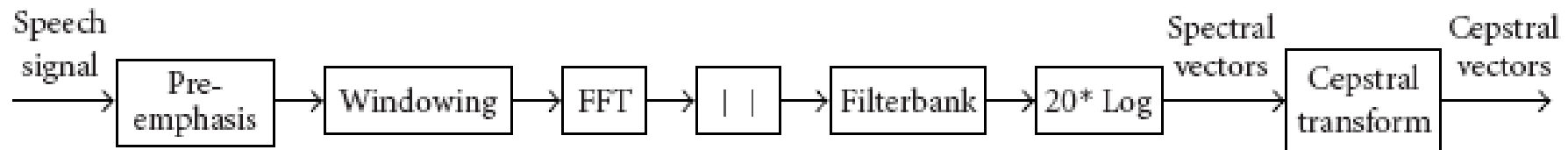
- “Pre-emphasising” soll die hohen Frequenzen im Sprachsignal verstärken die durch die Sprachgenerierung unterdrückt werden: $s_p(n) = s(n) - as(n - 1)$ mit a aus dem Intervall $[0.95, 0.98]$. Ob diese oft verwendete Vorverarbeitung Sinn macht, ist nicht völlig klar (probieren!).
- Aus den durch LPC Analyse pro Frame gewonnenen Vektoren a_k werden Featurevektoren berechnet, z.B. die Linear Predictive Cepstral Coefficients (LPCC, mit $c_1 = a_1$ und $K \leq P$ die Anzahl der gewünschten Koeffizienten):

$$c_n = a_n + \sum_{k=1}^{n-1} (1 - k/n) a_k c_{n-k}, \quad n = 1, \dots, K$$

- Von den LPCC wird oft ihr Mittelwert abgezogen (cepstral mean subtraction CMS – Beitrag von Hintergrundrauschen wird abgezogen), oft wird auch die Varianz auf eins normiert (reduction).
- Zusätzlich wird auch dynamische Information verwendet, in wie weit sich die gewonnenen Vektoren in der Zeit verändern (sog. delta cepstra, die durch Approximation der ersten und zweiten Ableitung berechnet werden)

Feature Extraction: MFCC

Mel Frequency Cepstral Coefficients sind das klassische Ergebnis einer Short Term Fourier Transform (STFT) Signal Analyse mit zusätzlicher psychoakustischer Modellierung.

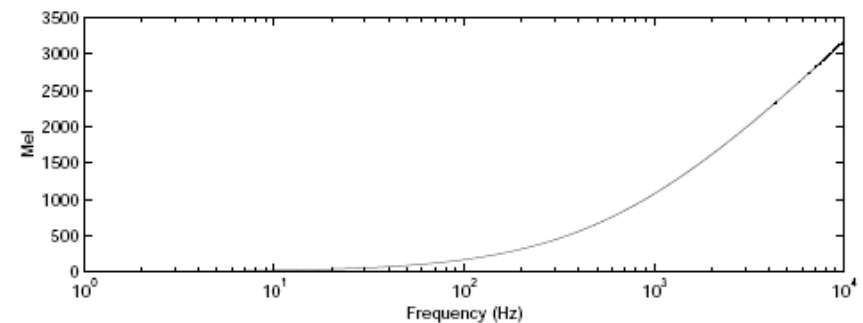
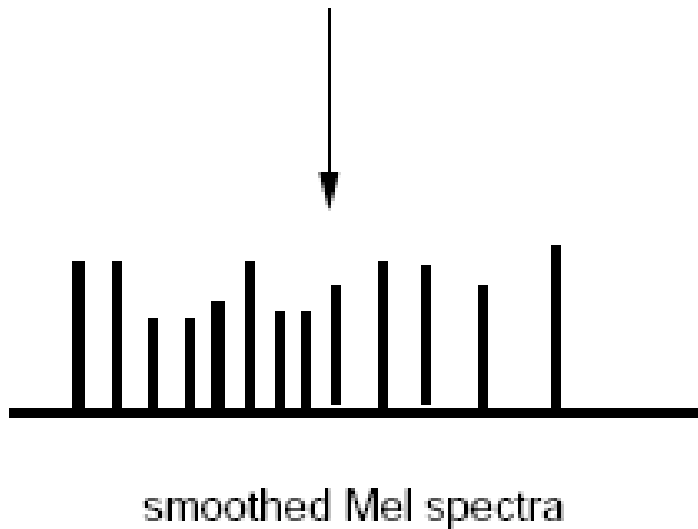
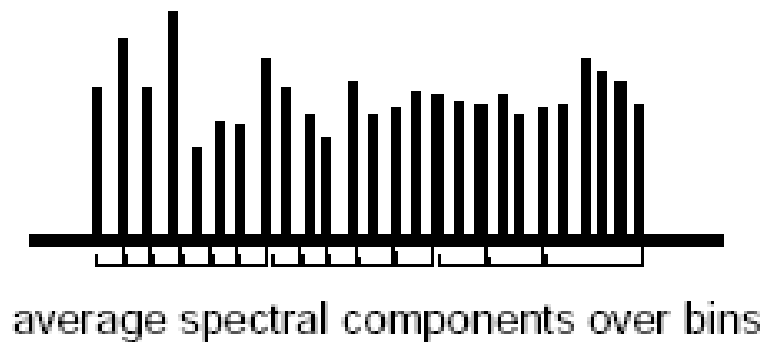


Nach der gefensterten Fourier Transformation werden die Spektralkomponenten nach einem psychoakustischen Modell einer non-uniform Quantisierung unterzogen. Hier wird die sog. "Mel" Frequenzskala verwendet, die die niederen Frequenzen eher bevorzugt (siehe nächste Folie) und eine teilweise Logarithmierung des Frequenzraums beinhaltet. Im Folgenden gibt es dann zwei Varianten der Konstruktion von Featurevektoren:

1. Die Mel-Frequenzvektoren sind hochkorreliert. Sie werden daher meist einer DCT unterzogen, von der wiederum nur die DC und niederen AC Koeffizienten behalten werden. Dieser Vorgang reduziert die z.B. ursprünglich 256 Spektralkomponenten zu z.B. 40 Mel-Spektalwerten und zu schlussendlich ca. 13 cepstral features pro Frame.
2. Die Mel-Frequenzvektoren werden einer inversen FFT unterzogen, das Ergebnis sind die sog. real cepstral coefficients (RCC).

MFCC – der Mel Frequenzraum

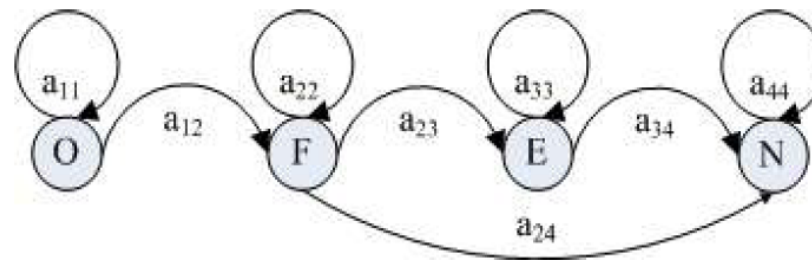
Die Mel Skala basiert auf einer Abbildung zwischen der physikalisch messbaren Frequenz und der analogen psychoakustischen Einheit, dem Pitch – Pitch (Einheit: ein phon, ein sone) wird nicht in linearer Abhängigkeit von Frequenz wahrgenommen und ist auch von der Signalintensität abhängig. Analoge Einheit für Lautstärke – physikalisch SPL sound pressure level in DB – ist die Loudness. Die Abbildung zwischen Frequenz und Pitch ist linear unter 1 kHz und darüber logarithmisch.



Sprechererkennung: Matching

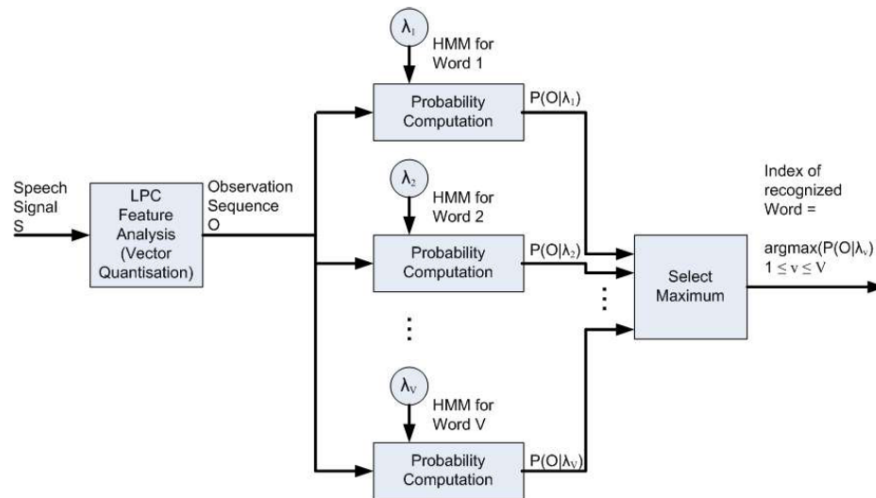
1. Template Modelle

- DTW: klar.
 - VQ: aus den Enrollment Daten wird durch klassische Verfahren aus der Vektorquantisierung VQ ein Codebook generiert das die Daten mit geringem Fehler approximiert. Der matching Score ist dann der minimale Abstand eines zu verifizierenden Frames zu einem Codebookeintrag.
 - Nearest Neighbour (NN): Es werden alle Abstände zwischen den Enrollment Frames und den zu verifizierenden Frames gemessen (z.B. auch durch DTW). Die NN Distanz eines Frames ist die kumulierte Distanz zu seinen k NN. Die NN Distanzen für alle zu verifizierenden Frames werden gemittelt und ergeben den Score.
2. Stochastische Modelle (HMM): Die hidden states entsprechen den Lauten eines Wortes, die Emissionssequenz entspricht den Featurevektoren der zugehörigen Frames. Wieder werden Links-Rechts Topologien verwendet.



Sprechererkennung: HMM

Durch enrollment stehen k Beobachtungssequenzen zur Verfügung, diese werden benutzt um mit der Baum-Welch Rekursion für jeden Sprecher ein HMM zu trainieren – entweder für jeden Sprecher das gleiche Wort oder ein Passwort (also verschiedenes Wort).



Bei der Verifikation/Identifikation wird dann mit dem Forward/Veterbi Algorithmus evaluiert, mit welcher Wahrscheinlichkeit die beobachtete Sequenz von dem behaupteten Sprecher stammt bzw. welcher Sprecher die höchste Wahrscheinlichkeit der Generierung der beobachteten Sequenz hat (wie in der Graphik dargestellt – wird nicht ein Passwort verwendet sondern eine identische Phrase für alle Sprecher ist Word1 durch Speaker1 zu ersetzen). Bei der Passwortphrasen Technik wird Spracherkennung und Sprechererkennung vermischt.

“Unusual” Biometrics: Geruch

Bei der Verwendung von “persönlichem” Geruch treten verschiedene Probleme auf, obwohl klar belegt ist dass Körpergeruch eine tatsächlich sehr individuelle Ausprägung hat und daher prinzipiell für biometrische Applikationen benutzbar ist:

- Intra-personal Variability: der Körpergeruch verändert sich signifikant durch verschiedene physiologische und pathologische Prozesse im Körper: physiologisch – z.B. Stress, hormonelle Umstellungen, Nahrungsaufnahme; pathologisch – z.B. Zuckerkrankheit, Erkrankungen des Magen- und Darmtrakts, Zahn- und Zahnfleischerkrankungen, Wundinfektion.
- Sensorproblem: im menschlichen Geruchsorgan stehen ca. 10000 Sensoren zur Verfügung, die Verarbeitung der eingehenden Signale wird von ca. 10 Millionen Sensorneuronen übernommen. Dieses Signalverarbeitungssystem analysiert dann Gerüche wie z.B. Kaffee, der aus ca. 670 Chemikalien besteht. Hier herrscht eine grosse Diskrepanz zu den Möglichkeiten der sog. ENoses, die 12-30 Sensoren für verschiedene Substanzen und entsprechend einfache Verarbeitungsmethoden bereitstellen.

Interessenten im Bereich Biometrie trotzdem vorhanden: Mastiff Electronic Systems und das Pentagon. Realistischere Applikationen im Bereich Opfersuche nach Unglücksfällen mit Verschütteten (zukünftig), gegenwärtig im Bereich medizinischer Diagnostik, Umweltüberwachung (Emissionskontrolle), Nahrungsmittelindustrie, Parfumherstellung und -kontrolle.

Table of Contents: Biometrie mit BVA Methoden

- Offline Unterschriftserkennung
- Iris Recognition,
- Retina Scan,
- Fingerprints,
- Palmprint, Hand- und Fingergeometrie,
- Face Recognition,
- Gait Recognition.

Off-Line Signaturen: Grundlagen

Vorteile: Es wird keine besondere Sensorik benötigt, da hier die klassische analoge Unterschrift einfach eingescannt werden kann. Ein Mindestmass an Scannerqualität ist aber erforderlich. Weiters wird die Akzeptanz sehr hoch sein, da für die meisten Menschen Unterschriftsleistung ein alltägliche Vorgang ist (habituated environment).

Nachteile: Die zeitliche Dynamik geht natürlich verloren und damit auch ein wichtiger Aspekt: die Fälschungssicherheit. Jedoch ist wegen der Erhöhung der intra-personal Variability deren Bedeutung ohnehin umstritten.

Viele der als “on-line” angepriesenen Features können durch Verwendung eines zur Zeit alternativen Parameters der Signatur (z.B. Bogenlänge des Unterschriftszugs) aus Off-line Signaturen gewonnen werden. Ausnahmen sind natürlich Druck, Stiftneigungen, Geschwindigkeit, Beschleunigung, Gesamtzeit, u.s.w.

Off-Line Signaturen: Normierung

Normierung ist wichtig für einen entsprechenden Vergleich verschieden grosser und unterschiedlich positionierter Signaturen. Je nach Art der verwendeten Features werden gänzlich verschiedene Strategien verwendet. Optimalerweise wird durch die Sensorik bereits eine Normierung durchgeführt (z.B. vorgegebene Zeile, beschränktes Schreibfeld).

- Daten in zeitreihenanaloger Form $X(t)$
 - ★ Durch DTW (wobei das dann nicht wirklich T ist) ohne fix definierten Anfangs- und Endpunkt kann ein optimales Alignment bestimmt werden und dann können zwei Unterschriften mit gleicher Sampleanzahl (falls benötigt) durch entsprechende Interpolation generiert werden.
 - ★ Einfacher ist eine Normierung auf einheitliche Bogenlänge (die allerdings Anfangs- und Endartefakte ausser Acht lässt)
- Daten als planare Kurve: Koordinaten (x_i, y_i)
 - ★ Fourier Methoden
 - ★ Spatial Domain Methoden

Off-Line Signaturen: Fourier Normierung für planare Kurven

Sei gegeben eine planare Kurve $\vec{z} = (z_1, z_2, \dots, z_N) = ((x_1, y_1), (x_2, y_2), \dots, (x_N, y_N))$. Die Fourier Transformation einer solchen Kurve ist gegeben durch

$$\vec{Z} = F\vec{z}$$

mit F als Fouriermatrix mit den Einträgen $F_{jk} = \omega^{jk}$ und $\omega = e^{2\pi i/N}$.

Die Folge der Fourierkoeffizienten $\vec{Z} = Z_0, Z_1, \dots, Z_{N-1}$ kann als eine alternative Darstellung genutzt werden (die, ebenso wie in der Bildverarbeitung Korrelationen entfernt und zu kompakterer Darstellung führt). In dieser Darstellung kann eine Normierung durchgeführt werden, die auf der Normierung von Fourierkoeffizienten beruht. Als erster Schritt wird $Z_0 = 0$ gesetzt, was einer Translation des Koordinatensystems in den Zentroiden der Kurve entspricht: $\vec{z} - Z_0 = \vec{z} - 1/N \sum_{k=1}^N z_k$. Dieser Vorgang entspricht in der Bildverarbeitung dem Abziehen des mittleren Grauwerts vom ganzen Bild (kommt z.B. im Bereich Iriserkennung vor, siehe weiter hinten).

Im zweiten Schritt wird der folgende Fourierkoeffizient Z_1 auf den Wert 1 normiert - das wird durch eine Division aller Z_i durch Z_1 erreicht (was durch die Linearität der Fouriertransformation einer Division von $\vec{z} - 1/N \sum_{k=1}^N z_k$ entspricht). Das entspricht geometrisch einer Skalierung und Rotation der ganzen Signatur.

Off-Line Signaturen: Positions-Normierung für planare Kurven

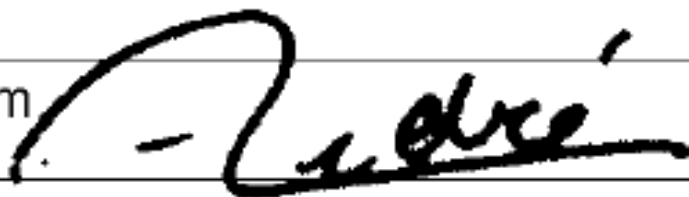
Zur **Positionierung** gibt es z.B. den Vorschlag alle Signaturen nach ihrem Schwerpunkt (siehe bei Features) auszurichten und dann optimale Rotation durch Pattern Matching zu bestimmen.

Eine andere Variante bestimmt bei der Unterschrift nach graphologischem Vorbild den Zentralteil und die unteren bzw. oberen Regionen. Die Trennungsgerade zwischen Zentralteil und unterer Region kann als x-Koordinatenachse verwendet werden, normal dazu durch den Punkt mit kleinster x-Koordinate die y-Koordinatenachse.

upper zone

medium
zone

lower zone



The signature 'Rudri' is written on a three-line grid. The 'u' and 'd' are in the upper zone, the 'r' is in the medium zone, and the 'i' is in the lower zone. The baseline of the signature is aligned with the bottom line of the grid.



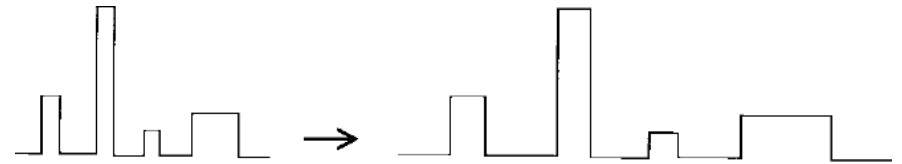
The signature 'Cardosa' is written on a three-line grid. The 'C' and 'a' are in the upper zone, 'r' and 'd' are in the medium zone, and 'o' and 's' are in the lower zone. The baseline of the signature is aligned with the bottom line of the grid.

Off-Line Signaturen: Skalierungs-Normierung für planare Kurven

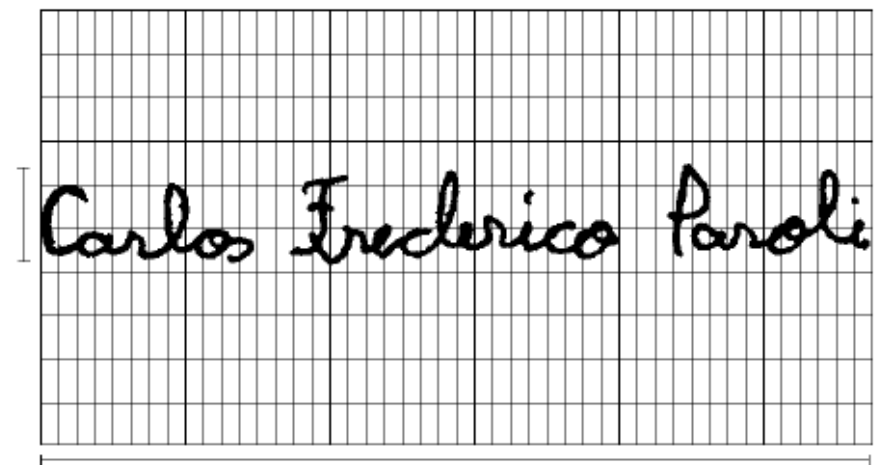
Zur **Skalierung** gibt es zwei unterschiedliche Strategien, je nachdem welche Variationsmöglichkeiten für eine authentische Unterschrift angenommen werden.

Die erste Strategie normiert den sog. Aspekt, das Verhältnis von Höhe zu Breite eines umschriebenen Rechtecks (tatsächlich wird der Bruch aus Summe aller vertikalen Displacements durch die Summe aller horizontalen Displacements gebildet und auf einen einheitlichen Wert gebracht).

Das Argument ist, dass man die Unterschrift wesentlich länger machen kann ohne die Höhe zu verändern (die Frage ist ob hier nicht wesentliche Eigenschaften verloren gehen - die Form des umschriebenen Rechtecks ist jedenfalls auch ein – sogar verwendetes – globales Feature).



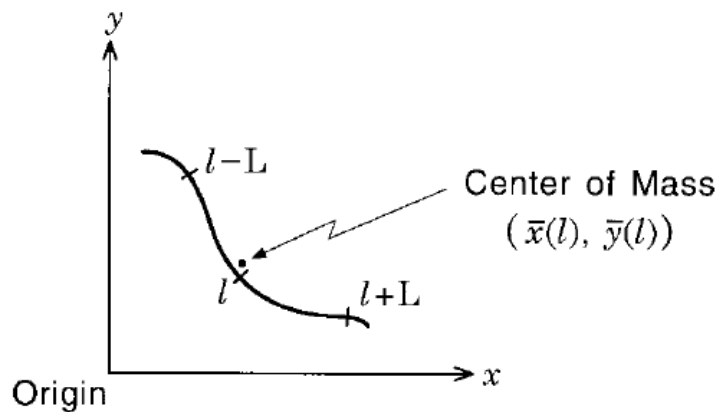
Genau dieser Ansatz wird bei der zweiten Variante verfolgt, der nur eine Längennormierung vornimmt. Hier ist dann schon die Anzahl der frei bleibenden Kästchen ein verwendbares Feature.



Off-Line Signaturen: “Zeitreihen” Features I

Die im folgenden beschriebenen Features werden oft auch im Kontext mit On-line Methoden verwendet, hier als Features einer Folge parametrisiert nach der Bogenlänge der Signatur.

Seien $x(l)$ und $y(l)$ die Koordinaten parametrisiert nach l und $g(\lambda)$ ein Gewichtungsfester (typischerweise Gaussfunktion) der Weite $+ - L$.

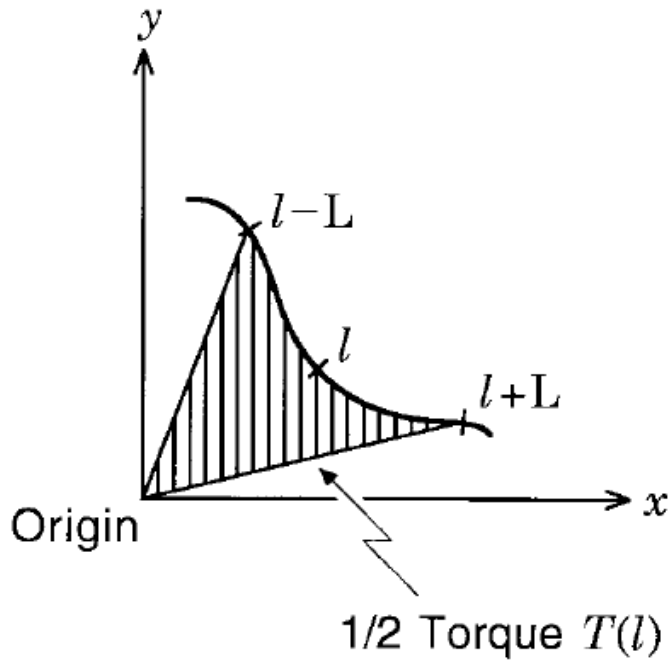


Die Koordinaten des **Schwerpunkts** $(X(l), Y(l))$ sind gegeben durch $(Y(l)$ analog):

$$X(l) = \int_{-L}^L g(\lambda)x(l + \lambda)d\lambda$$

Die Folge der Schwerpunktkoordinaten ist wesentlich robuster als die Originalkoordinaten.

Off-Line Signaturen: "Zeitreihen" Features II

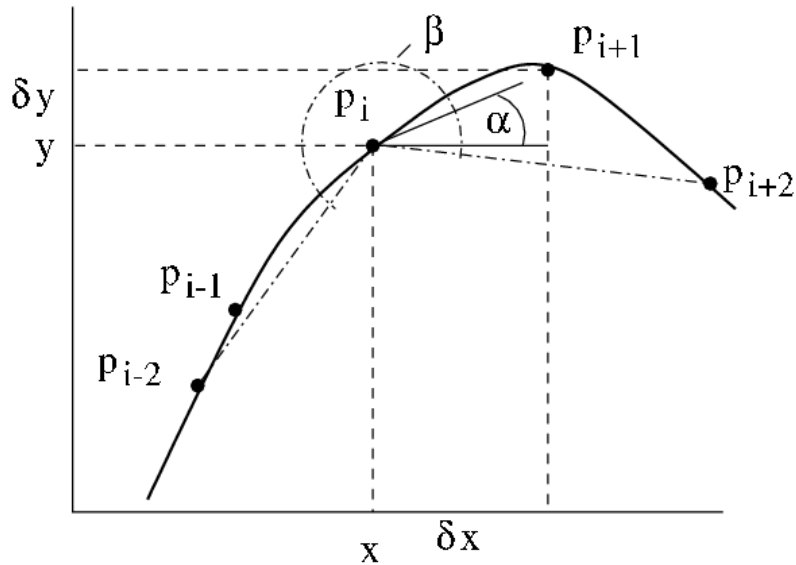


Die **Drillung/Drehmoment - Torque** ist gegeben durch:

$$T(l) = \int_{-L}^L g(\lambda) (y(l+\lambda) dx(l+\lambda) - x(l+\lambda) dy(l+\lambda)) d\lambda$$

mit $dx(l+\lambda)$ die Veränderung in x-Richtung an der Stelle $(l+\lambda)$ bei Veränderung von λ . Hat physikalische Bedeutung! (positives $T(l)$ ist Rotation entgegen dem Uhrzeigersinn).

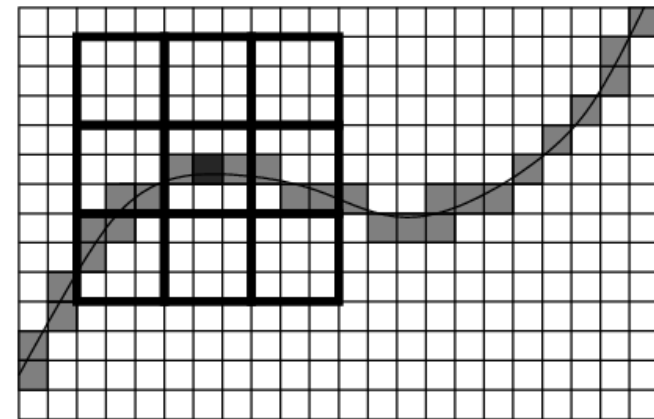
Off-Line Signaturen: "Zeitreihen" Features III



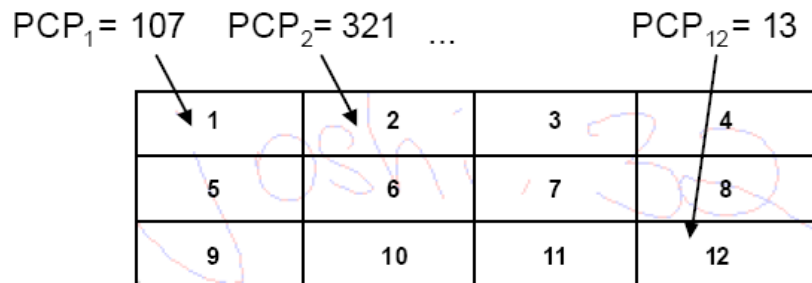
Typische weitere Features sind die Krümmung $\beta(i)$ als Winkel zwischen den Geraden $p_{i-2}p_i$ und $p_i p_{i+2}$, sowie der Winkel $\alpha(i)$ (zwischen der x-Achse und der Geraden $p_i p_{i+1}$) sowie dem Differenzwinkel $\delta\alpha(i) = \alpha(i) - \alpha(i-1)$.

Die Folge der Tangentenwinkel an p_i wird ebenso betrachtet – eine Variante ist das Anwenden einer DFT auf einen Vektor der Länge 10 dieser Reihe.

Als letztes Zeitreihen Feature wird noch das "sliding Bitmap Window" erwähnt, das in jedem Punkt ein z.B. 9×9 Pixel grosses Fenster betrachtet das in $9 \times 3 \times 3$ Pixel grosse Blöcke unterteilt wird in denen die Pixel gezählt werden. dies gibt einen 9-elementigen Vektor pro Signaturpunkt.



Off-Line Signaturen: Pixel Feature



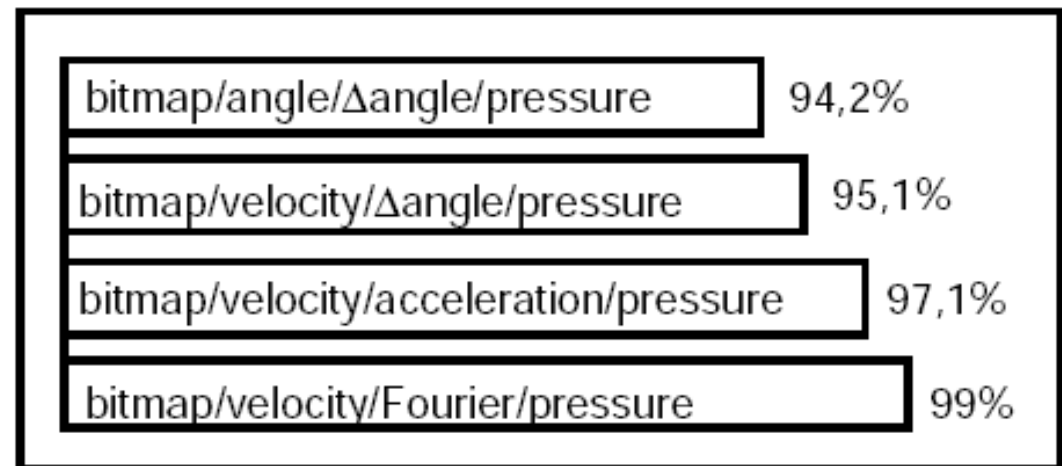
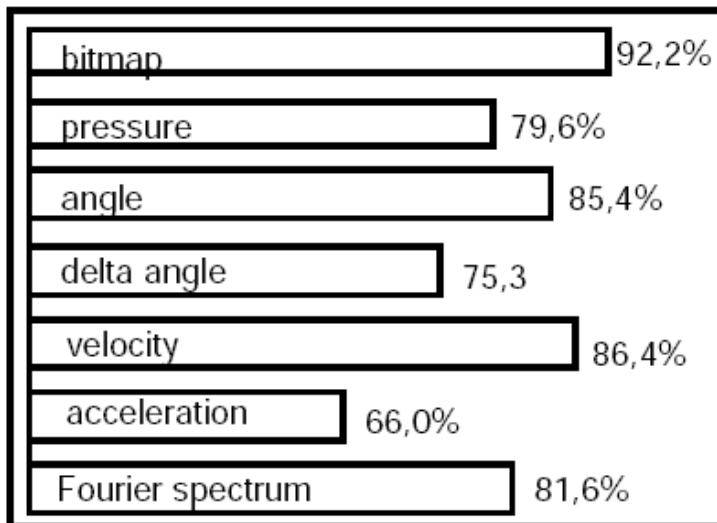
Das "Bitmap" Feature ist eines der einfachsten und braucht keine parametrisierte Kurvendarstellung: die Signatur wird von einem minimalen Rechteck umschrieben das skaliert wird und in eine fixe Anzahl von Quadraten aufgeteilt wird. In diesen Quadraten werden dann die Pixel gezählt, was den sog. Pixel Count Parameter (PCP) ergibt.

Besonders in Regionen über und unter dem Hauptkörper der Unterschrift zeigt sich dass eine feiner Unterteilung oft Vorteile bringt.

Hier können dann Multiresolution Methoden eingesetzt werden, z.B. mit einer Quadtree Struktur.

Off-line vs. On-line Signature Verification

Die Graphiken zeigen einen experimentellen Vergleich bzgl. Korrektheit (in Prozent, Anzahl der korrekterweise akzeptierten Originalunterschriften und Anzahl der korrekterweise zurückgewiesenen Fälschungen).



Interessanterweise sind die Off-line Features zumindest ebenbürtig, hohe Korrektheitsraten werden nur mit Kombinationen erzielt.

Signature Verification Contest SVC 2004

SVC ist ein offener alle zwei Jahre stattfindender Wettbewerb unter genormten Bedingungen. Es werden Trainingsdaten zur Verfügung gestellt, alle eingereichten Verfahren werden dann anhand eines Standardtestsets evaluiert und ihre Leistungsparameter ermittelt.

Es gibt zwei Tasks: einer mit zeitabhängigen Koordinatendaten (eher off-line) und einer mit zusätzlichen Druck und Stiftorientierungsdaten (on- und offline). Die Datenbanken beinhalten eine grosse Menge an skilled forgeries, und es wurden nicht die echten Unterschriften der Probanden verwendet (aus Privacy Gründen – obwohl trainiert wurde, ist das nicht als habituated einzustufen – eine echte Designschwäche !) Es blieben 15 funktionsfähige Verfahren für Task 1 und 12 Verfahren für Task 2 übrig – alles akademische Gruppen, die industriellen Teilnehmer zogen trotz der Möglichkeit zur Anonymisierung ihre Teilnahme zurück.

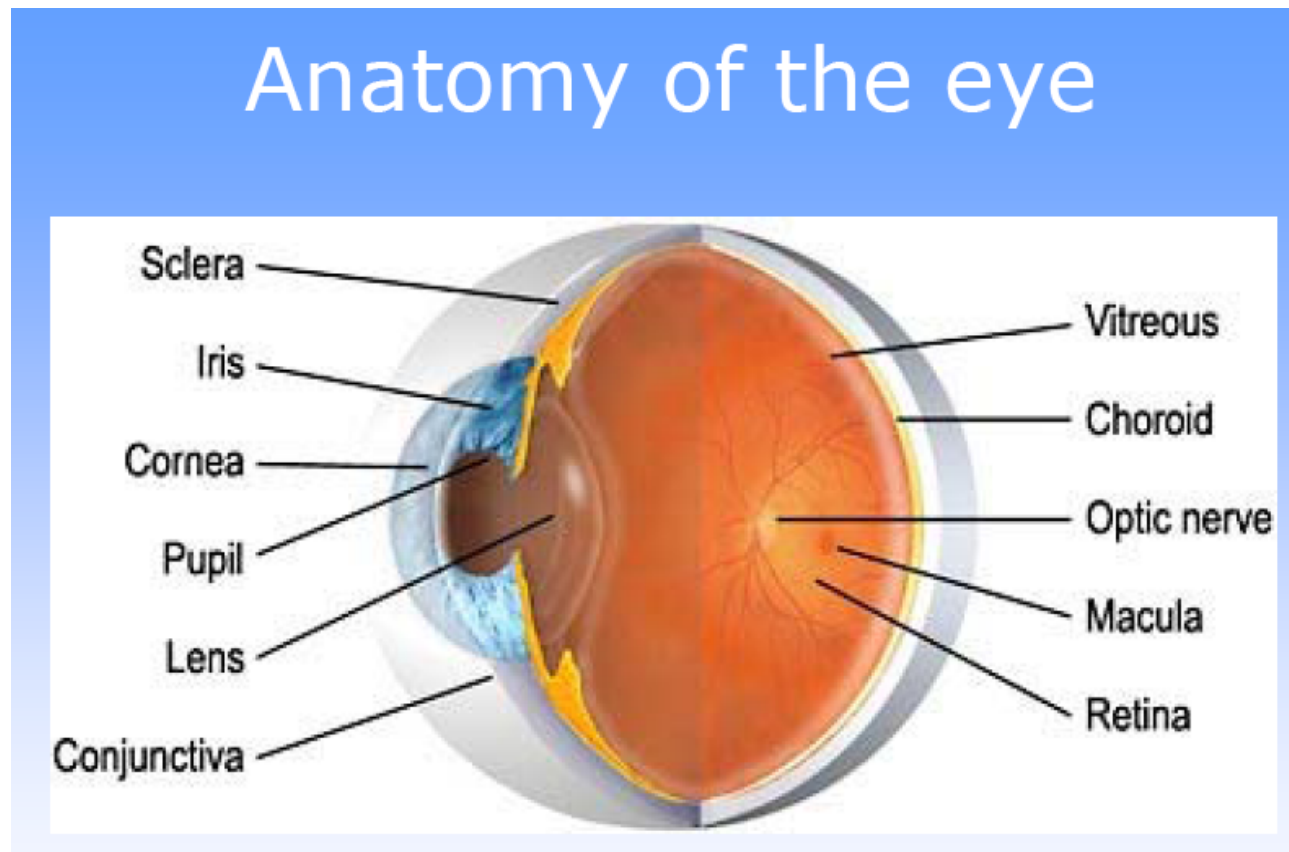
Es ergaben sich sehr grosse Leistungsunterschiede zwischen den Einreichungen, das Team der Sabanci Universität (Türkei) zeigte die besten Ergebnisse für beide Tasks. Interessanterweise waren die Ergebnisse für Task 1 besser als die von Task 2, was wiederum zweifel an der Sinnhaftigkeit der (aufwändigeren) On-line features aufkommen lässt. Diese Ergebnisse werden den “untypischen” Signaturen zugeschrieben.

Signaturen: Produkte

- <http://www.cybersign.com/> On-line und Off-line Features, vermischt mit Digitalen Unterschriften
- <http://www.onclickbiometrics.com> (→ Products/Business) On-line und Off-line Features, vermischt mit Digitalen Unterschriften
- <http://www.signplus.com/en> Reine Signatur Verifikation (z.B. für Login – SignSecure) aber auch fokussiert auf Kombination mit digitalen Unterschriften (SignDoc)
- <http://www.valyd.com/> (→ Products/Document Security) Erst ab eSign Advanced Biometrische Funktionalität, darunter “nur” digitale Unterschrift
- <http://www.bio-pen.com/> On-line Signatur Verifikation, Sensorik sitzt im Stift (analog <http://www.smartpen.net/>)
- <http://www.penflow.com/> On-line und Off-line Features, vermischt mit Digitalen Unterschriften (nette Beschreibung des kryptographischen Szenarios → Products/Technological Overview)

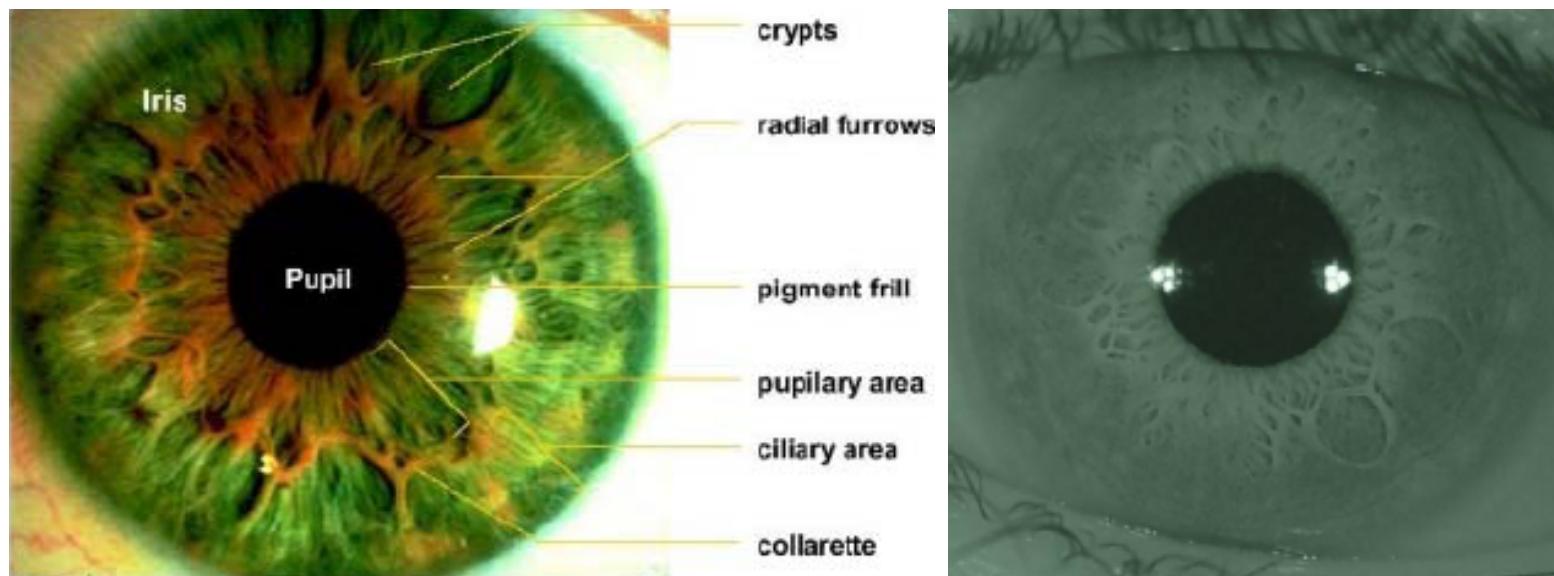
Iris Erkennung: Grundlagen - Anatomie

Die Iris liegt geschützt im Körperinneren und kann doch gut “ausgelesen” werden. Das typische Muster ist spätestens nach dem ersten Lebensjahr stabil und bleibt es auch im wesentlichen. Durch die Muskulatur zur Kontraktion der Pupille verändert die Iris ihre Grösse und Form und zeigt ein regelmässiges Pulsieren (Hippus) – Möglichkeit zur Liveness Detection (z.B. Reaktion auf Licht, Hippus-messung).

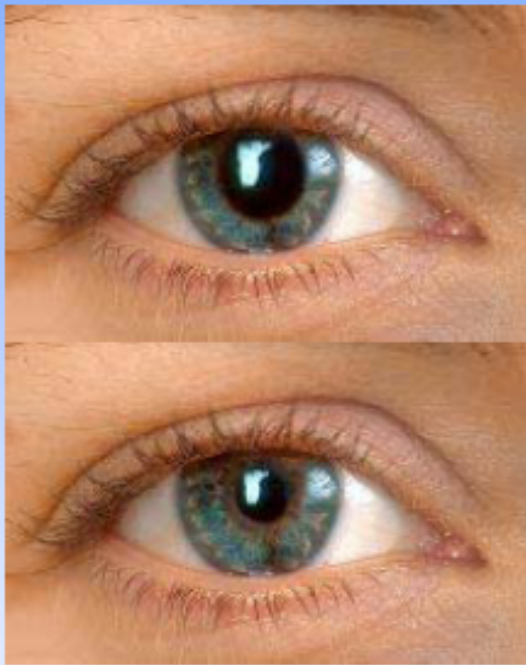


Iris Erkennung: Grundlagen - Bildmaterial

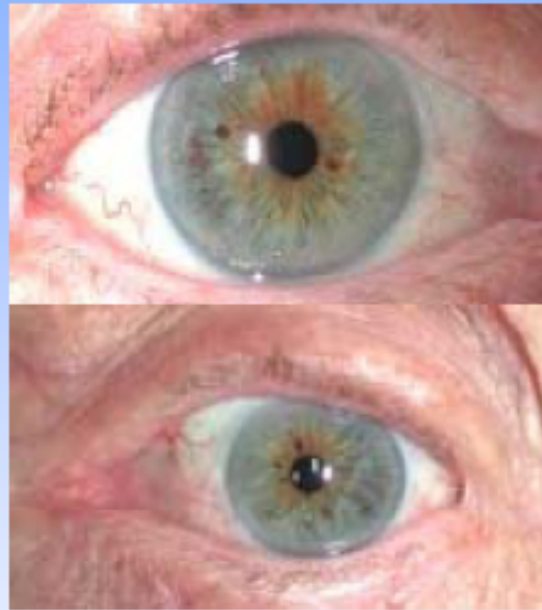
Die Irismusterung zeichnet sich durch hohen Detailreichtum aus. Die Farbinformation wäre für eine Grobklassifikation bei Identifikationsanwendungen nutzbar, wird aber nicht verwendet. Im near-infrared range (NIR) treten die Textureigenschaften besser zutage, sodass diese Art von Bildern eigentlich immer benutzt wird. Günstig ist auch die geringer Benutzerbeeinträchtigung, da NIR Beleuchtung nur als rotes Glimmen wahrgenommen wird.



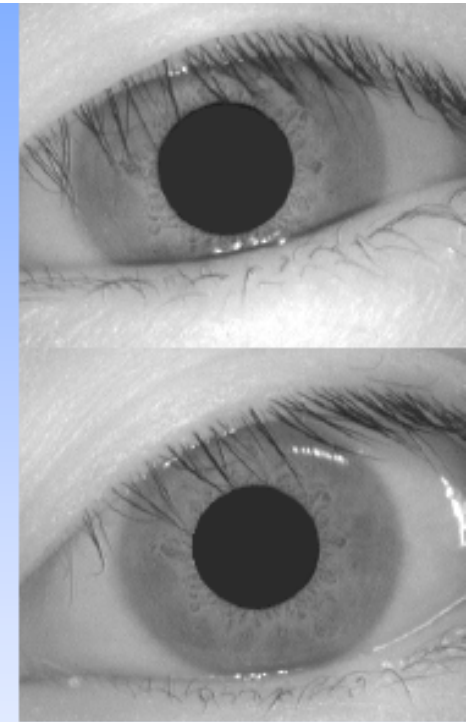
Iris Erkennung: Grundlagen - Intrapersonal Variability



Pupil Dilation
(lighting changes)

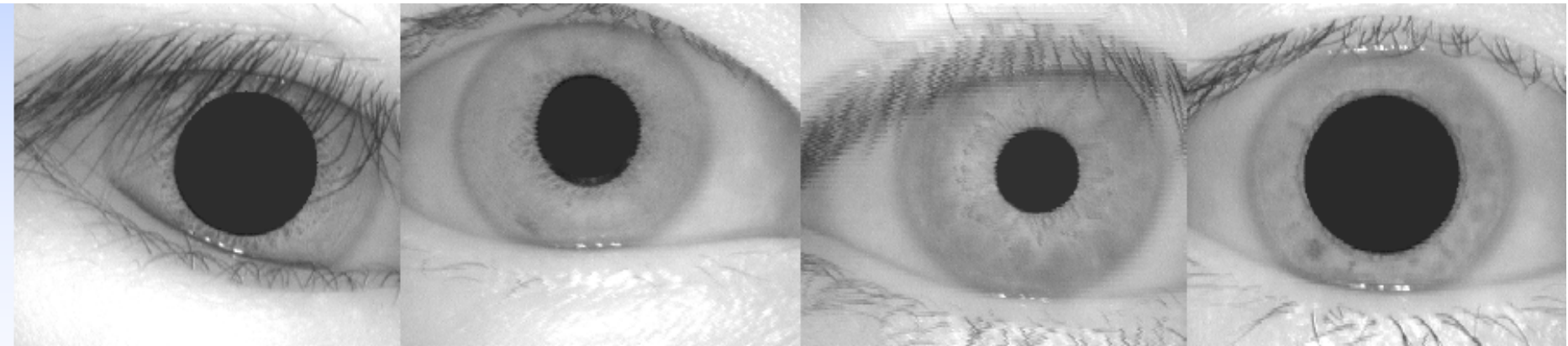


Inconsistent Iris Size
(distance from the camera)



Eye Rotation
(head tilt)

Iris Erkennung: Grundlagen - Probleme



Occlusion (eyelids/eyelashes)

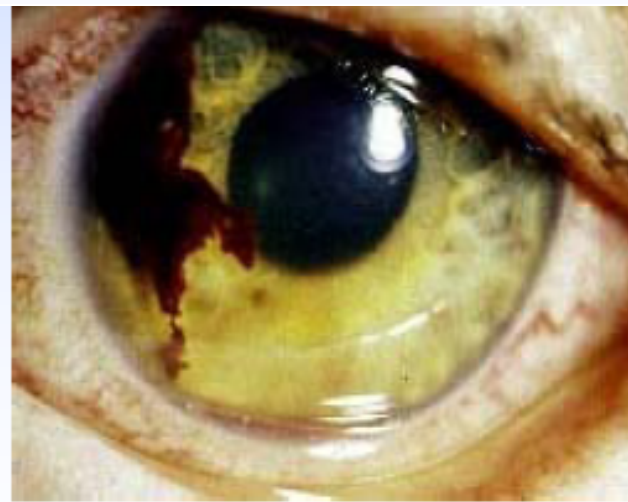
Defocus

Motion blurred

Large pupil



cataract surgery



hyphaema (blood clot)



iridodialysis

Iris Erkennung: Ablauf I

Die meisten Verfahren ähneln sich in den Punkten 1) - 3) sehr stark, die grossen Unterschiede gibt es dann im Bereich der extrahierten und für das Matching verwendeten features und bei den Methoden des Matching selbst.

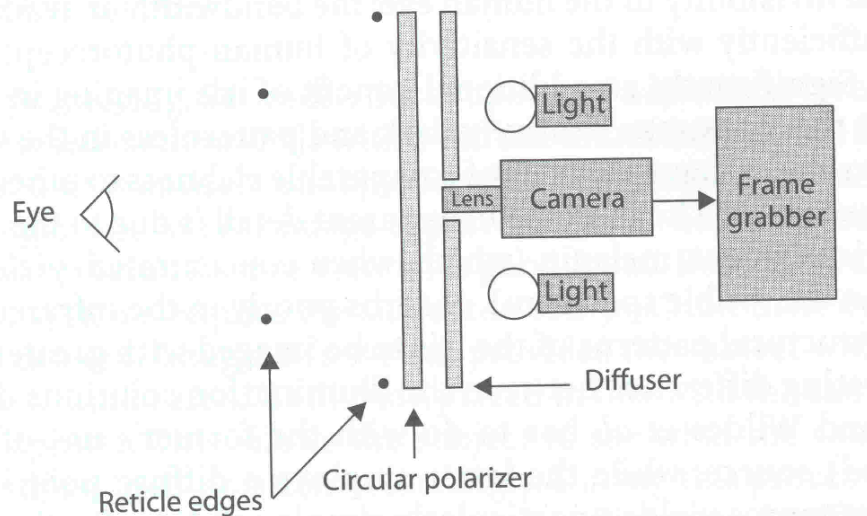
- 1) Aufnahme: alle verwendeten Verfahren benutzen NIR Beleuchtung. Typisch ist eine stark von der Benutzerkooperation abhängige Aufnahmetechnik, die aktive Benutzerpositionierung verlangt und einen relativ kleinen Abstand zur Kamera vorschreibt. Die exakte Positionierung ist durch die Positionsabhängigkeit mancher Verfahren sehr wichtig. Neuere Technologien versuchen dieses restriktive Szenario zu vermeiden und Iris Erkennung im Bereich Surveillance einzusetzen (Iris Aufnahmen aus 10m Entfernung mit Teleobjektiv und Iris Lokalisation und Tracking – schwierig !).
- 2) Iris Lokalisation: aus dem Grauwertbild wird das Objekt "Iris" segmentiert.
- 3) Koordinatentransformation: der resultierende Kreisring ist nicht angenehm zu verarbeiten und weist darüberhinaus unterschiedliche Radien auf – es wird rechteckiges Datenmaterial generiert durch eine Transformation auf Polarkoordinaten.

Iris Erkennung: Ablauf II

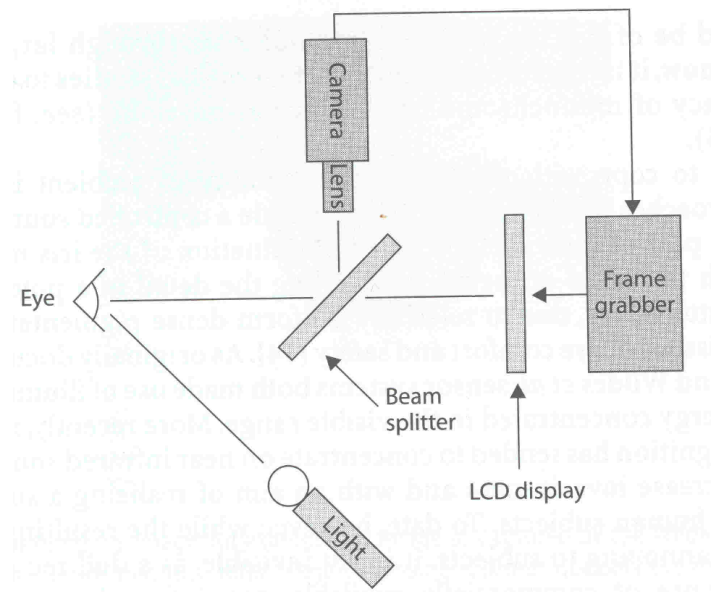
- 4) Image Enhancement: durch Verschiedene Methoden wie Kontrastverbesserung, Histogrammequalisierung oder Abziehen des Mittelwerts werden die Irisstrukturen deutlicher herausgehoben.
- 5) Feature Extraction: hier unterscheiden sich die verschiedenen Verfahren fundamental – gemeinsamkeit ist jedoch, dass praktisch ausschliesslich Wavelet-basierte Verfahren zum Einsatz kommen. Der Grund ist dass die durch Fourier-methoden ermittelbare globale Frequenzverteilung zu wenig spezifisch ist und Lokalisation der verschiedenen Features ein wesentliches Unterscheidungsmerkmal ist.
- 6) Matching: entsprechend den unterschiedlichen Features wird auch das Matching verschieden realisiert.
- 7) Decision: klassische meist threshold-basierte Entscheidungsstrategien.

Iris Erkennung: Aufnahme

Lokalisation wird im ersten System (Wildes) durch zwei versetzte Quadrate unterschiedlicher Größe erreicht, die durch den Benutzer zur Überdeckung gebracht werden müssen. Ist dies erreicht, befindet sich das Auge an den richtigen Koordinaten und im Fokus der Kamera (die Schärfentiefe ist wegen der rel. geringen Beleuchtung und der höheren Brennweite entsprechend gering). Offensichtlich nur für cooperative environments !

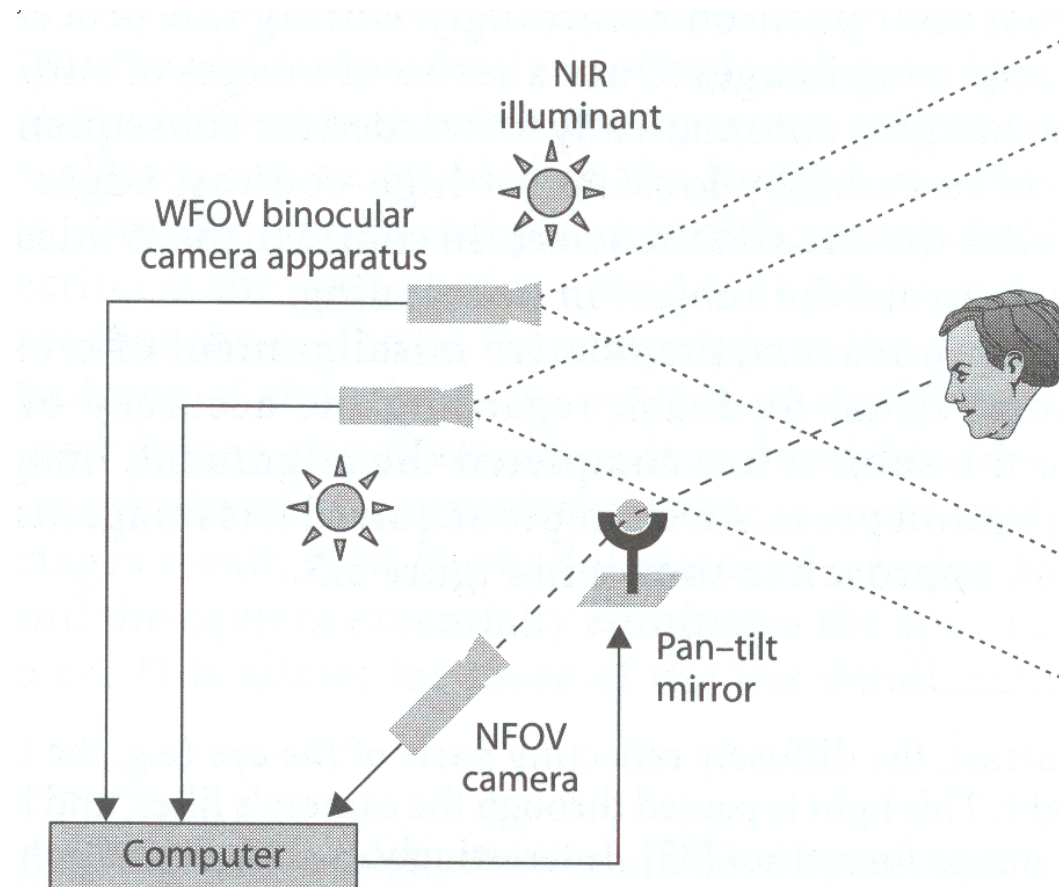


Im zweiten System (Daugman) wird die Positionierung durch Rückkoppelung des Bildes erreicht: die Position des Auges relativ zu den gewünschten Zielkoordinaten wird dem Benutzer über einen Bildschirm angezeigt, sodass er interaktiv eine optimale Ausrichtung erreichen kann. Wieder ist die Notwendigkeit zur Kooperation offensichtlich.



Iris Erkennung: Aufnahme mit Active Vision

Hier ist der Aufnahmeabstand auf 40 - 80 cm erhöht, die Person muss in einem grösseren Bereich (also nicht exakt lokalisiert) mit erlaubter Kopfneigung in beide Achsen stehen. Die Augen werden durch die bi-okularen Weitwinkelkameras erkannt und lokalisiert, die Irisaufnahme wird dann mit einer Telekamera gemacht. Das ganze dauert 2-10 Sekunden (Sarnoff, Sensor Inc.)



Iris Erkennung: Iris Lokalisation I

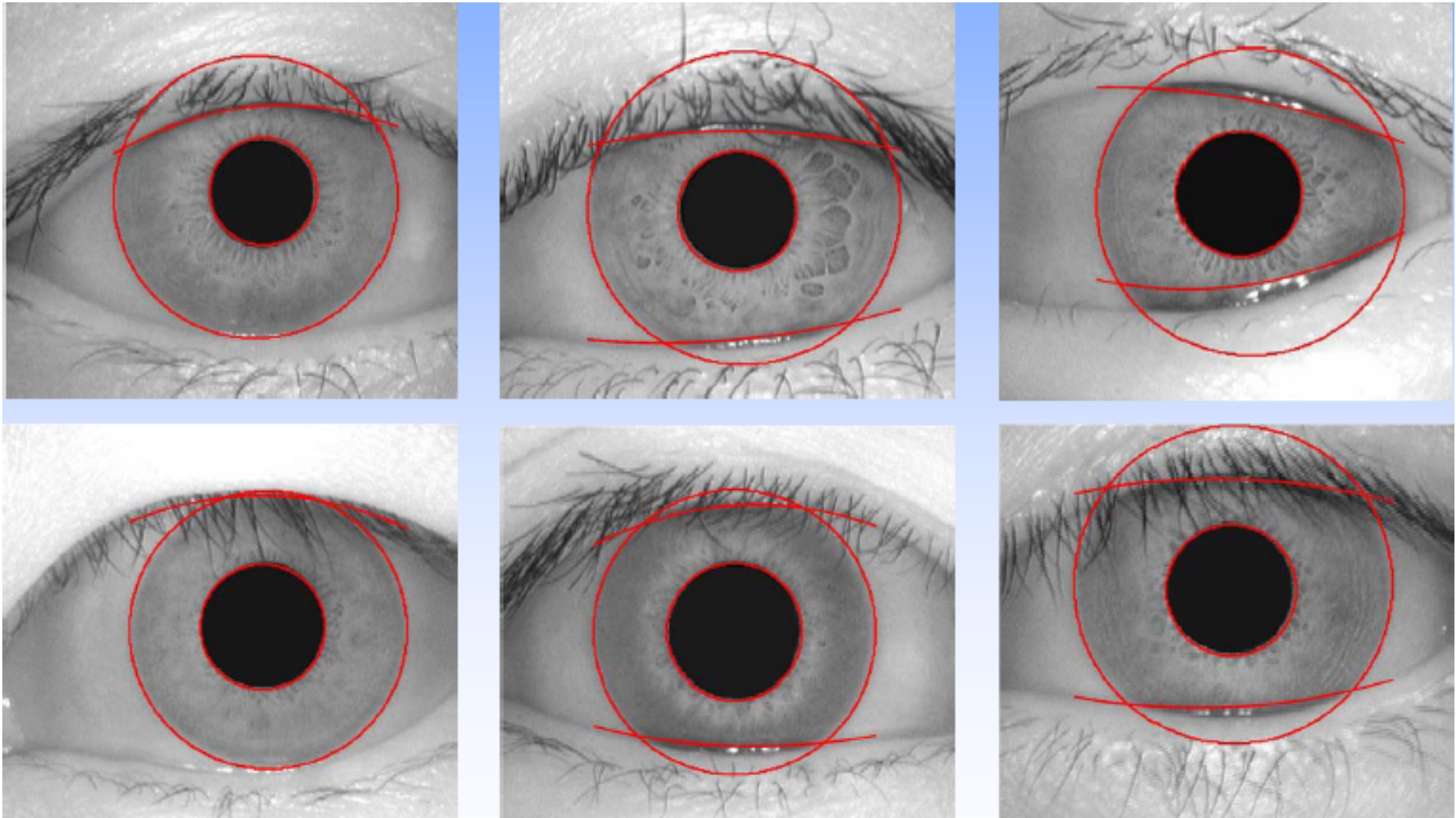
Die Iris ist durch verschiedene Objekte eingegrenzt bzw. teilweise verdeckt: die Pupille von “innen” und der weisse Augenhintergrund von “ausen”, durch Ober- und Unterlid sowie durch die Wimpern. Da alle diese Objektgrenzen relativ deutlich ausgeprägten Kanten entsprechen, ist ein kantenbasierter Segmentierungsansatz vielversprechend. Es ist allerdings zu beachten, dass die Kanten zur Pupille und zum Augenhintergrund hin verschiedene Schärfe (und damit Auflösung) aufweisen. Idealerweise sind daher unterschiedliche adaptierte Kantenerkennungsverfahren optimal.

Nach der Kantenerkennung werden die erhaltenen Kantenpixel auf ein parametrisiertes Kantenmodell abgebildet – eine klassische Anwendung für die Hough Transformation. Die Grenzen zur Pupille und zum Augenhintergrund werden durch Kreise dargestellt wobei zu beachten ist dass diese Kreise nicht notwendigerweise konzentrisch sind.

Einige Verfahren setzen explizite Liderkennung ein und modellieren die entsprechenden Grenzen z.B. durch Parabeln oder Kreisbögen.

Es gibt bisher keine Methoden die sich explizit mit Wimpernartefakten beschäftigen.

Iris Erkennung: Iris Lokalisation II



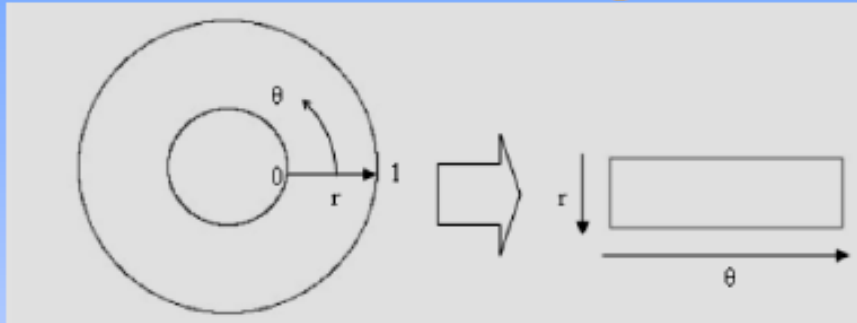
Iris Erkennung: Koordinatentransformation

Der radiale Wertebereich zwischen der Pupillengrenze und der Augenhintergrundgrenze wird auf die Koordinate r im Bereich $[0, 1]$ abgebildet, die Position auf dem Kreisbogen durch den Winkel Θ . Seien (x_i, y_i) die ursprünglichen Koordinaten der Irispixel, (x_0, y_0) der Pupillenmittelpunkt und r_0 der Pupillradius, weiters M der Zielabstand zwischen Pupillen- und Augenhintergrundgrenze in Pixel und L der tatsächliche Abstand im Bild.

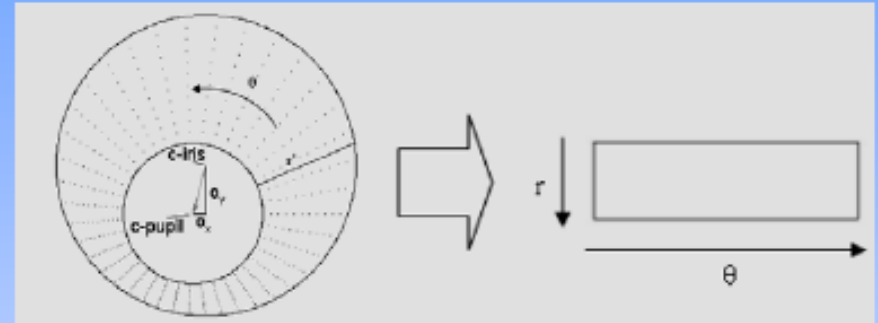
$$r_i = \frac{M}{L} \left(\left[(x_i - x_0)^2 + (y_i - y_0)^2 \right]^{1/2} - r_0 \right)$$

$$\Theta_i = \arcsin \left(\frac{y_i - y_0}{x_i - x_0} \right) \text{ für } y_i \geq y_0 \text{ sonst plus Pi.}$$

Daugman's Rubber Sheet Model

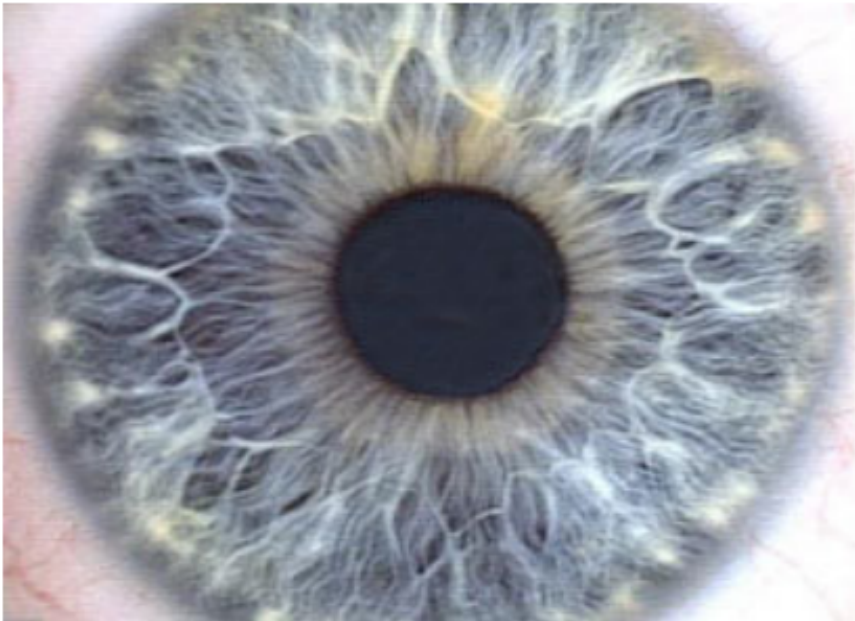


Centers of iris and pupil coincide



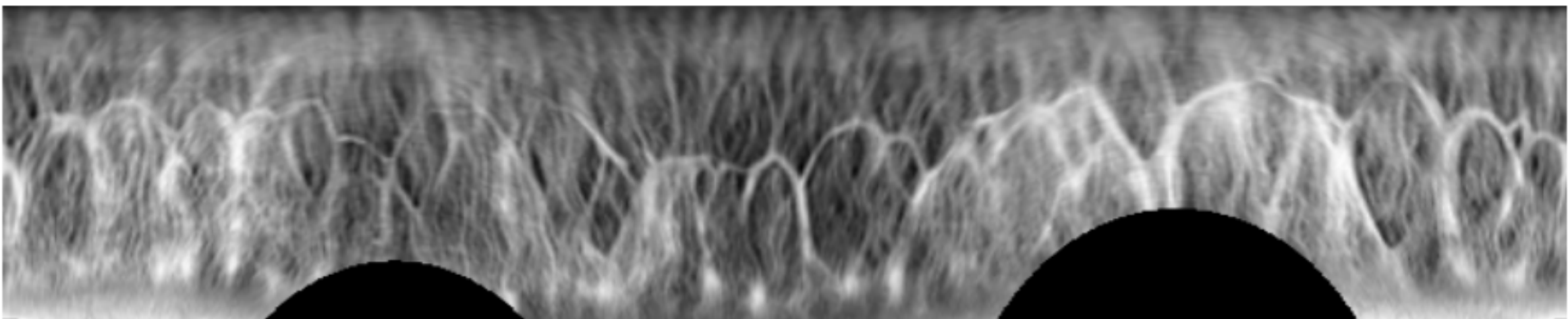
Centers of iris and pupil do not coincide

Iris Erkennung: Koordinatentransformation Beispiel



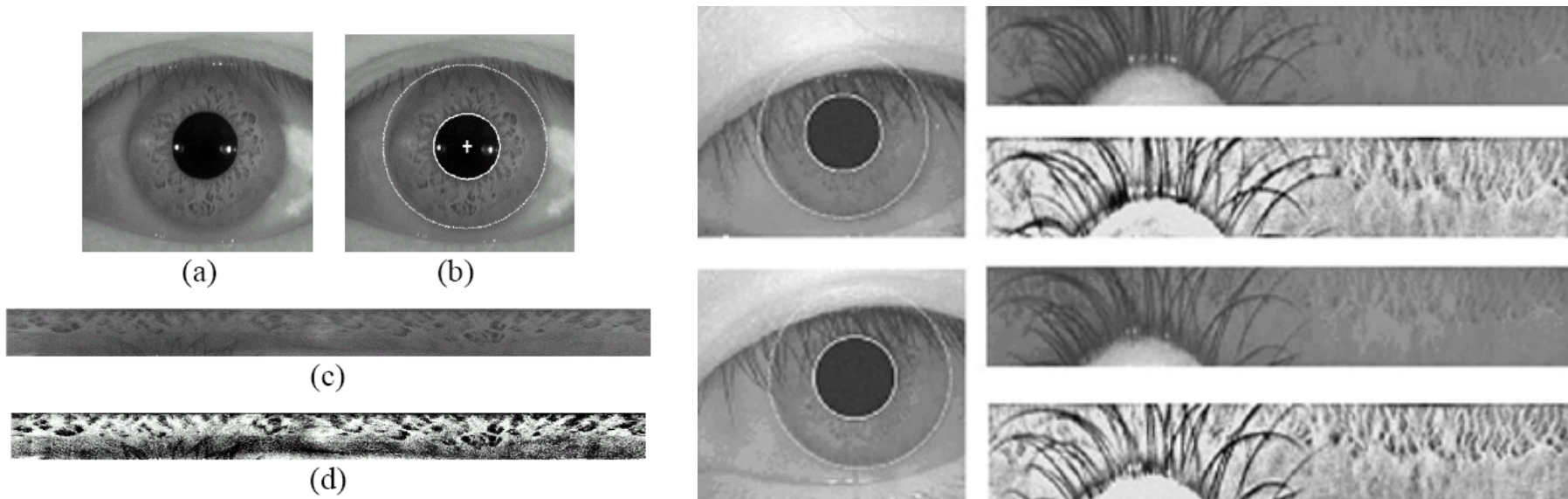
Infra-red illuminated image of iris

Iris image is 'unwrapped'

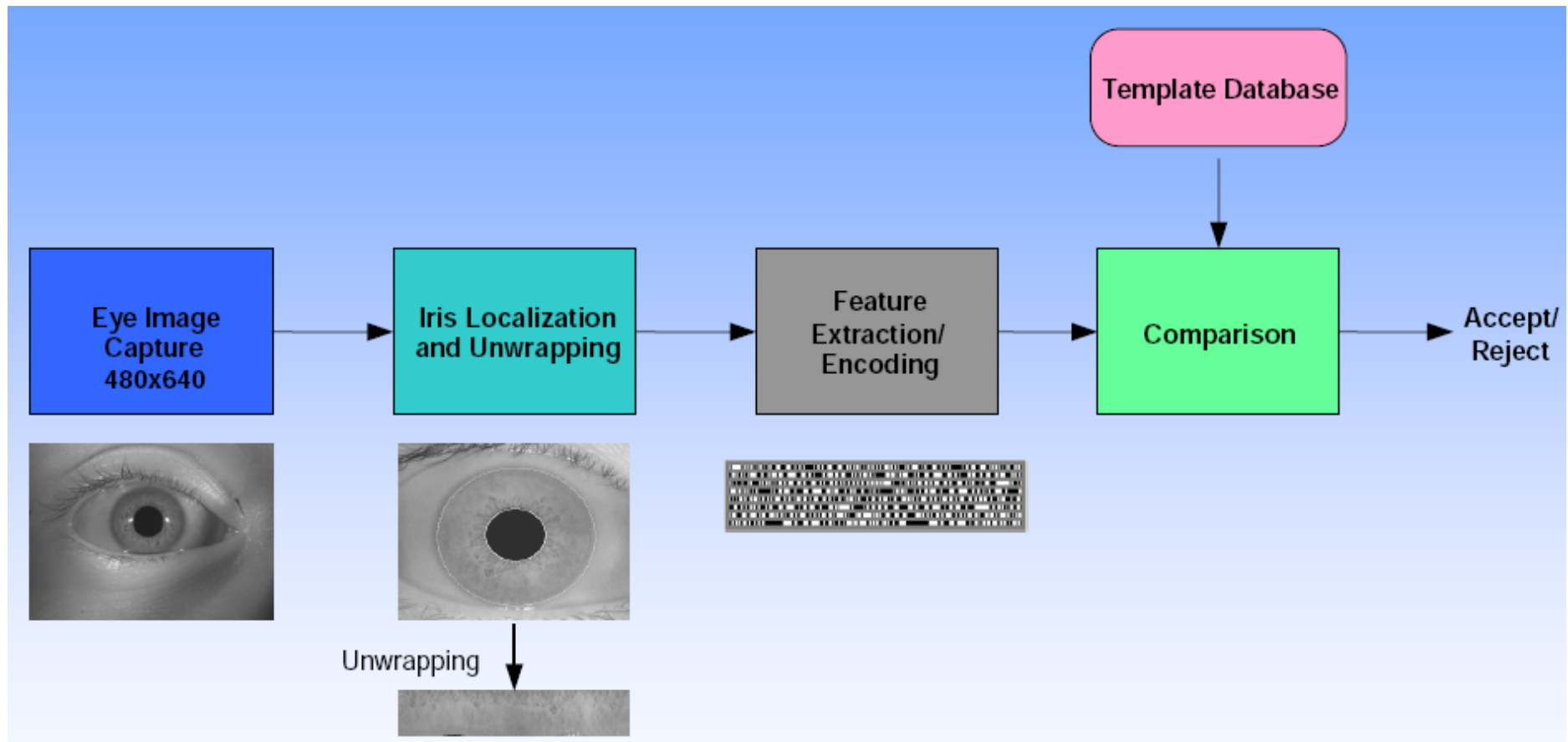


Iris Erkennung: Koordinatentransformation in der Praxis

Durch Verdeckung grosser Teile der Iris kann es in der Praxis durchaus zu Problemen kommen. In der linken Graphik ist der wünschenswerte Idealfall dargestellt, rechts sieht man starke Artefakte durch Augenlid- und Wimpernüberdeckung. Kann durch Benutzerkooperation vermindert werden.



Iris Erkennung: Das Daugman Verfahren



Iris Erkennung: Das Daugman Verfahren – Gabor Features I

Es werden 2D Gabor Funktionen eingesetzt:

$$\Psi(x, y) = e^{-\pi[(x-x_0)^2/\alpha^2 + (y-y_0)^2/\beta^2]} e^{-2\pi i[u_0(x-x_0) - v_0(y-y_0)]}$$

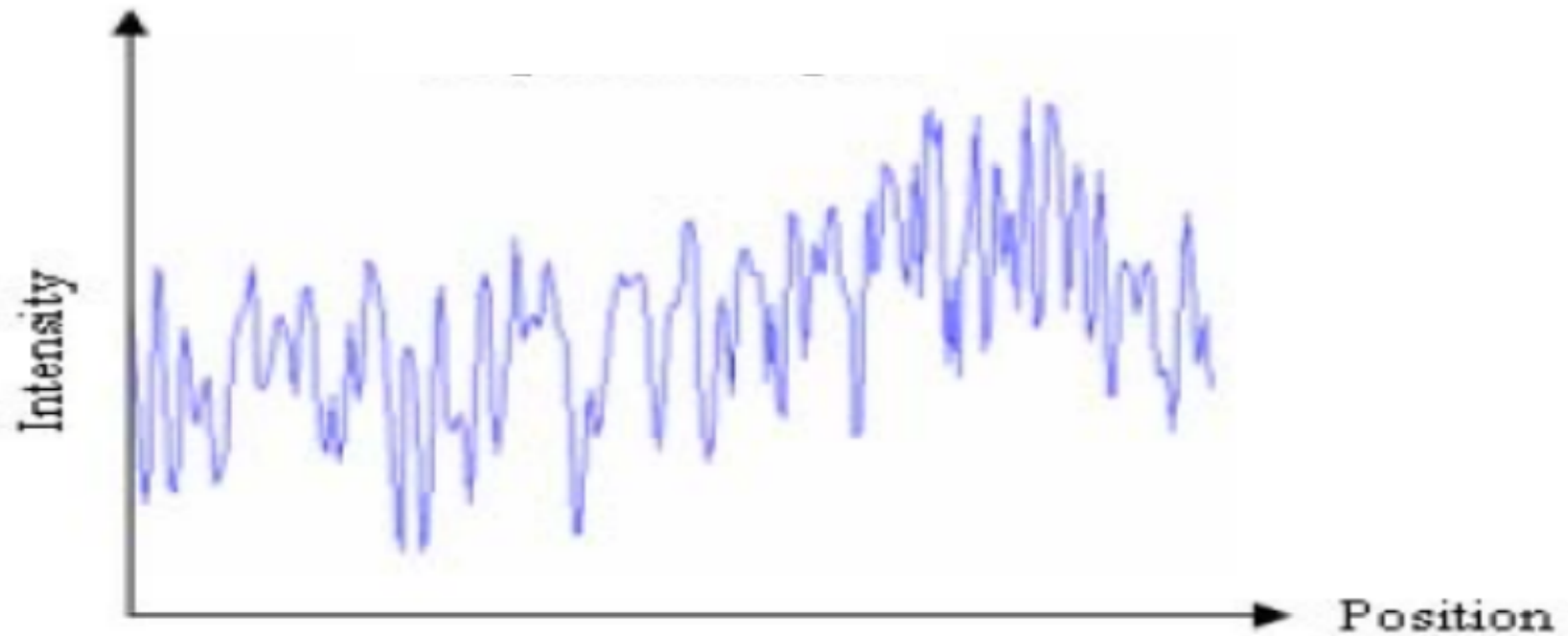
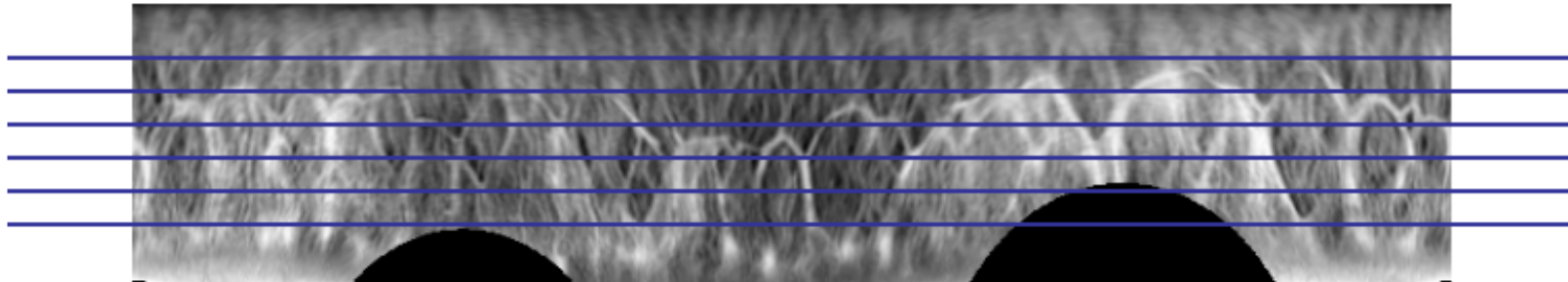
mit (x_0, y_0) die Position des Wavelet, (α, β) dessen Länge und Breite, (u_0, v_0) sind Modulationsparameter die in Polarkoordinaten die Frequenz $\omega = (u_0^2 + v_0^2)^{1/2}$ und die Orientierung $\Theta = \arctan(v_0/u_0)$ angeben.

Da diese Wavelets komplexwertig sind kann man Imaginär- und Realteil ihrer Faltung * mit einem Bild $I(x, y)$ als Beschreibung des Bildes bezüglich lokaler Amplitude und Phase verwenden. Der entsprechende Phasenwinkel (Phase Modulation) ist

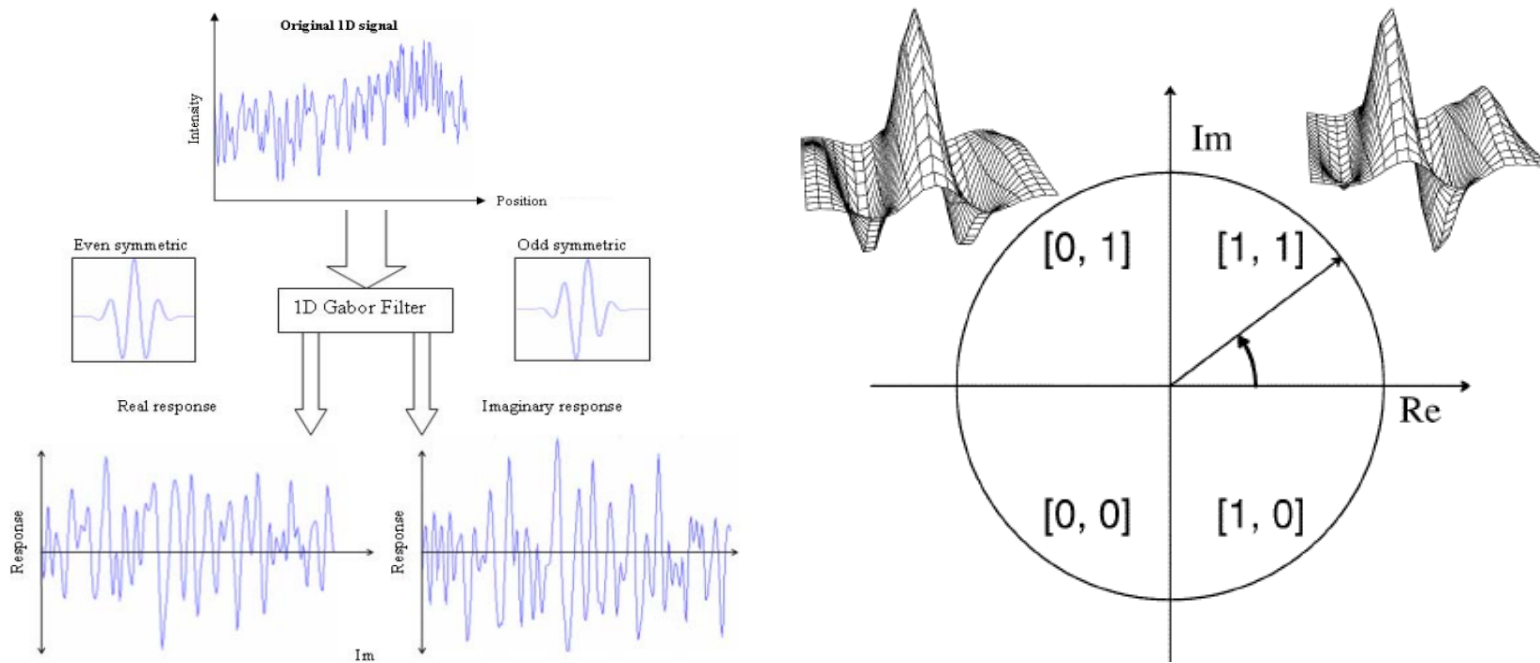
$$\phi(x, y) = \tan^{-1} \frac{\text{Im}\{\Psi(x, y) * I(x, y)\}}{\text{Re}\{\Psi(x, y) * I(x, y)\}} .$$

Dieser Phasenwinkel auf eine 2 Bit Darstellung quantisiert wird als “Iris Code” bezeichnet.

Iris Erkennung: Das Daugman Verfahren – Gabor Features II



Iris Erkennung: Das Daugman Verfahren – Gabor Features III

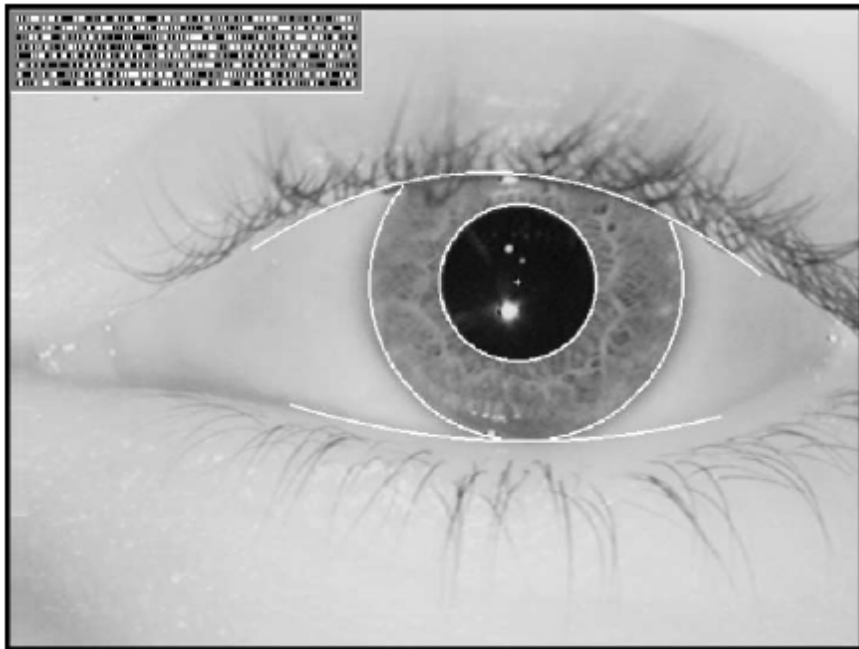


Die auf das in Polarkoordinaten gegebene Irismuster $I(r, \Theta)$ angewendete Faltung $* = G(r_0, \Theta_0, \alpha, \beta, \omega) = G$ ist definiert wie folgt (beachte dass die Gabor Funktionen nur bzgl. ihrer Frequenz nicht aber bzgl. ihrer Orientierung wechseln):

$$G = \int_r \int_{\Theta} e^{-i\omega(\Theta_0 - \Theta)} e^{-(r_0 - r)^2 / \alpha^2} e^{-(\Theta_0 - \Theta)^2 / \beta^2} I(r, \Theta) r d\Theta dr .$$

Iris Erkennung: Das Daugman Verfahren – Iris Code

Ist $Re\{G\} \geq 0$ wird das erste Bit 1 gesetzt (ansonsten 0), ist $Im\{G\} \geq 0$ wird das zweite Bit 1 gesetzt (ansonsten 0). Auf diese Art werden 2048 solcher Phasen Bits (i.e. 256 bytes) berechnet (8 Grössen und entsprechende Frequenzen werden verwendet, 2 Bits pro Position, an 128 Positionen (r, Θ)).

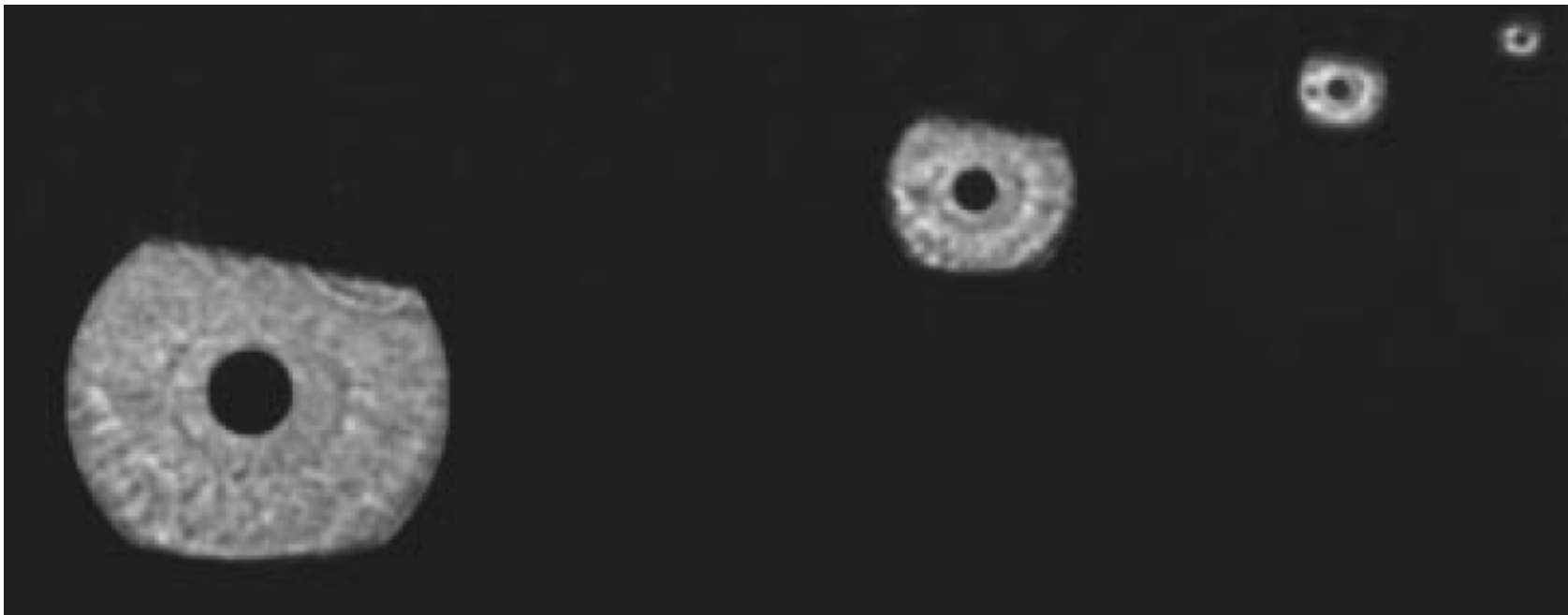


In neueren Implementierungen werden zusätzliche 2048 Maskierungsbits angegeben die anzeigen wenn der entsprechende Teil des Codes wegen Occlusions oder geringer Qualität unberücksichtigt bleiben sollte. Zum Matching wird einfach die Hamming Distanz zwischen zwei Iris Codes berechnet (i.e. die Anzahl der ungleichen Bits).

Rotationsinvarianz wird durch zyklisches verschieben von zwei Iriscodes gegeneinander erreicht. Die niedrigste Distanz gibt dann den tatsächlichen Abstand an. Es ist zu beachten dass durch das Verwenden der Vorzeichen bzw. der Vorzeichenänderung in der Phase implizit die Position von Zerocrossings in Θ -Richtung kodiert wird.

Iris Erkennung: Das Wildes et al. Verfahren

Wurde in etwa zur selben Zeit entwickelt (Mitte der 90er), ist aber deutlich weniger ausgefeilt. Nach der Irissegmentierung wird die Iristextur also sog. Laplace Pyramide dargestellt: das Iris Bild wird rekursiv mit einem 2D Gaussfilter gefaltet und in jedem Schritt einem Downsampling unterzogen. Gespeichert wird die kleinste Pyramidenebene und die drei Differenzbilder zwischen je zwei Auflösungen (wobei die gröbere Auflösung zuvor interpoliert wird um die gleiche Auflösung zu erhalten (vgl. z.B. hierachical progressive JPEG)).



Iris Erkennung: Das Wildes et al. Verfahren - Matching

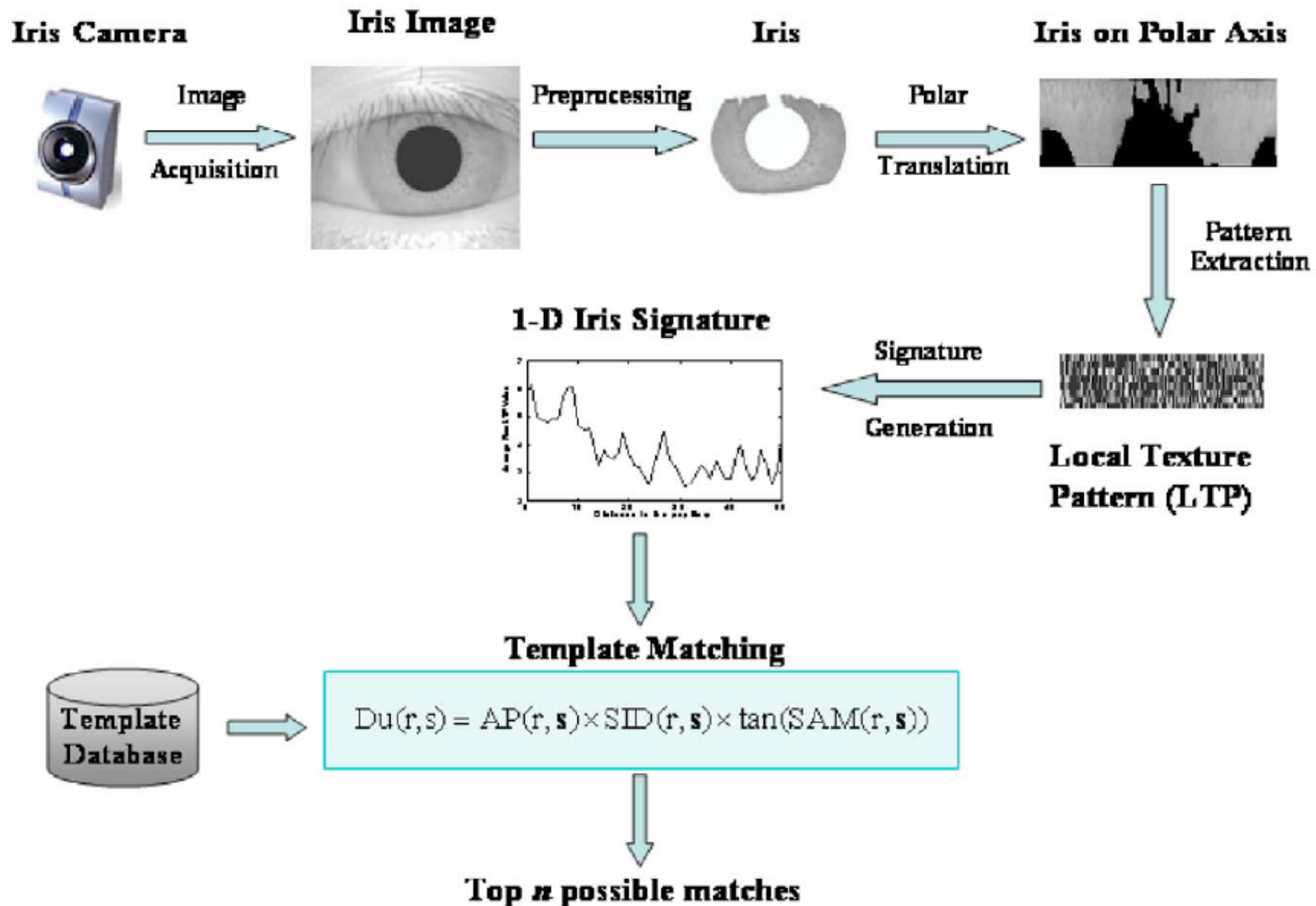
Alignment wird durch die Ausrichtung nach dem Pupillenmittelpunkt und durch entsprechende Rotationen beim Matching erreicht. Matching selbst wird durch normierte Korrelation zwischen zwei Bildern $p_1(i, j)$ und $p_2(i, j)$ berechnet (wobei μ_1 und σ_1 Mittelwert und Varianz des Bildes p_1 sind, n, m sind die Bilddimensionen):

$$\frac{\sum_{i=1}^n \sum_{j=1}^m (p_1(i, j) - \mu_1)(p_2(i, j) - \mu_2)}{nm\sigma_1\sigma_2}$$

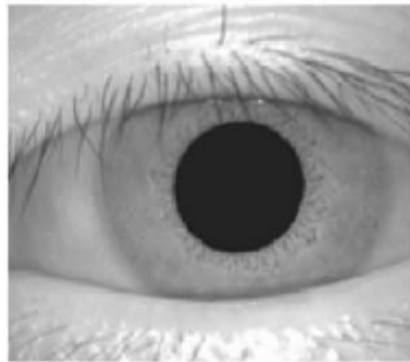
In der Implementierung wird die Korrelation auf 8×8 Pixel Blöcken berechnet in jedem der 4 Frequenzbänder. In jedem Band wird der Median der Blöcke verwendet, was insgesamt zu vier Matchwerten führt, die entsprechend kombiniert werden müssen.

Insgesamt macht dieser Ansatz einen wesentlich unkonkreteren Eindruck, der fix definierte Iris Code des Daugman Systems ist wesentlich effizienter handhabbar als 4 Korrelationswerte, wo auch die Kombination und Gewichtung nicht klar definiert ist. Dementsprechend sind die kommerziellen Anwendungen auf das Daugman System konzentriert.

Iris Erkennung: Das Du et al. Verfahren: Übersicht



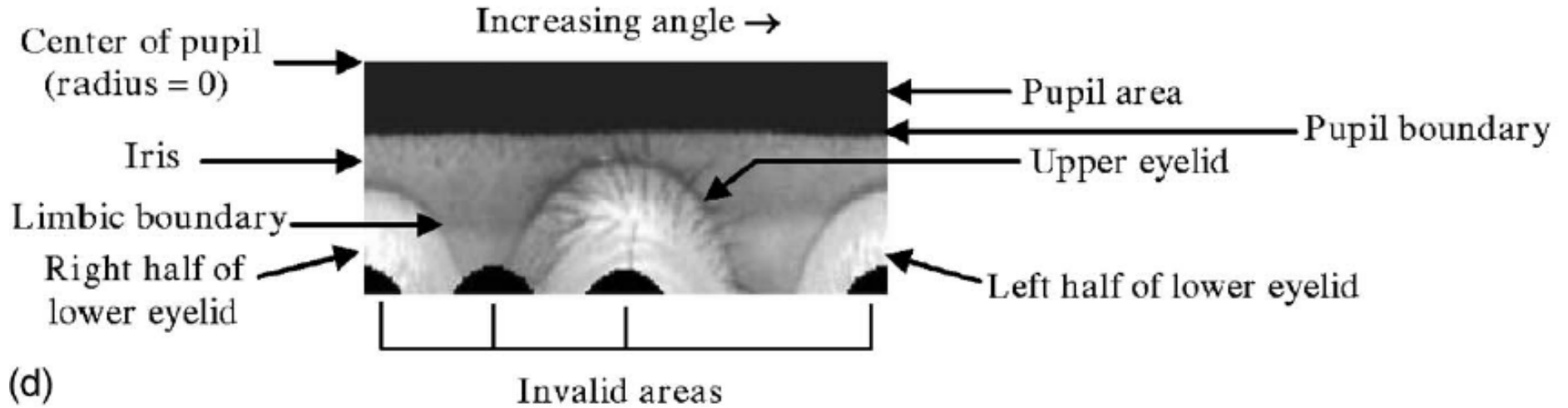
Iris Erkennung: Das Du et al. Verfahren: Vorverarbeitung



(b)



(c)

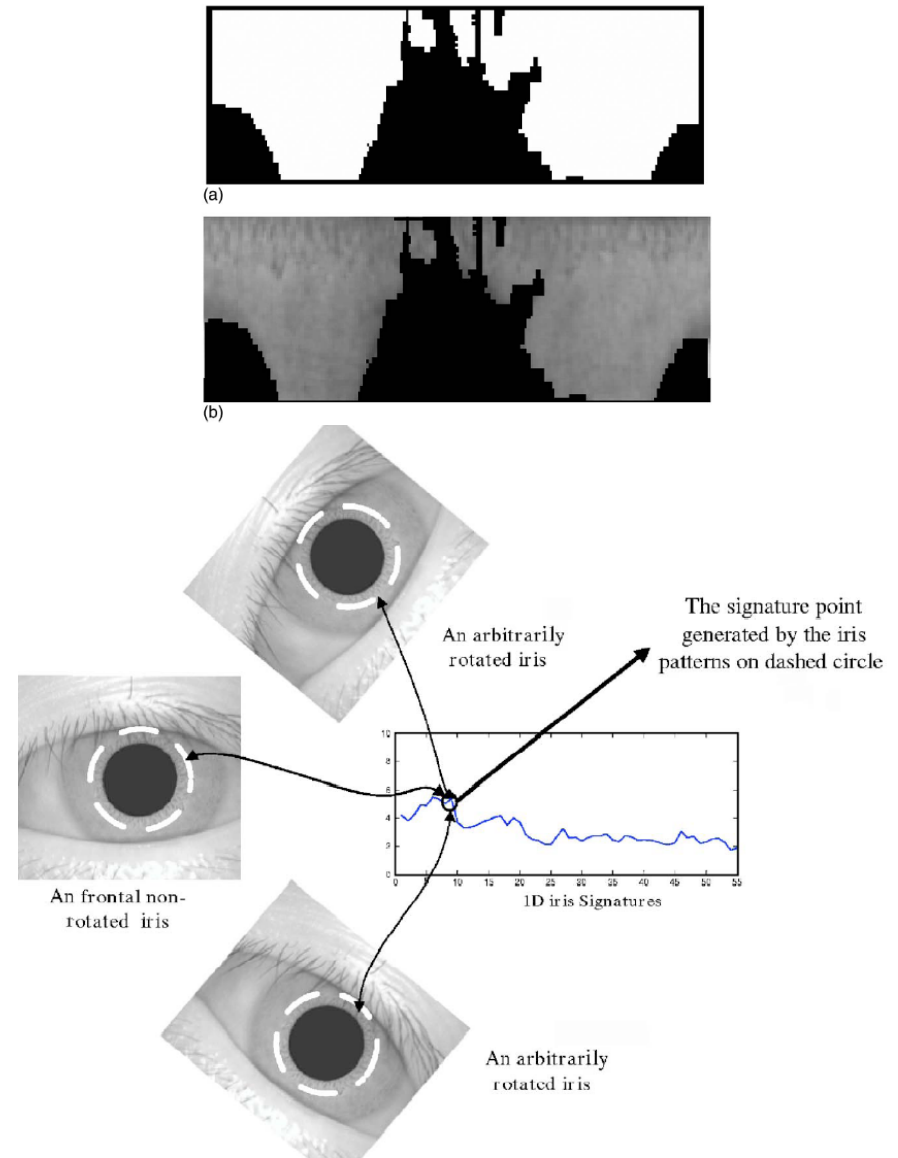


(d)

Iris Erkennung: Das Du et al. Verfahren: Features

Zuerst wird die Iristextur in Polarkoordinaten einer lokalen Subtraktion eines Umgebungsmittelwerts in einem Fenster unterzogen – es bleiben die Schwankungen um den Mittelpunkt (mehr nahe der Pupille - siehe Graphik). Anschließend werden die Werte einer Zeile (entspricht den Werten in einem konzentrischen Ring in den Originaldaten) aufsummiert, nur wenn über 60% der Pixel Iristextur (und nicht occlusion) sind. Diese Werte in Abhängigkeit von der Entfernung zum Pupillenmittelpunkt ergeben die 1D Signatur. **ACHTUNG**: diese ist vollständig rotationsinvariant !!

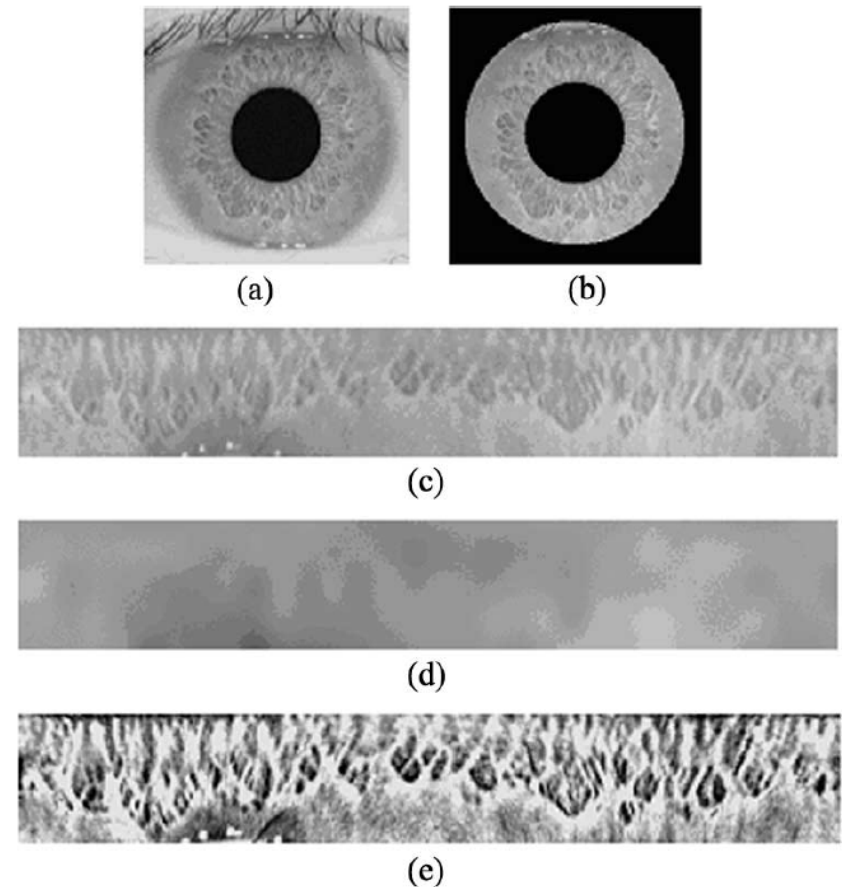
Möglicherweise werden die Daten hier zu sehr verdichtet – es wird die Kombination mit einer anderen biometrischen Modalität vorgeschlagen.



Iris Erkennung: Die Verfahren von Ma und Boles et al.: Vorverarbeitung

Vorverarbeitung ist wieder ähnlich mit Konvertierung zu Polarkoordinaten und Abziehen des lokalen Mittelwerts. Ma verwendet dann noch eine Histogrammequalisierung in einem 32×32 Pixel Fenster (siehe Bild).

Im weiteren werden die Daten in einem Kreisring gemittelt, d.h. M untereinanderliegende Zeilen in Polarkoordinaten werden pixelweise gemittelt und es entsteht ein 1D Signal pro Kreisring. Es werden 78% der Daten (Ringe näher zur Pupille) verwendet um 10 1D Signaturen zu erzeugen, die extreme Verdichtung des Du Verfahrens unterbleibt. Boles verwendet nur einen Kreisring der Breite 3. Rotationsinvarianz wird in beiden Fällen durch Translation der Iristextur beim Matching erreicht.

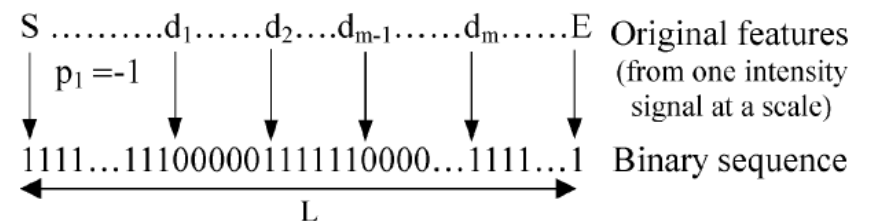


Iris Erkennung: Die Verfahren von Ma und Boles et al.: Features

Die extrahierten Features der beiden Verfahren unterscheiden sich dann etwas, aber nicht wesentlich. Beide verwenden multiscale representations basierend auf diskreter Wavelettransformation und heben v.a. starke Variationen heraus, beide benutzen zwei Skalen eher geringer Auflösung einer nicht-downgesampelten DWT.

Ma et al. suchen die lokalen Extremwerte der beiden Detailsignale und betrachten immer benachbarte Paare von Extrempunkten (mit ausreichend grossem Amplitudenabstand um signifikante Extremaleigenschaft zu garantieren) – ein lokales Maximum und ein lokales Minimum. Die Positionen dieser Extremwertpaare werden gespeichert, für beide betrachteten Skalen und für alle betrachteten 1D Signaturen – dies ergibt den vorläufigen Featurevektor.

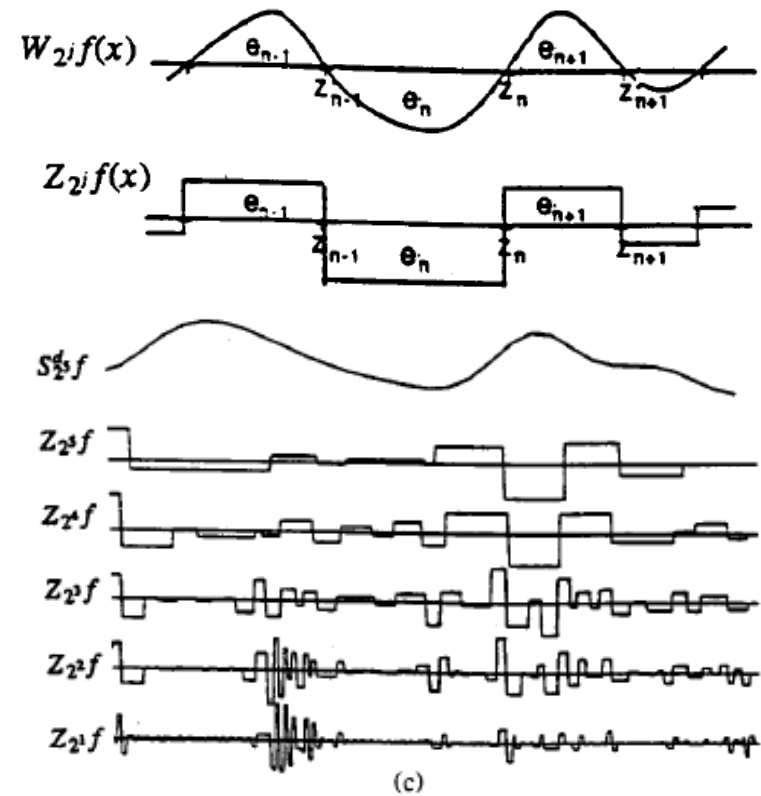
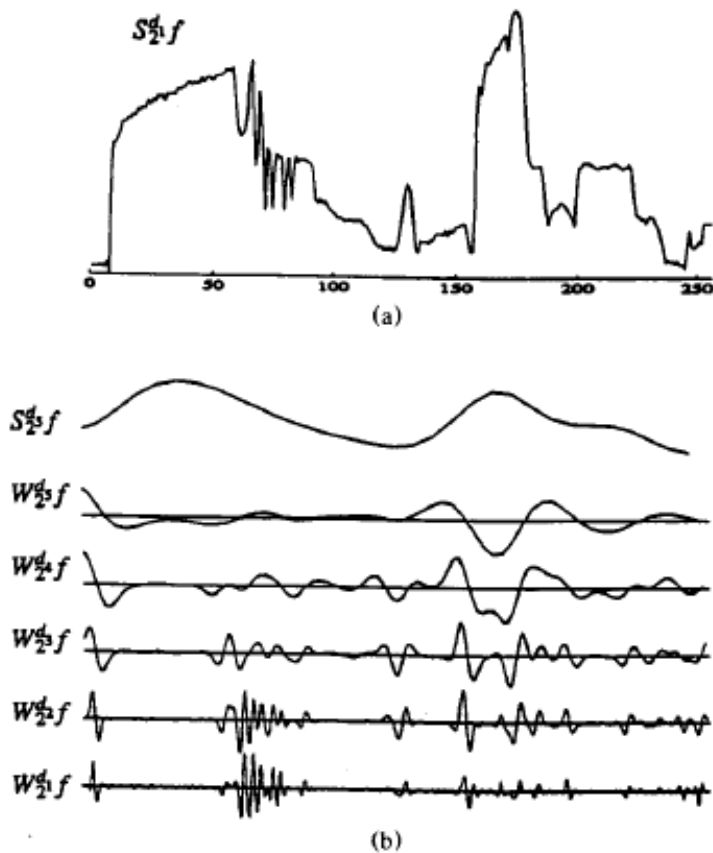
Nun werden diese Featurevektoren in binäre Vektoren umgewandelt um XOR oder Hammingdistanz anwenden zu können: An jeder Position die im Featurevektor angegeben ist, wird von 0 auf 1 gewechselt (oder umgekehrt), begonnen wird mit dem entsprechenden Typ des Extremwerts; d.h. der binäre Vektor ist gleich lang wie 2x das Originalsignal (2 Skalen).



S: the starting point of the sequence (i.e., 1)
E: the ending point of the sequence (i.e., L)

Iris Erkennung: Boles et al.: Wavelet Zero Crossings

Wavelet Zero Crossings ist ein von S. Mallat entwickeltes Verfahren um Wavelet Transformationskoeffizienten effizienter zu beschreiben und auch die Möglichkeit einer exakten Rekonstruktion zu bieten. Dabei werden nur die Positionen der Nullstellen der Detailsignale und ein konstanter Wert zwischen diesen Nullstellen behalten (der Wert wird so gewählt dass das Integral zwischen zwei Nullstellen erhalten bleibt).



Iris Erkennung: Das Verfahren von Boles et al.: Matching

Sei die Zero-crossing Darstellung auf Scale j einer 1D Iris-Signatur f gegeben als $Z_j f$. $Z_j f$ kann eindeutig dargestellt werden durch eine Menge von geordneten komplexen Zahlen deren Imaginärteile die Position der Nullstellen und deren Realteile die Grösse von $Z_j f$ zwischen benachbarten Nullstellen angeben. Um hier ein Abstandsmass direkt anwenden zu können muss die Anzahl der Nullstellen auf dem betreffenden Scale identisch sein – es wird vorgeschlagen diese Strategie nur zu verwenden wenn zwei benachbarte Scales gefunden werden, bei denen diese Bedingung erfüllt ist.

Alternativ wird vorgeschlagen direkt den Abstand zwischen den $Z_j f$ der zu matchenden 1D Signaturen zu berechnen (was wesentlich aufwändiger ist) oder die Zero-crossing Darstellung nachzubearbeiten um auf die gleiche Anzahl von Nullstellen zu kommen (Elimination von “falschen” Nullstellen).

Beide Strategien werden mit gegeneinander verschobenen 1D Signaturen durchgeführt und der minimale Wert wird verwendet (um Rotationsinvarianz zu erreichen).

Iris Erkennung: Das Zhu et al. Verfahren

Die Vorverarbeitung entspricht den restlichen Verfahren bis schlussendlich die rechteckige Iristextur in Polarkoordinaten gegeben ist. Im Anschluss daran werden Gabor und DWT Filterung (Daub4) angewendet, aus den resultierenden Subbands werden Mittelwert und Standardabweichung berechnet und als Features verwendet – hier werden also klassische Verfahren der Texturklassifikation eingesetzt. Die verwendeten Gaborfilter gegeteiliger Symmetrie $h_e(x, y)$ und $h_o(x, y)$ sind gegeben als

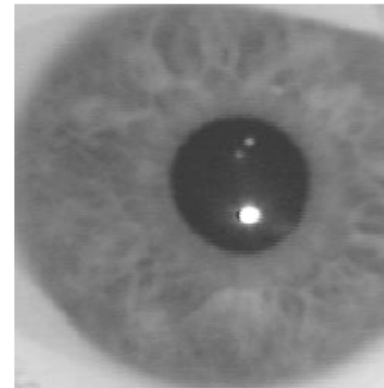
$$h_e(x, y) = g(x, y) \cos[2\pi\omega(x\cos\Theta + y\sin\Theta)]$$

mit $g(x, y)$ eine 2D Gauss Funktion und $h_o(x, y)$ mit $\sin[]$ anstelle von $\cos[]$. Es werden 24 verschiedene Kombinationen von ω und Θ verwendet was zu 48 Features führt. Im Wavelet Fall werden 13 Subbands verwendet (26 Features).

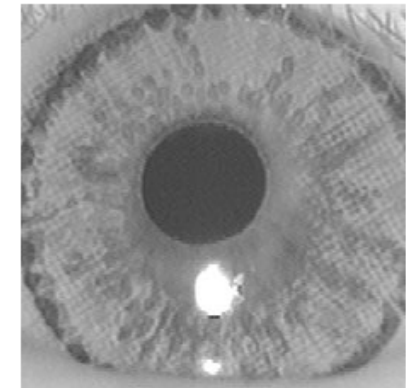
Im Gegensatz zu den meisten anderen Verfahren wird hier durch die ausschliessliche Verwendung von statistischen Textur Deskriptoren Rotations- und Translationsinvarianz erreicht. Das individuelle tatsächliche Iris Muster geht allerdings verloren und die Ergebnisse sind dementsprechend deutlich schlechter, die reine Texturcharakteristik ist hier zu wenig diskriminativ – offensichtlich gibt es hier einen Tradeoff !

Iris Erkennung: Täuschungsmöglichkeiten

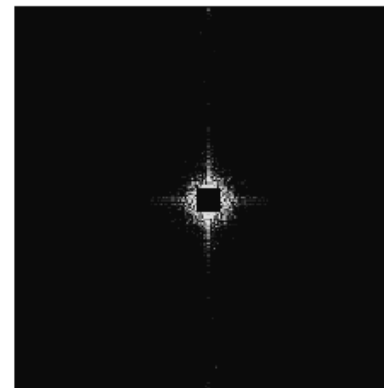
Eine offensichtliche Täuschungsmöglichkeit ist die Verwendung von bedruckten Kontaktlinsen. Eine einfache (aber ev. unangenehme) Möglichkeit ist die schnelle Veränderung der Beleuchtungsverhältnisse um Veränderung der Pupillenweite messen zu können (Abdunklung bringt durch die Notwendigkeit zur Öffnung die unterbleiben muss klarere Ergebnisse). Andere Möglichkeiten sind die Messung des Hippus (hohe Auflösung in spatial und temporal domain nötig) und die Erkennung der Kontaktlinsenränder durch Kantenerkennung. Die Abbildung zeigt dass die Amplitude der DFT Druckartefakte aufweist, ob das bei höherer Qualität auch so wäre ist zu bezweifeln.



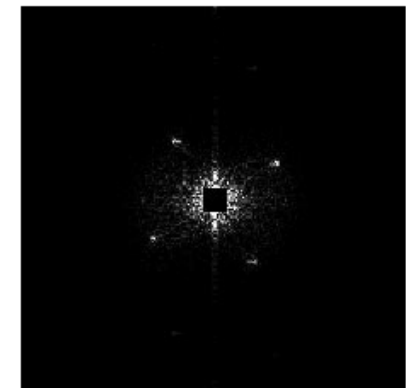
Natural iris



Fake iris printed on a contact lens



2D Fourier spectrum of natural iris



2D Fourier spectrum of fake iris

Iris Erkennung: Produkte

Die Daugman Technologie dominiert den kommerziellen Bereich, sie wird durch Iridian (<http://www.iridiantech.com/>) lizenziert und vertrieben. LG Electronics / Iris Technology Division hat die grösste Verbreitung mit über 1000 Sites installiert.

Einreise Watchlistcheck in die Vereinigte Arabische Emirate (IrisGuard Technologie (<http://www.irisguard.com/>) an allen 17 Grenzstationen; jeder der täglich ca. 7000 Einreisenden wird gegen 544000 Einträge der Watchlist über Internetverbindung zu einem zentralen Server geprüft.



Iris Erkennung: Produkte II – Frequent Flyers

Shipol (NL), Narita (Japan) und Frankfurt (dieses System von Byometric Systems AG, verwendet aktive Kamera, passt sich automatisch der Passagiergröße an).



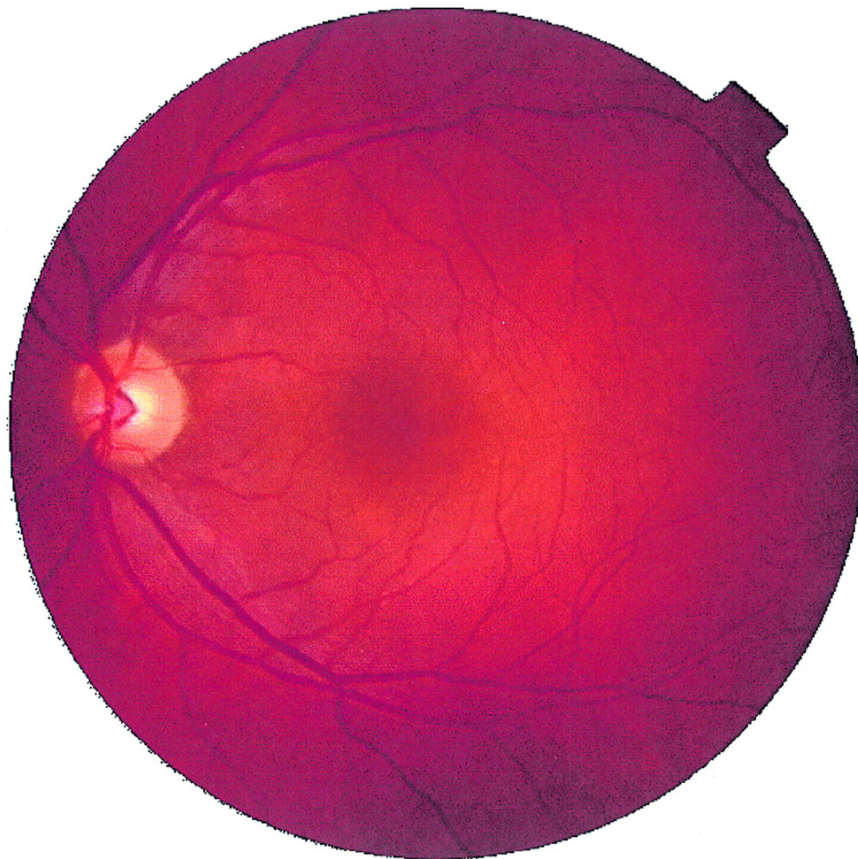
Iris Erkennung: Produkte III

Flughafen Albany (Zugangskontrollen in gesperrte Bereiche nur für Berechtigte),
Securimetrics (<http://www.securimetrics.com/>) produziert tragbare Geräte für
Enrollment & Verifikation (auch multimodal), v.a. militärischer Einsatzbereich.



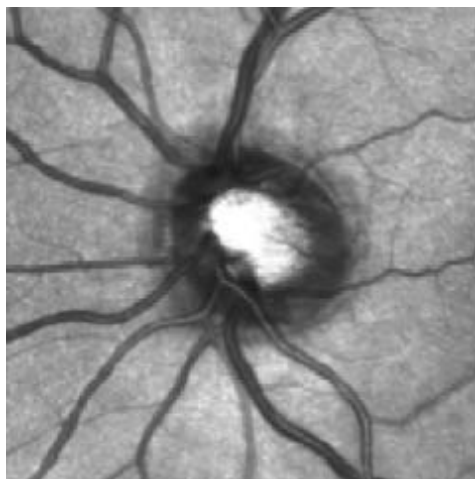
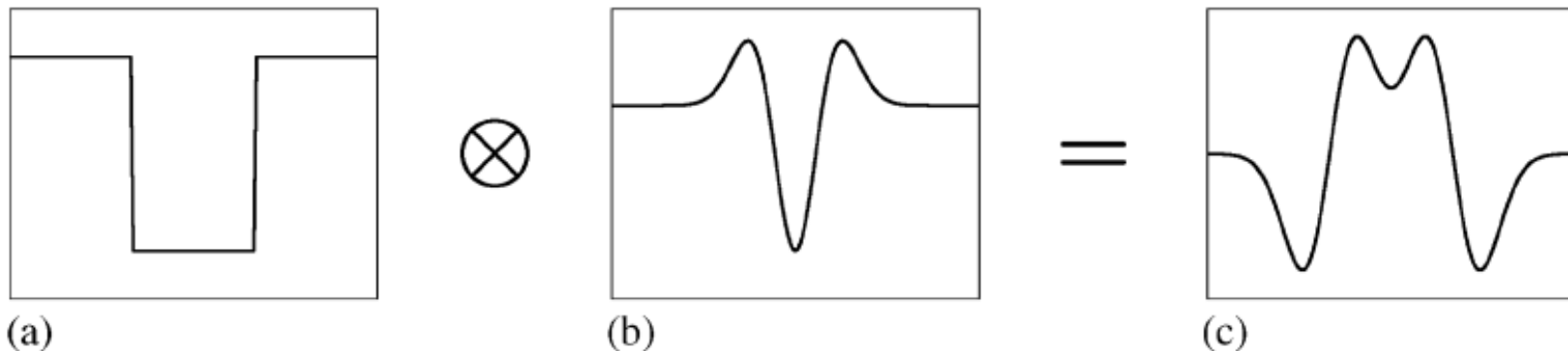
Retina (Netzhaut) Erkennung: Grundlagen

Da der Augenhintergrund beleuchtet werden muss, ist die Sensorik wesentlich aufwändiger als im Iris Fall. Anwendungen sind im Moment auf High Security beschränkt. Besonders charakteristisch ist bei Retina Bildern die Struktur der Adern. Hier ist es jedenfalls wichtig diese Adern gut darzustellen, im Beispiel ein Originalbild mit anschließender lokaler Kontrastverbesserung.



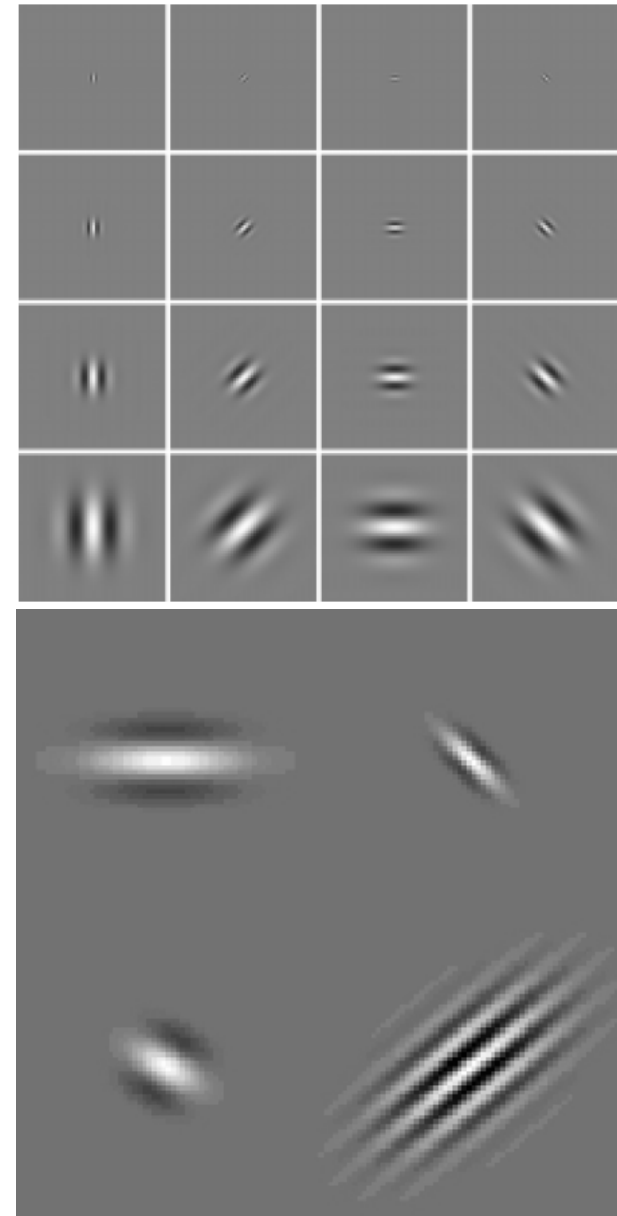
Retina Erkennung: Adernsegmentierung

Ein klassischer Ansatz verwendet Kantenerkennung mit LoG (je nach Kantenbeschaffenheit mit 1 oder 2 Responses) mit entsprechender morphologischer Nachbearbeitung zum "Befüllen" der Adern zwischen ihren Grenzkanten (Closing, wobei Auffüllungseffekte bei Bifurkationspunkten abgefangen werden müssen).
Ergebnisse noch nicht sehr zufriedenstellend.

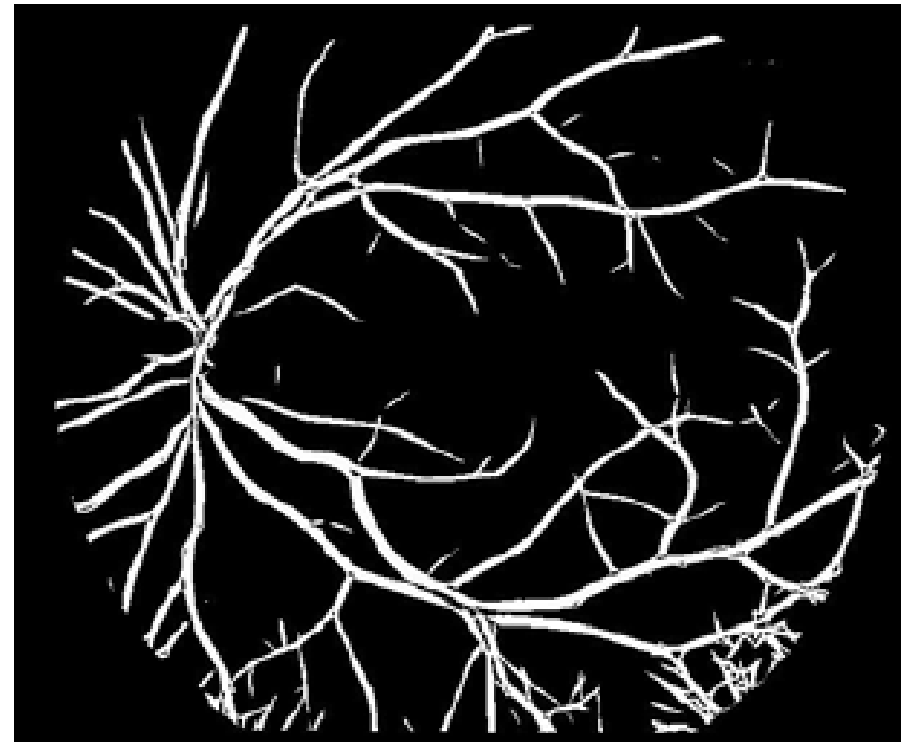
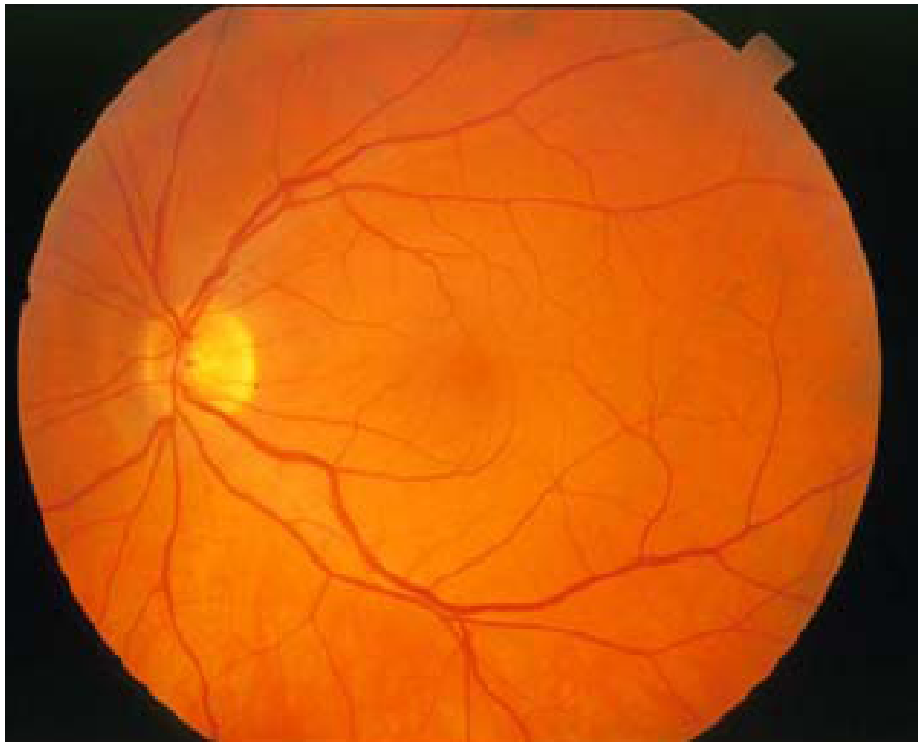


Retina Erkennung: Verbesserte Adernsegmentierung

In diesem Verfahren werden nach der lokalen Kontrastverbesserung sowohl LoG Kantenerkennung als auch Gabor Filterung (hier auch als Morlet-Wavelets bezeichnet) verwendet die zu einem finalen Kantenbild kombiniert werden. Bei der Gabor Filterung werden aus verschiedenen Orientierungen einer Skala die maximalen Filterantworten verwendet, die verschiedenen Skalen werden in einer Art Feature Synthesis kombiniert. Zusätzlich wird adaptives Thresholdung durchgeführt (Schranke verändert sich kontextabhängig).

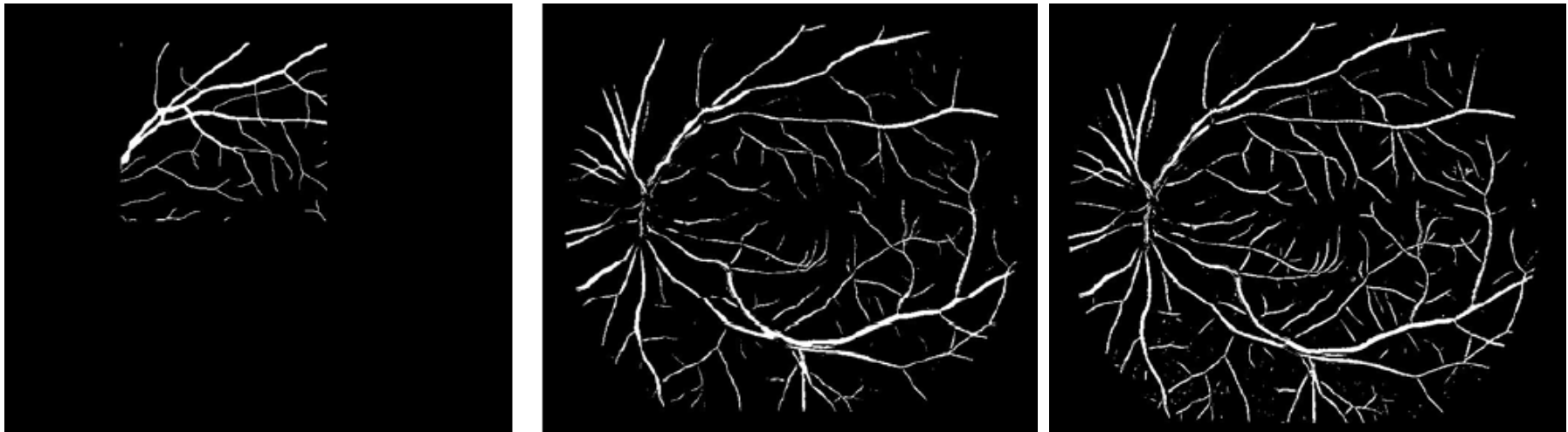


Retina Erkennung: Adererkennung durch Adaptive Thresholding



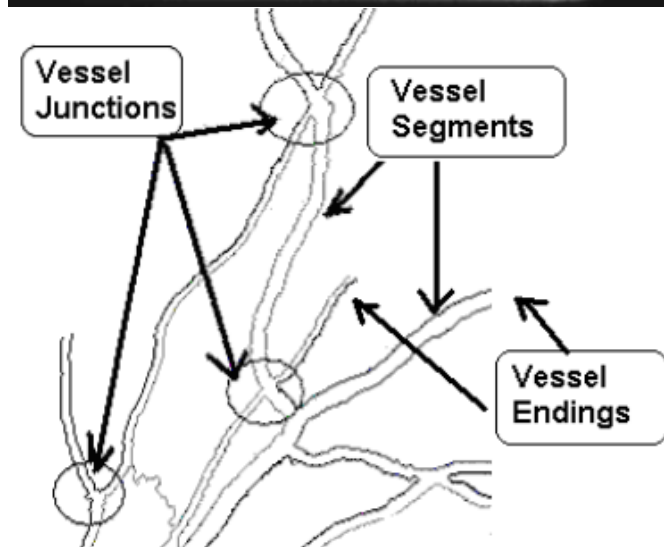
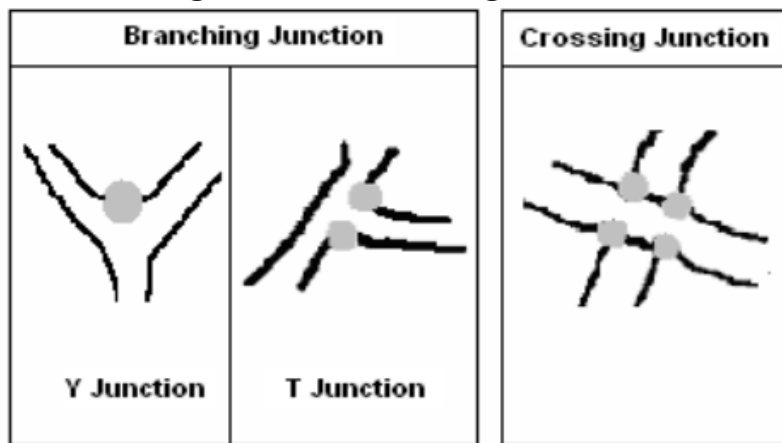
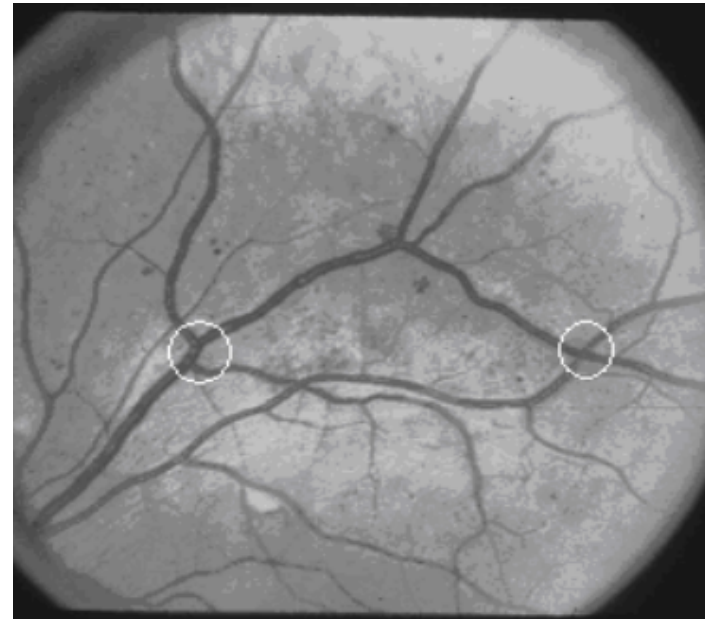
Retina Erkennung: Adernerkenkung durch Klassifikation

Eine Verbesserung der Ergebnisse wird durch Training eines Classifiers erreicht: in einem Teil des Bildes wird durch einen Experten eine "perfekte" Adernkennzeichnung durchgeführt. Mit diesen Daten wird trainiert und die den gelernten Features entsprechende (Features kommen aus dem durch adaptives Thresholding gewonnenen Bildes) Klassifikation angewendet. Es zeigen sich wesentlich bessere Ergebnisse bei der Verwendung eines Bildteiles als beim Training anhand anderer Bilder (leider – optimale Ergebnisse also nur im semi-automatischen Verfahren).



Retina Erkennung: Matching von Adernkreuzungen I

Von jedem Retinabild werden die Koordinaten der Adernkreuzungen als Features verwendet. Es ist per se kein Ausrichten der Bilder nötig, Matching geschieht durch optimales Matching von zwei Punktwolken: "point pattern matching" (hier grosse Ähnlichkeit zu Minutienbasierten FP-Systemen !). Verbesserte Verfahren könnten zusätzlich etwa Kreuzungstypen und die Richtungen der Adern in den Kreuzungen verwenden. Hier gibt es wenig Literatur, im Gegensatz zur Adernerkennung die für medizinische Diagnostik wichtig ist.



[Quit](#)

[Full Screen](#)

[Previous Page](#)

[Next Page](#)

[GoTo Page](#)

[Go Forward](#)

[Go Back](#)

Retina Erkennung: Matching von Adernkreuzungen II

Klassische Vorgehensweise ist Binarisierung (z.B. durch Thresholding des Adernbildes) und anschliessendes Thinning. Anschliessend müssen die Adernkreuzungen erkannt werden: ein Kreuzungspixel hat typischerweise drei oder vier benachbarte Pixel, im Gegensatz zum normalen Adernpixel mit nur zwei Nachbarpixel (immer ein Thinning auf 1 Pixelbreite vorausgesetzt).

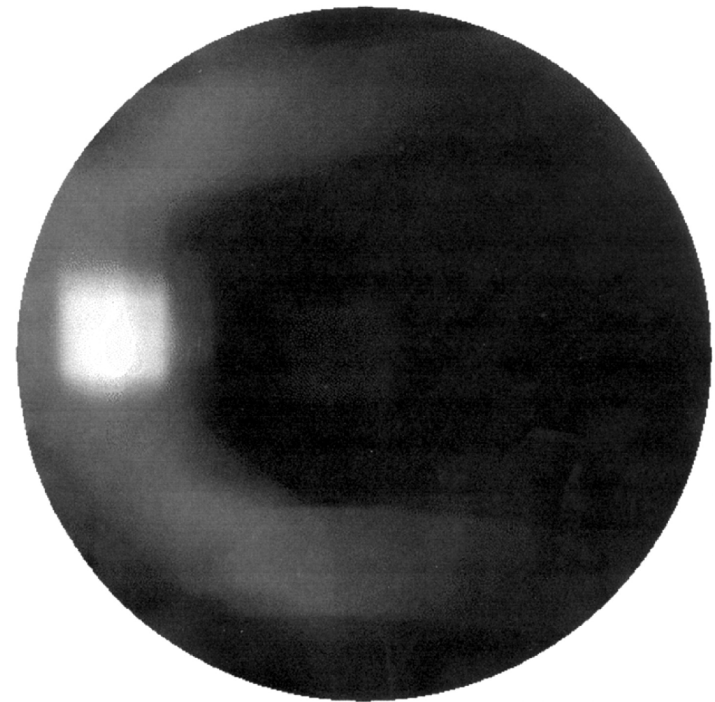
Es gibt kaum Skalierungsprobleme, da der Aufnahmeabstand durch die technischen Gegebenheiten recht fixiert ist. Rotation und Translation müssen allerdings berücksichtigt werden. Um einen Vergleich der Kreuzungen zu ermöglichen, wird im einzigen verfügbaren Paper (!) folgende Bild-Registrierungsprozedur vorgeschlagen: es wird der Schwerpunkt der erkannten Kreuzungen berechnet und die am weitesten vom Schwerpunkt entfernte Kreuzung bestimmt. Der Vergleich von zwei "Kreuzungsmengen" geschieht durch Translation auf identische Schwerpunkte und Rotation auf identische Position der am weitesten entfernten Kreuzung (blöd ist wenn gerade die am weitesten entfernte Kreuzung fehlt – Bildrand !).

Retina Erkennung: Matching von Adernkreuzungen III

Anschliessend wird bewertet, wieviele Kreuzungen übereinstimmen (und ein gewisser Toleranzabstand ist erlaubt): sei N_U und N_V die Anzahl der Kreuzungen in den Bildern U und V und $N_{V|U}$ die Anzahl der Kreuzungen in V die in U ein Pendant finden und $N_{U|V}$ analog (muss nicht identisch sein !). Im folgenden wird die Wahrscheinlichkeit einer Kreuzung im Bild V bestimmt, gegen eine Kreuzung im Bild U gematched zu werden: $P(V|U) = N_{V|U}/N_V$, analog wieder $P(U|V)$. Die Ähnlichkeit S zwischen den beiden Bildern wird dann bestimmt durch: $S = (P(V|U)P(U|V))^{\frac{1}{2}}$. Schwache experimentelle Validierung !

Retina Erkennung: Retina Code

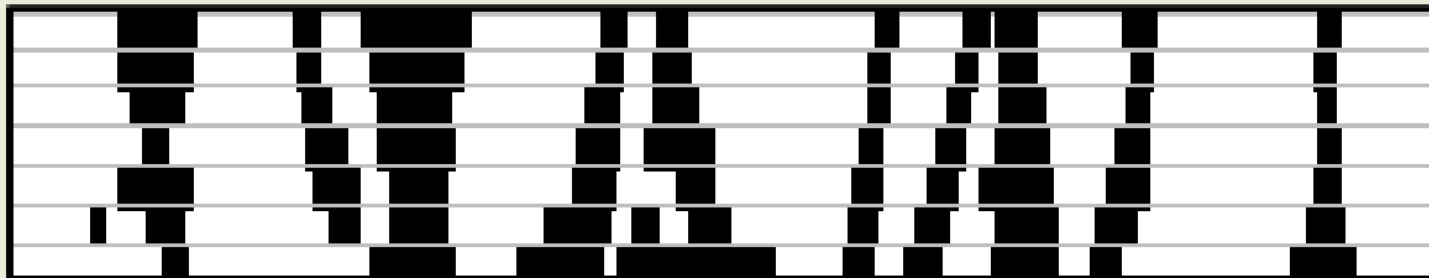
Um einen Code in Analogie zum Iris Code entwickeln zu können, muss ein Referenzpunkt in Analogie zum Mittelpunkt der Pupille definiert werden. Hier bietet sich die “opical disc” an, wo der opische Nerv die Netzhaut verlässt. Diese Scheibe zeichnet sich durch höhere Hintergrundhelligkeit als die Umgebung aus. Zusätzlich weist sie auch eine höhere lokale Varianz auf als der Rest der Retina (durch den hohen Kontrast dunkle Adern gegen helleren Hintergrund (siehe nebenstehendes Bild: lokale Varianz)). Der Mittelpunkt dieser Scheibe kann als Referenzpunkt dienen.



Retina Erkennung: Produkt

Die Firma Retica (<http://www.retica.com/>) bietet ein Retina Standalone Produkt sowie ein Kombinationsprodukt mit Irisscan an. Ausgehend von der optical disc werden konzentrische Kreise betrachtet und entsprechende 1D Signaturen generiert. Vermutlich wird ähnlich wie im Iris Fall eine Mittelung über mehrere angrenzende Kreise und eine Transformation in Polarkoordinaten durchgeführt. Positionen wo sich Adern befinden werden durch umgebungsabhängiges Thersholding mit schwarz (1) kodiert, Hintergrundpositionen mit weiss (0). Durch zirkuläres Shiften beim Matching (Hammingdistanz kann eingesetzt werden) wird Rotationsinvarianz erreicht, Skalierungsinvarianz könnte z.B. durch Normierung des Radius der optical disc erreicht werden.

Multi-Radius Digital Pattern



Fingerabdruck: Grundlagen

Fingerprints sind die am häufigsten verwendete biometrische Modalität, die mit der längsten Historie, die am besten untersucht und auch die mit Abstand höchstem monetären Umsatz (ca. 50% des gesamten Umsatzes im Bereich Biometrie wird mit Fingerprint Systemen gemacht). FPs entstehen im 7. Schwangerschaftsmonat und zeigen bei näherem Verwandtschaftsverhältnis eine grössere Ähnlichkeit.

Durch Massenproduktion sind die Sensoren im niederen Preissegment zu finden (z.B. FP-Maus, FB-USB Stick), Anwendungen gibt es natürlich nach wie vor im forensischen Bereich, im staatlichen Bereich (Grenzkontrollen – Pass, Sozialversicherungen) und stark wachsend im rein kommerziellen Bereich (Zutrittskontrolle – FP Schloss an der Haustür und Zugangskontrolle – PDA, ATM, Handy).

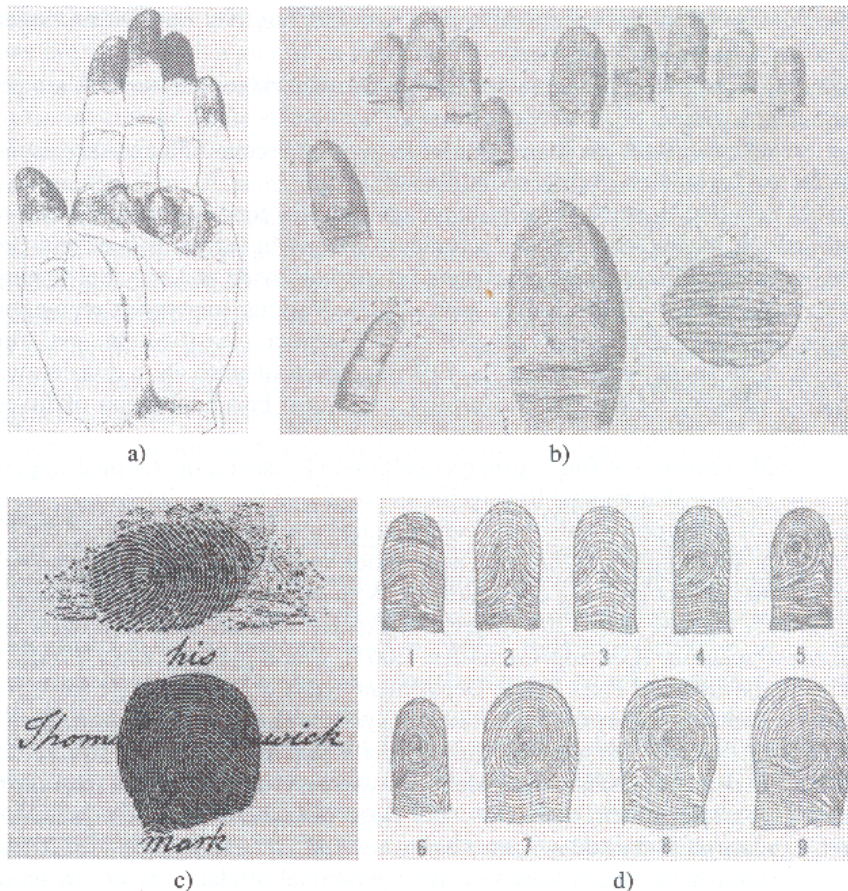
Vorteil ist natürlich die hohe Akzeptanz da ein Finger als deutlich weniger kritisch empfunden wird als z.B. das Auge, auch stehen mehrere Finger zur Verfügung, falls verschiedene Systeme bedient werden sollen oder ein Finger kompromittiert werden sollte. Ebenfalls als Vorteil wird gesehen dass nur mit Zustimmung und aktiver Mithilfe die Daten gewonnen werden können (im Gegensatz z.B. zur Gesichtserkennung)
Nachteil ist die enge Verknüpfung mit kriminaltechnischen Vorgängen, die potentiellen Probleme bei Verletzungen und jede Menge Verfahren zum Überlisten von FP-Verfahren.

Fingerabdruck: Ur-Geschichte



FP Muster auf archäologischen Artefakten, hier z.B. 5000 und 2000 v.Chr. – unklar ist ob die Bedeutung für die Individualität klar war.

Fingerabdruck: Geschichte



1684 erste wiss. Arbeit eines Botanikers über versch. Eigenschaften von FPs (a).
1788 detaillierte anatomische Beschreibungen von FP Eigenschaften (b).

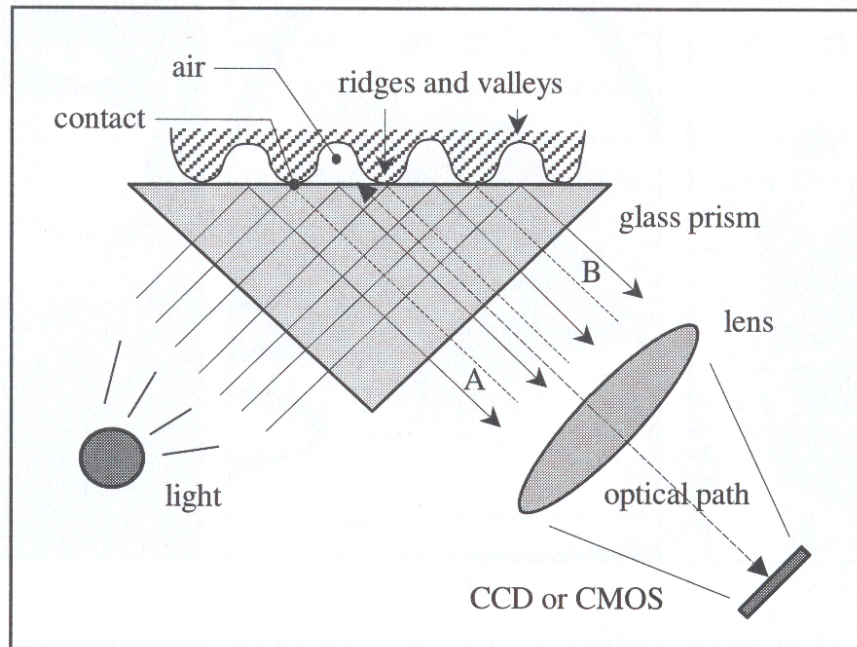
1809 erstmalige Verwendung eines Fingerprints als Trademark durch Thomas Bewick (c). 1823 entstand das erste Verfahren zur Klassifikation von FPs basierend auf den Grundstrukturen der Ridges, 9 Klassen werden hier unterschieden (d). 1888 führt F. Galton den Begriff der Minutien ein und definiert verschiedene Features, 1899 wird das Henry-System zur Klassifikation definiert (das heute im wesentlichen verwendet wird). Endpunkt der historischen Entwicklung ist der Beginn der FBI FP Datenbank mit 810000 FPs, heute über 200 Mio FPs. Das stetige Anwachsen dieser Datenbanken machte schliesslich automatisierte Verarbeitung notwendig.

FP-Sensing: Off-line Aquisition

Obwohl es die ersten FP Sensoren bereits vor 30 Jahren gab, wird auch heute noch z.T. im forensischen Bereich mit dem "Tintenverfahren" gearbeitet: der Finger wird mit Tinte benetzt und gegen einen Karton gepresst. Dieses Verfahren ermöglicht es, den Fingerabdruck von Nagel zu Nagel durch Abrollen abzunehmen was verglichen mit "flachen" FPs mehr Informationsgehalt liefert. Allerdings ist die Abdruckqualität stark von einer gleichmässigen Tintenaufbringung und dem Fingerzustand abhängig (Schweiss, Fett, etc.). Häufig werden hier ten-FP genommen (im Gegensatz zu kommerziellen Applikationen). Die so gewonnenen Abdrücke werden dann mit Scannern oder CCD Kameras digitalisiert um in einem AFIS verwendbar zu sein. Gegenwärtige forensische FP Datenbanken beinhalten daher sowohl wie eben beschrieben gewonnene Off-line FPs als auch On-line FPs, was eine nicht einfache Aufgabe für automatisiertes Matching bedeutet.

Wieder im forensischen Bereich sind die sog. "latenten" FPs eine Besonderheit. Berührt ein Finger ein Objekt, wird ein Feuchtigkeits- bzw. Fettfilm aufgebracht der die Struktur der Ridges wiedergibt. Mit verschiedenen Verfahren werden diese FPs verbessert um sie aufnehmen und verarbeiten zu können (z.B. Puder).

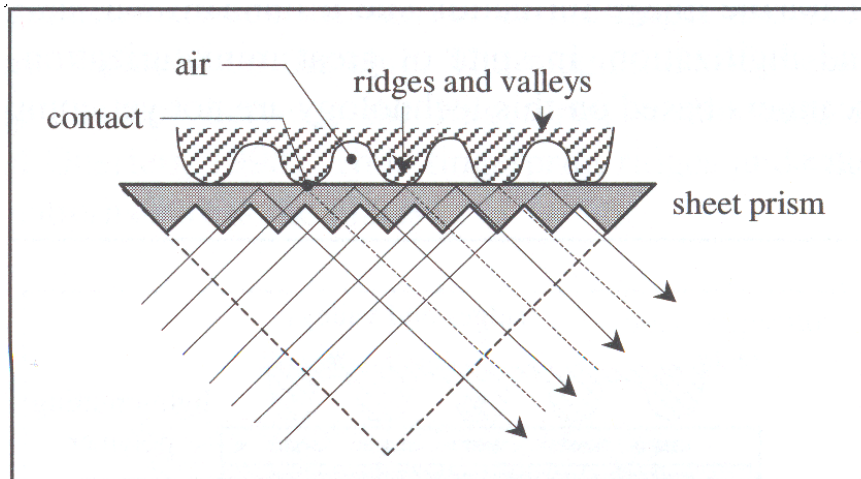
FP-Sensing: On-line Aquisition – Optische Sensoren I



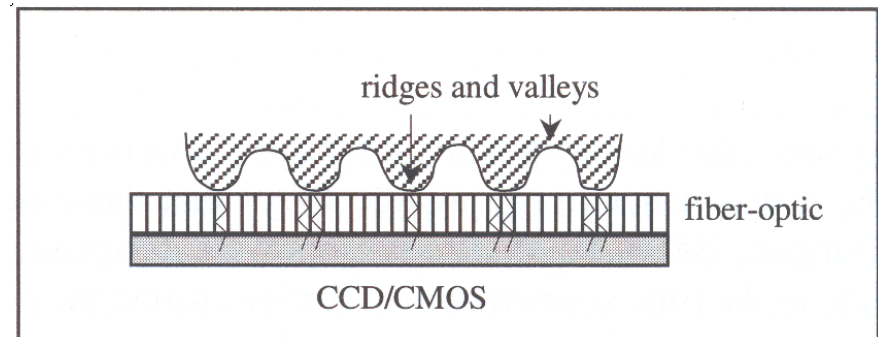
FTIR (Frustrated Total Internal Reflection) ist die älteste und am meisten verwendete Technik. Der Finger berührt ein Glasprisma – während die Ridges das Prisma berühren bleiben die Valleys in einem Abstand. Die linke Seite des Prismas wird beleuchtet (LEDs), das Licht wird an den Ridges absorbiert und an den Valleys reflektiert – so werden die Valleys hell dargestellt, die Ridges dunkel. Das austretende Licht wird durch eine Linse auf einen CCD Sensor fokussiert. Die entstehende Verzerrung muss optisch oder rechnerisch korrigiert werden (A und B sind nicht gleich lang!). Miniaturisierung ist ein Problem bei sehr kleinen geräten – Handys oder PDAs.

FP-Sensing: On-line Aquisition – Optische Sensoren II

FTIR mit einem zusammengesetzten Prisma ("sheet prism") reduzieren die Grösse, allerdings auf Kosten der Qualität.

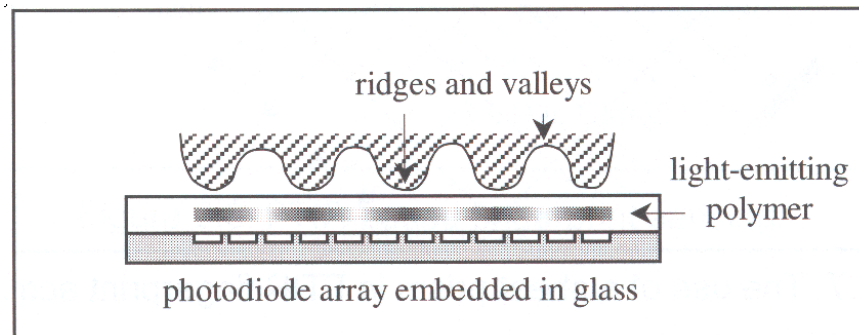


Prisma und Linse können auch durch eine Fiberglasplatte ("optical fiber") und direkt verbundenem CCD Sensor ersetzt werden. Fingerestlicht wird durch das Fiberglas übertragen und aufgenommen (ohne Beleuchtung). Durch rel. grosse CCD Sensorfläche entstehen rel. hohe Kosten.



FP-Sensing: On-line Aquisition – Optische Sensoren III

Elektro-optische Sensoren bestehen aus zwei Schichten: ein Polymer das bei entsprechend angelegter Spannung Licht emittiert, dessen Stärke vom auf einer Seite angelegten Potential abhängt. Da Ridges das Polymer berühren und Valleys nicht, ist das Potential unterschiedlich und damit das emittierte Licht. Die zweite Schicht ist ein Array aus Photodiode die das emittierte Licht in ein digitales Bild umwandeln. Werden sehr klein, Qualität aber schlechter als bei FTIR.

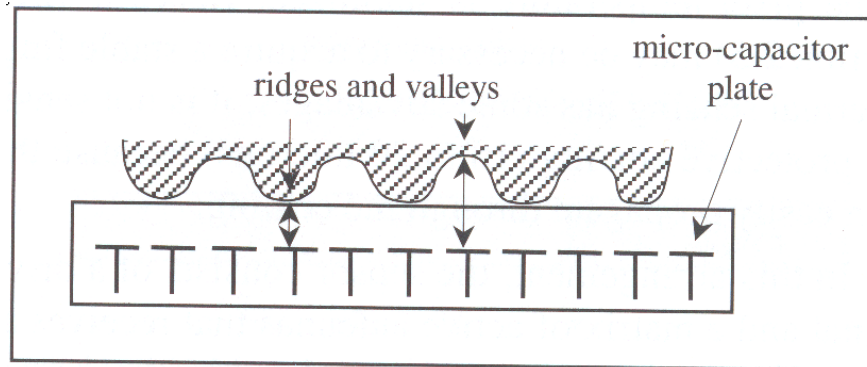


Direkte photographische Aufnahmen leiden meist unter zu geringem Kontrast. Weiters sind solche Verfahren extrem anfällig gegenüber wechselnden Beleuchtungsverhältnissen und können durch Fotos von Fingern oder FPs leicht getäuscht werden.

Weiters muss eine gute Ausrichtung des Fingers gewährleistet werden. Direkte Fotografie wird daher praktisch nicht verwendet.

FP-Sensing: On-line Aquisition – Silikon Sensoren I

Auch bezeichnet als “Solid-state Sensors”. Gibt es erst seit Mitte der 90er; alle Typen bestehen aus einem Array von Pixeln wobei jedes Pixel ein kleiner Sensor ist, der Benutzer berührt direkt die Silikonoberfläche. Es gibt hier verschiedene Varianten, je nachdem wie die physikalische Information in elektrische Signale umgewandelt wird.



Kapazitiv: der Sensor besteht aus in einen Chip eingebetteten Mikro-Kondensatoren, die zweite Kondensatorplatte ist der Finger selbst. Beim Berühren des Sensors entstehen kleine elektrische Ladungen zwischen Finger und Silikonplatte, deren Grösse vom Abstand zur Kondensatorplatte abhängt. Können wie die optischen Sensoren nicht durch ein Bild des Fingers getäuscht werden. Wichtig ist der Schutz der Oberfläche (der nicht zu dick sein darf) und Resistenz geg. elektrostatische Entladungen.

FP-Sensing: On-line Aquisition – Silikon Sensoren II

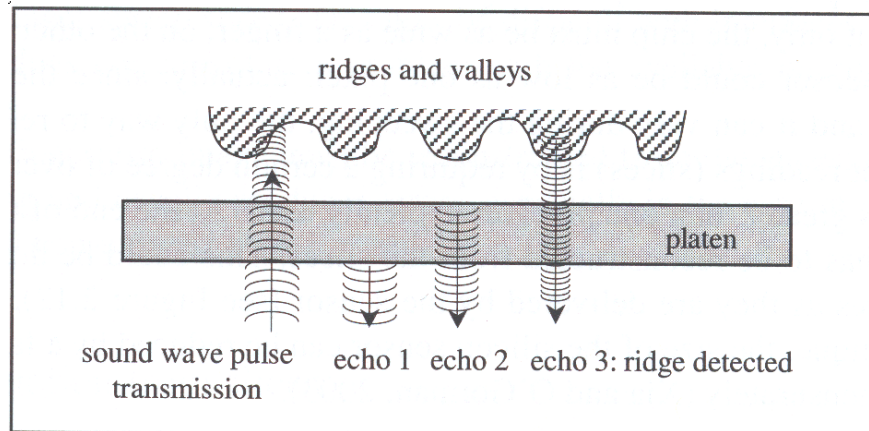
Thermal: Die Sensoren bestehen aus pyro-elektrischem Material das Strom in Abhängigkeit von Temperatur Unterschieden erzeugt. Die Sensoren werden typischerweise warm gehalten um einen grossen Unterschied zur Ridge Temperatur zu gewährleisten und werden im Bereich der Ridges abgekühlt. Das erzeugte Bild verschwindet allerdings durch den Temperatúrausgleich rasch, sodass eine dynamische Aufnahmemethode (“sweeping”) von Vorteil ist. Positiv ist die Robustheit geg. Entladungen und weniger Empfindlichkeit bei stärkerer Schutzschicht.

Piezoelektrisch: Die Sensoren sind hier druckempfindlich und erzeugen ein elektrisches Signal in Abhängigkeit vom angelegten Druck. Ridges geben mehr Druck – die verfügbaren Materialien sind (noch) zu wenig sensitiv, das Schutzschicht Problem ist nicht gelöst und Ergebnis ist nur ein binäres Bild.

Elekrisches Feld: der Sensor legt ein elektrisches Feld an das durch die Stuktur der Hauptoberfläche beeinflusst wird (was dan gemessen werden kann). Ebenfalls anfällig geg. elektrostatische Entladungen.

FP-Sensing: On-line Aquisition – Ultraschall Sensoren

Bei Ultraschall werden akustische Signale ausgesendet, die an den Ridge und Valley Strukturen reflektiert werden. Die reflektierten Wellen werden von einem Sensor erfasst und durch die unterschiedlichen Zeitpunkte des Empfangens kann der zurückgelegte Weg bestimmt werden. Wellen die an Ridges reflektiert werden haben offenbar einen kürzeren Weg hinter sich.



Die Qualität der erhaltenen Bilder ist sehr gut, auch können die Aufnahmen durch Handschuhe, Schmutz etc. nicht beeinträchtigt werden.

Problematisch ist jedoch die Grösse und der Preis eines solchen Systems, die Aufnahmezeit beträgt einige Sekunden – insgesamt ist die Technologie noch nicht reif für den praktischen Einsatz.

FP-Sensing: Touch

Das einfache Legen des Fingers auf die Sensorfläche (obwohl hier kein Benutzertraining erforderlich ist) hat einige Nachteile:

- Der Sensor kann schnell verschmutzen was die Benutzerakzeptanz und die Aufnahmegenauigkeit verringert.
- Ein latenter FP bleibt auf dem Sensor zurück der u.U. “gestohlen” werden kann.
- Der Finger wird ev. rotiert auf den Sensor gelegt, was je nach Matching Verfahren problematisch sein kann.
- Die Kosten des Sensors sind relativ hoch und durch die Grösse steigt auch die Anzahl der fehlerhaften Chips.

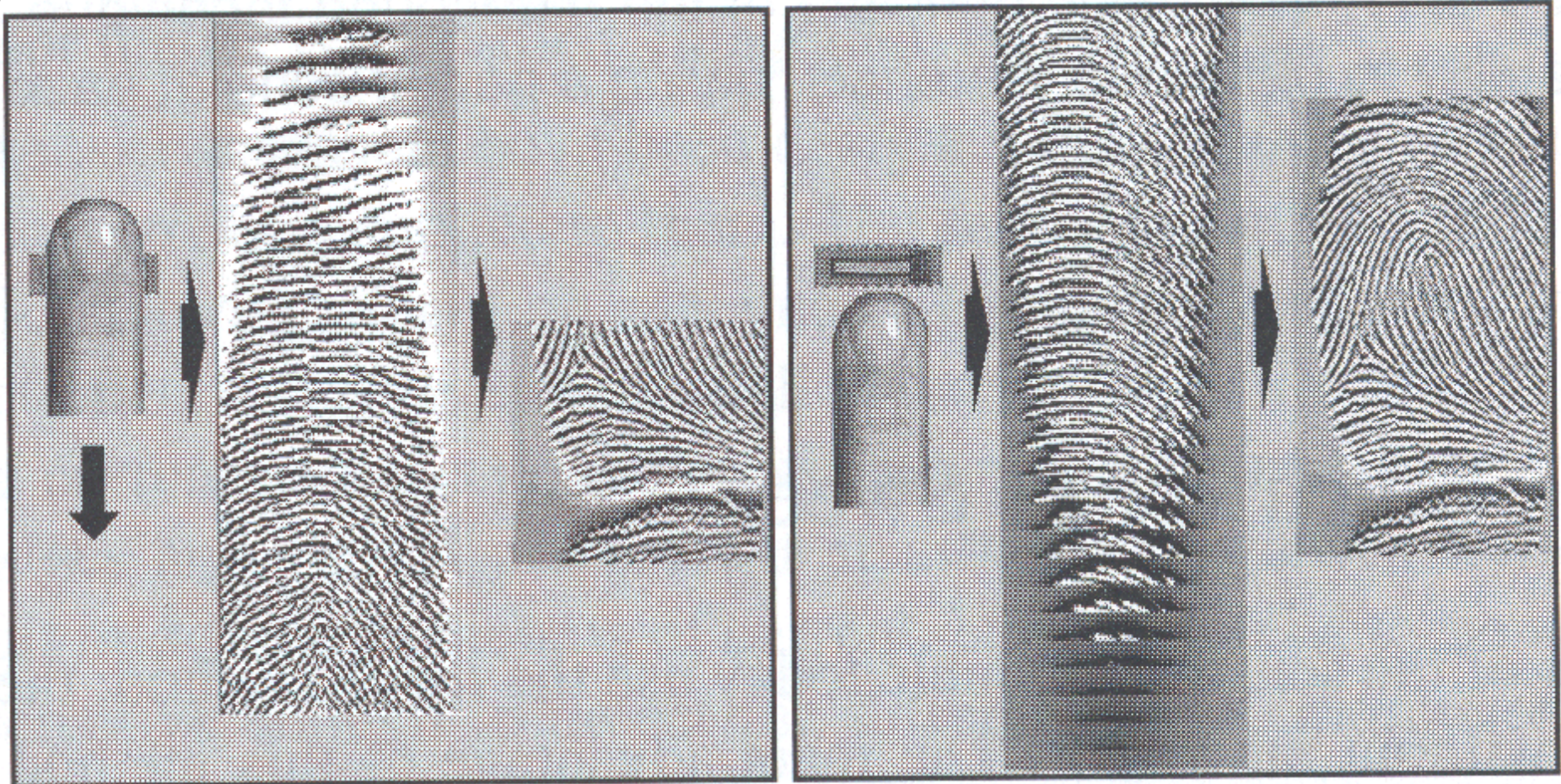
FP-Sensing: Sweep

Als Alternative gibt es aus diesen Gründen Sensoren, über die der Finger gezogen wird – durch die vertikale Bewegung muss der Sensor nur so breit sein wie ein Finger. Die Höhe muss mehrere Pixel betragen, da sonst eine robuste Kombination der aufgenommenen Streifen nicht möglich ist. Ursprünglich für thermale Sensoren entwickelt, da hier das “Liegenlassen” zu Qualitätsproblemen führt.

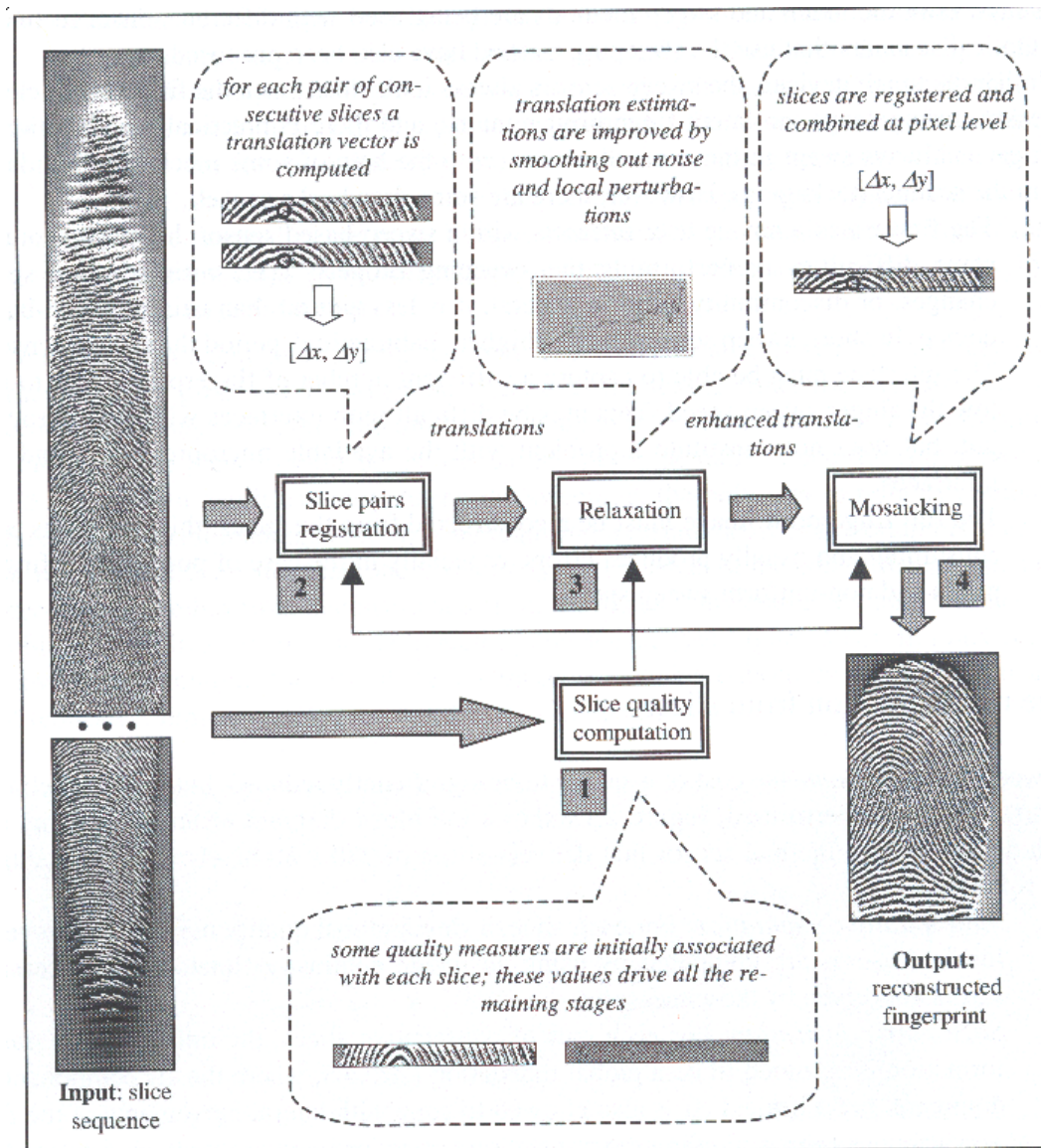
Neben den vorringerten Kosten, besserer Sauberkeit durch die stetige Reinigung beim Drüberziehen und Vermeidung der Rotationsproblematik gibt es aber auch Nachteile:

- Benutzer benötigen eine Einlernphase um gleichmässiges sweeping mit der richtigen Geschwindigkeit durchführen zu können
- Der Sensor muss schnell genug arbeiten um der Geschwindigkeit des Darüberziehens folgen zu können.
- Das FP-Bild muss aus den Slices rekonstruiert werden was Zeit kostet und Fehler einführen kann.

FP-Sensing: Sweep Visualisierung

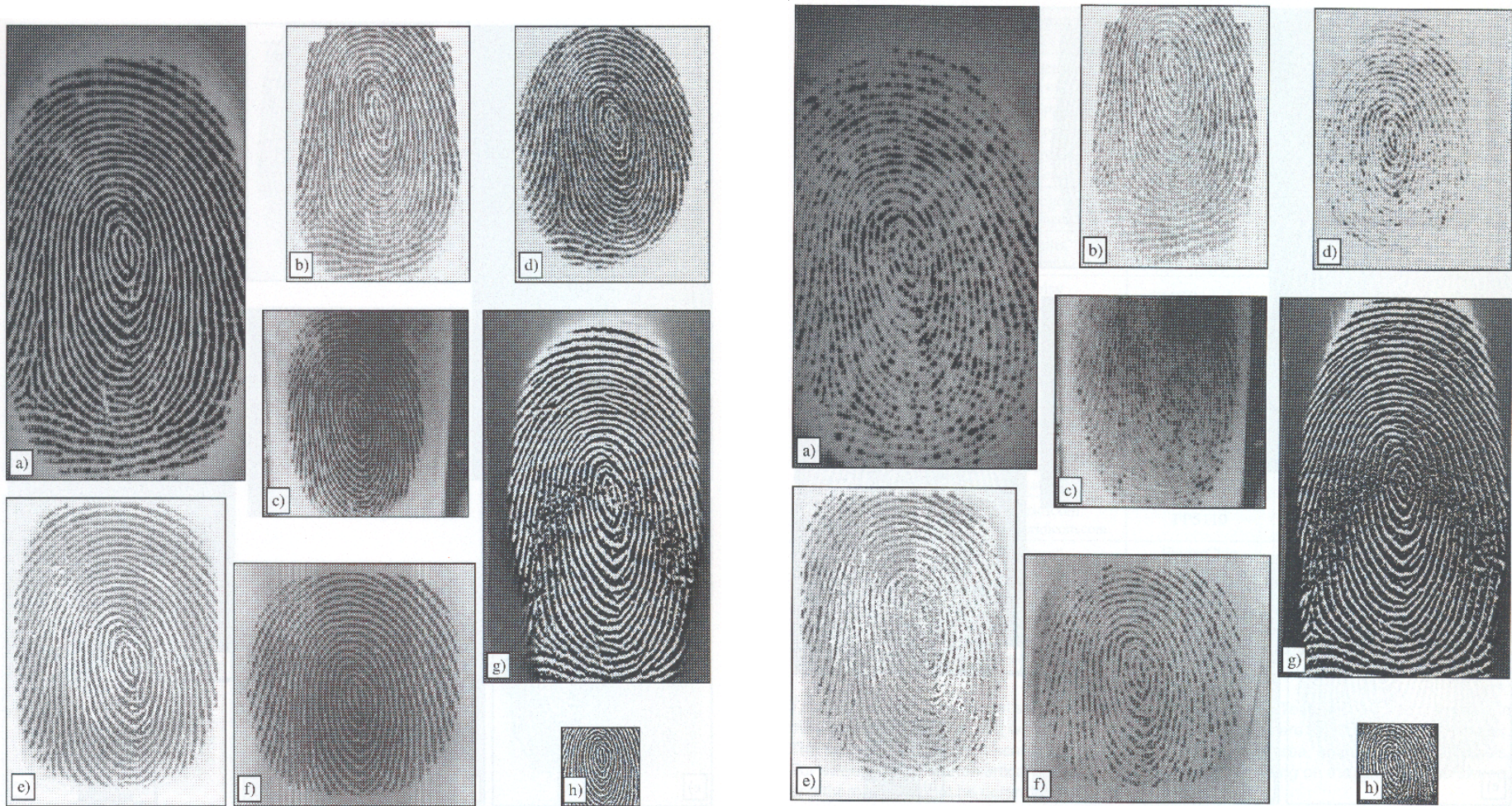


FP-Sensing: Bildrekonstruktion aus FP-Slices



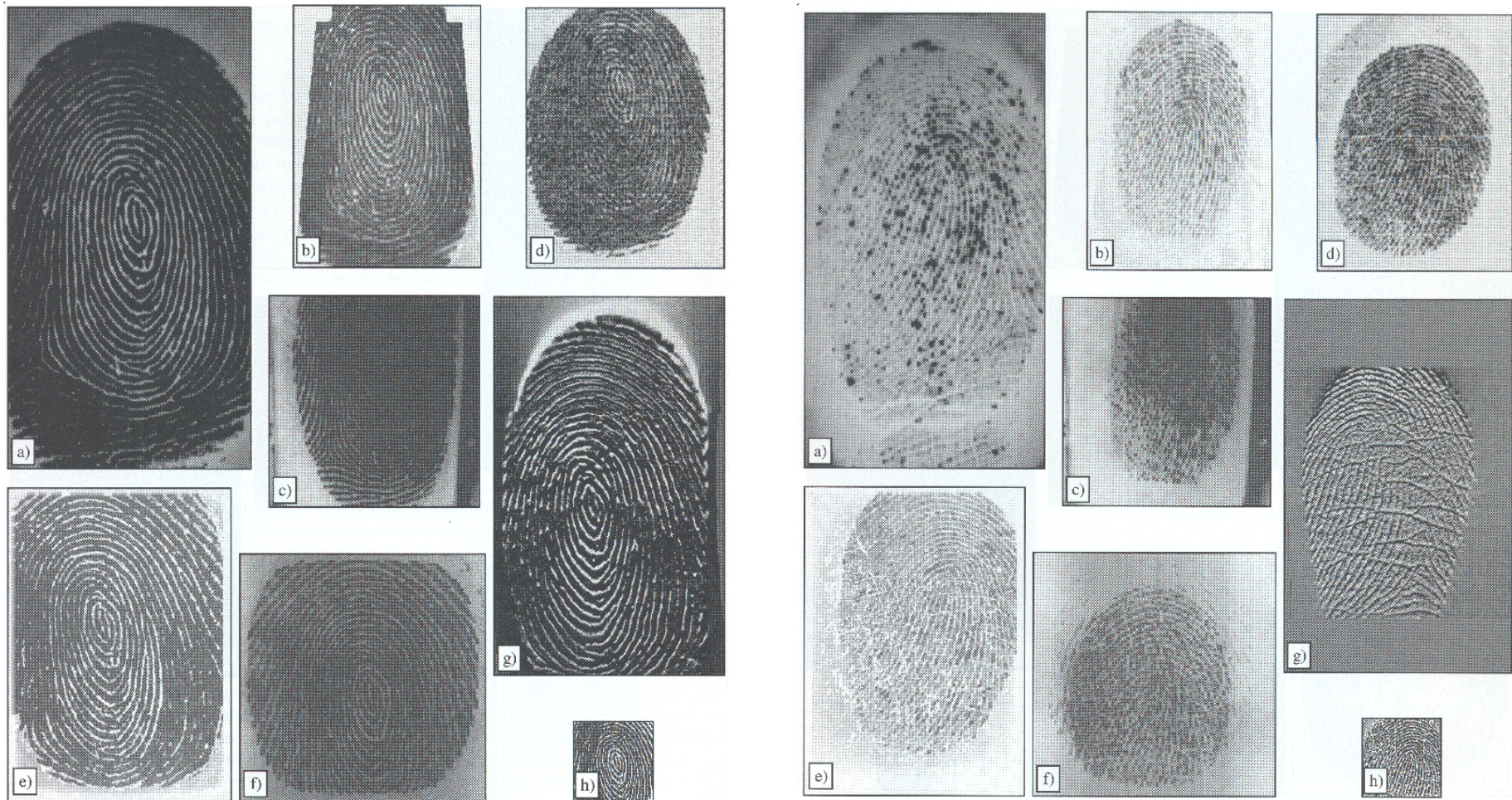
Typisch: Slices mit 280 x 30 Pixel; nach der Bestimmung der Qualität jedes Slice wird eine Registrierung durchgeführt, d.h. es wird der globale Translationsvektor zwischen zwei aufeinanderfolgenden Slices durch vollständige Suche bestimmt, auch horizontale Translation ist in geringem Ausmass erlaubt. Durch Stetigkeitsannahmen bzgl. der Bewegung (und damit der Translationsvektoren) werden Ausreisser korrigiert. Schlussendlich wird dann der gesamte FP durch eine gewichtete Summe der Pixel in den einzelnen Slices gebildet.

FP-Sensing: Visuelle Beispiele I – normal vs. trocken



a) FTIR b) FTIR c) Sheet prism d) Elektro-optisch e) Kapazitiv f) Kapazitiv g) Thermal (sweep) h) Electric field

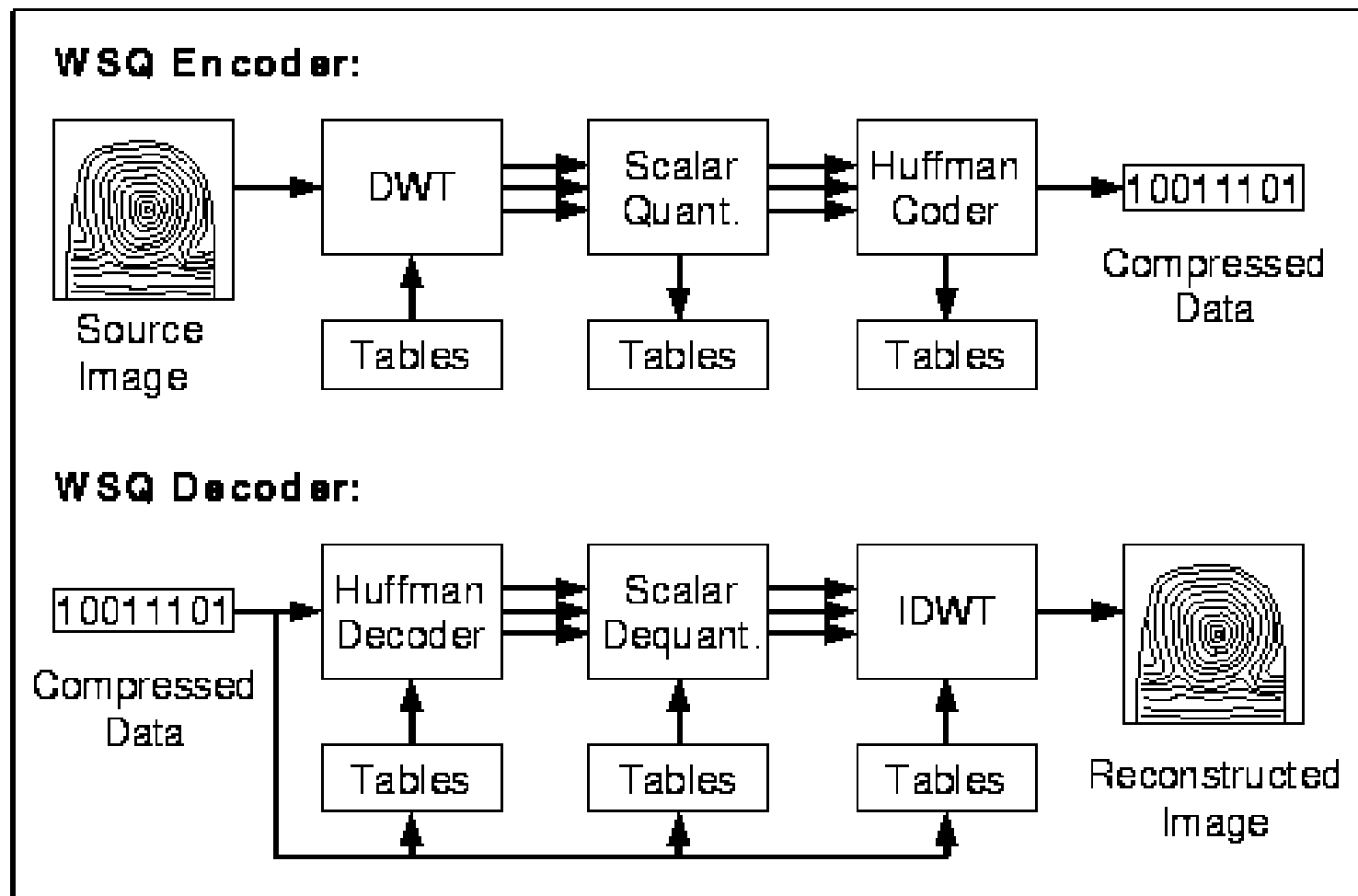
FP-Sensing: Visuelle Beispiele II – nass vs. low quality



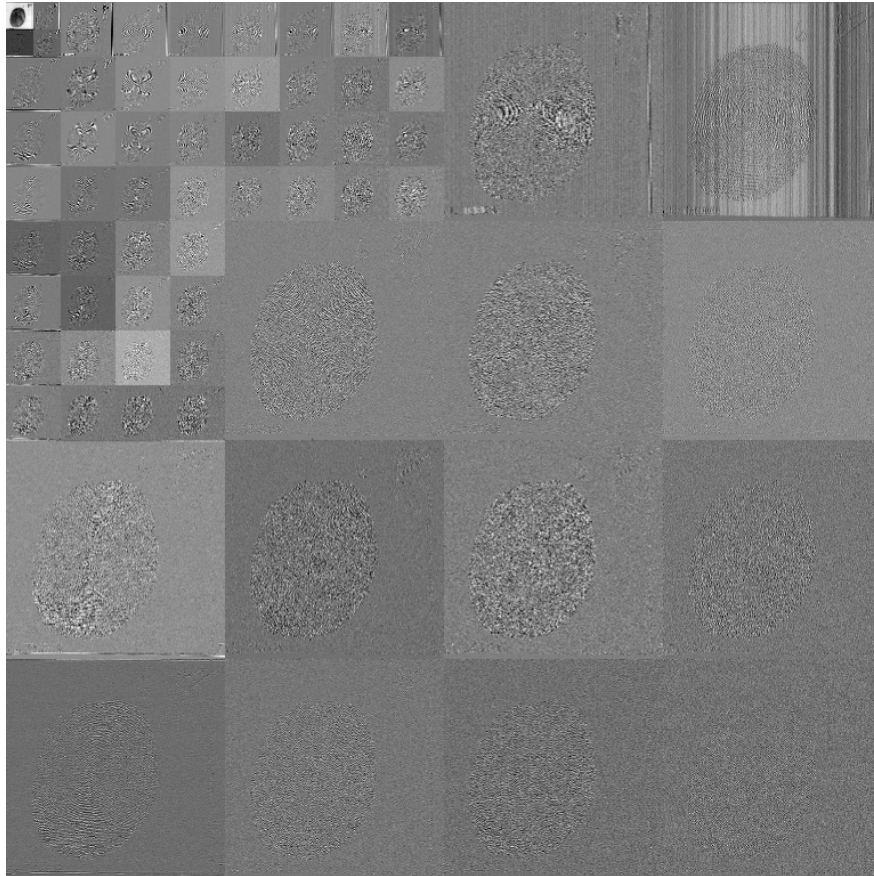
a) FTIR b) FTIR c) Sheet prism d) Elektro-optisch e) Kapazitiv f) Kapazitiv g) Thermal (sweep) h) Electric field

FP - Kompression: FBI WSQ Standard I

ISO/IEC 19794-4 standard on Biometric Data Interchange Formats definiert für FPs drei Varianten die Sensordaten verlustbehaftet zu speichern: JPEG, JPEG2000 und **Wavelet Scalar Quantization (WSQ)**.



FP - Kompression: FBI WSQ Standard II



WSQ ist ein Wavelet-basierter Koder der ersten Generation, d.h. keine Inter-Subband Korrelationen werden ausgenutzt. Besonderheit ist eine festgelegte Wavelet Packet Subbandstruktur die signifikant von der standard pyramidalen DWT abweicht und eine höhere Frequenzauflösung im mittleren Frequenzbereich ergibt. Der Grund ist der durch die Ridge-Valleystruktur stark vorhandene mittlere Frequenzbereich.

Analoge Subbandstrukturen wurden auch für Wavelet-Koder der zweiten Generation verwendet mit gutem Erfolg. Auch JPEG2000 Teil 2 erlaubt solche Strukturen, Ergebnisse siehe PS ! Auch für JPEG Kompression ist ein Tuning der Quantisierungsmatrix ev. gewinnbringend.

FP - Kompression: WSQ vs. JPEG (Ratio 19)



Interessant ist in diesem Zusammenhang die Frage, inwieweit Qualitätsmessungen in PSNR und visueller Eindruck überhaupt Einfluss auf das Matchingverhalten von AFIS haben. Dazu gibt es praktisch keine Literatur !

FP - Qualität

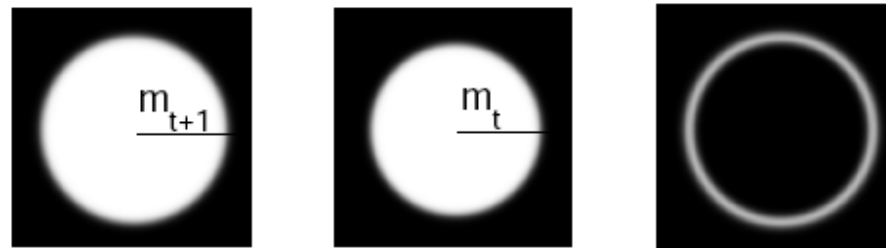
In neueren Untersuchungen wurde gezeigt dass FP Qualität ein ganz entscheidender Faktor bei der Matchinggenauigkeit von FP Systemen ist. Eine automatisierte Beurteilung der Qualität ist wichtig bei der Abnahme des FP (ob der Sensor-Vorgang wiederholt werden muss), bei der Beurteilung ob Enhancement Verfahren sich positiv auswirken und auch bei der Frage welches Verfahren bei welcher Qualität angewendet wird (Hintergrund: die besten Verfahren bei guter FP Qualität sind nicht notwendigerweise die besten bei niederer Qualität).

Im Gegensatz zu gewöhnlichen Bildern gibt es bei FPs spezielle Features, deren Ausprägung für eine Qualitätsbeurteilung herangezogen werden kann, unspezifische Masse wie Glattheit sind hier vermutlich schlechter. Auch zur Beurteilung von komprimierten FPs könnten solche Masse herangezogen werden, Korrelation zu PSNR dann von Interesse.

Neben den behandelten Methoden werden auch diverse Energiebelegungen der WSQ subbands als Qualität vorgeschlagen, diese Themen sind aktuellstes Forschungsgebiet.

Globale FP - Qualität I

Idee ist hier, das DFT eines FP zu analysieren. Ein FP mit guter Qualität sollte eine hohe Energiekonzentration im Bereich der dominierenden Ridge-Frequenz aufweisen. Um dies zu untersuchen, werden wiederholte Bandpassfilterungen durchgeführt und der entsprechende Energiegehalt der Frequenzbänder wird aufgezeichnet. Die Bandpassfilterung kann durch die Differenz von zwei Butterworth-LPF Masken im Frequenzbereich realisiert werden.



Die Energiekonzentration wird dann gemessen durch die Entropie (die minimal ist bei gleichmässiger Verteilung und wächst bei ge-peakter Verteilung), deren Abweichung vom Minimalwert gemessen wird, das ganze auf $[0, 1]$ normiert. Die FPs auf der folgenden Seite haben $Q = 1.0, 0.6, 0.3$. Kritik: da eigentlich nicht eine beliebige Abweichung der Energieverteilung von der Gleichverteilung gesucht wird, könnte der Peak im interessanten Frequenzbereich wesentlich expliziter modelliert werden.

Globale FP - Qualität II



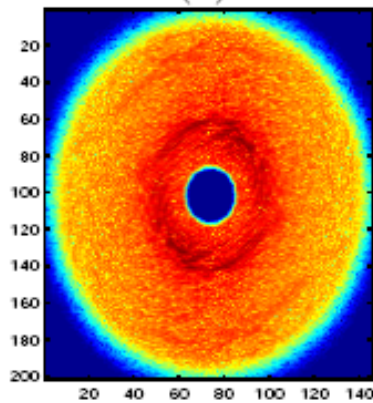
(a)



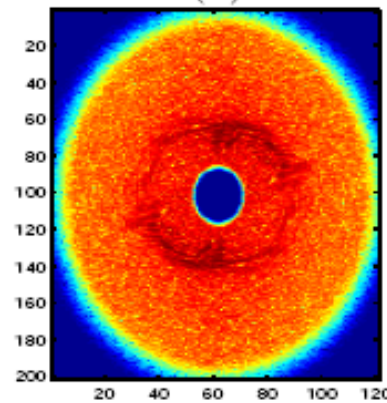
(b)



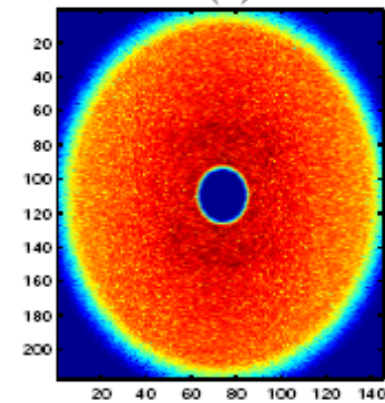
(c)



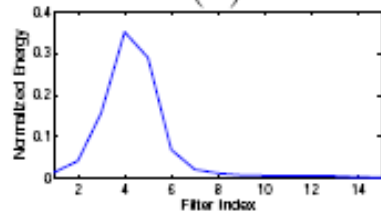
(d)



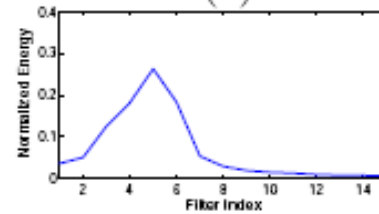
(e)



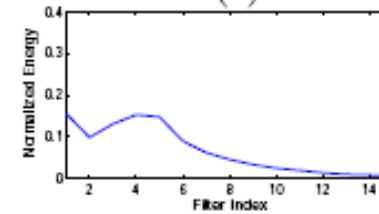
(f)



(g)



(h)



(i)

Lokale FP - Qualität I

Das Bild wird in $b \times b$ Pixel Blöcke aufgeteilt. Für jeden dieser Blöcke ist $g_s = (g_s^x, g_s^y)$ der Gradient an der Stelle s . Die 2×2 Kovarianz Matrix der Gradienten aller b^2 Positionen pro Block ist gegeben durch

$$J_{i,j} = \frac{1}{b^2} \sum_{s \in B} g_s g_s^T$$

Die so gewonnene symmetrische Matrix hat die Eigenwerte

$$\lambda_1 = \frac{1}{2}(\text{trace}(J) + (\text{trace}^2(J) - 4\det(J))^{\frac{1}{2}}) \text{ und}$$

$\lambda_2 = \frac{1}{2}(\text{trace}(J) - (\text{trace}^2(J) - 4\det(J))^{\frac{1}{2}})$, wobei $\text{trace}(J) = J_{1,1} + J_{2,2}$ und $\det(J) = J_{1,1}J_{2,2} - J_{1,2}^2$, sowie $\lambda_1 \geq \lambda_2$. Ein normiertes Kohärenzmass $0 \leq k \leq 1$ ist definiert als

$$k = \frac{(\lambda_1 - \lambda_2)^2}{(\lambda_1 + \lambda_2)^2}$$

das die Deutlichkeit der Ridge-Valley Orientierung in jedem Block wiedergibt (dann ist $\lambda_1 \gg \lambda_2$ und k ist nahe 1).

Ein overall Qualitätsmass wird durch eine gewichtete Summe über alle Blöcke erreicht, in der zentrumsnahe Blöcke stärker gewichtet werden.

Lokale FP - Qualität II



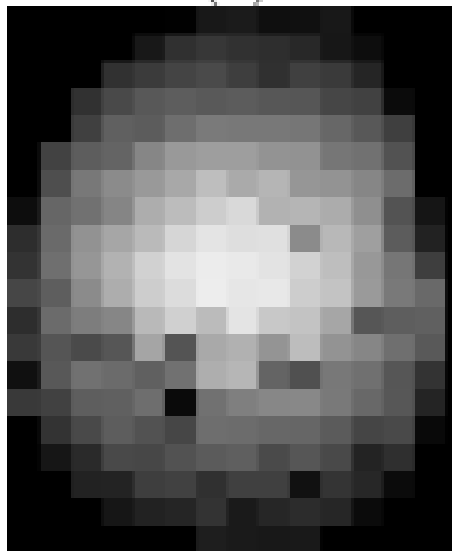
(a)



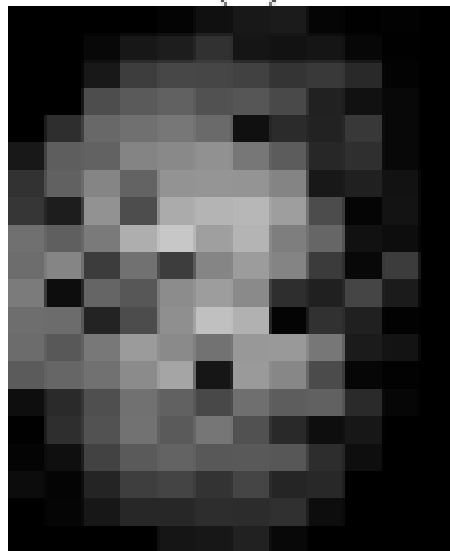
(b)



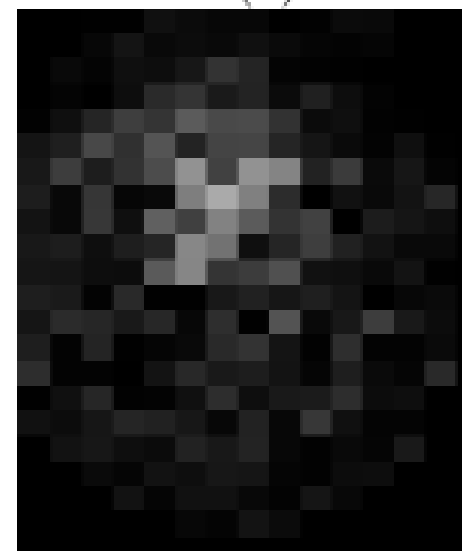
(c)



(d)

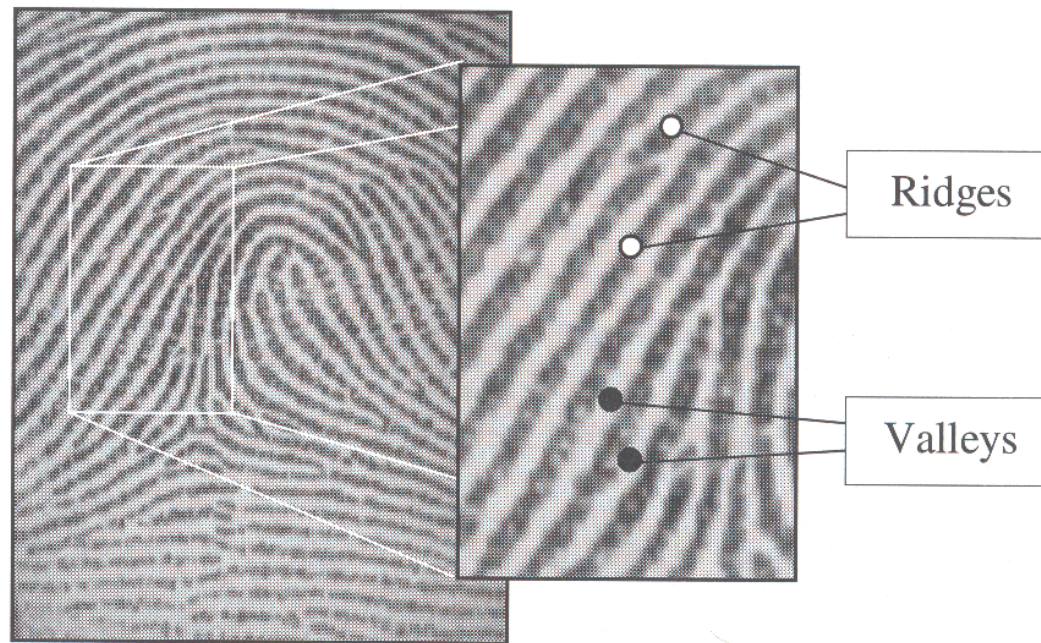


(e)



(f)

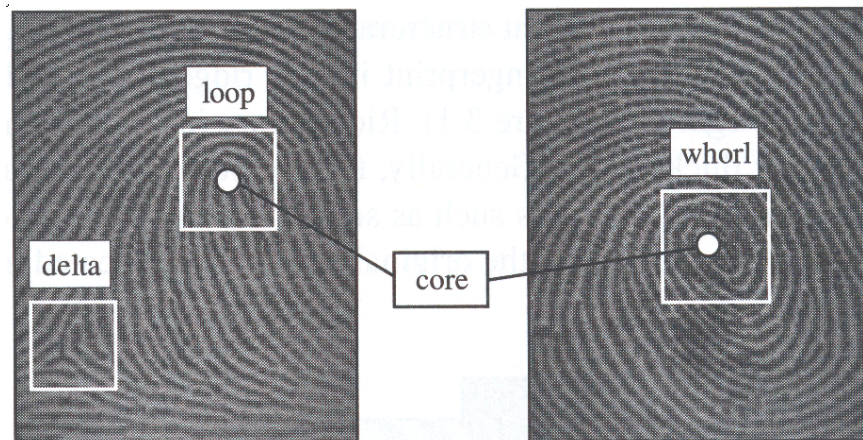
FP Features: Ridges and Valleys



Ridge Linien sind meist dunkel und Valleys hell, Ridges sind 100 - 300 Micrometer breit. Oberflächliche Verletzungen wie Schnitte oder leichte Verbrennungen verändern dieses Muster nicht da es durch nachwachsende Haut dupliziert wird. Ridges und Valleys laufen meist parallel, teilen sich und enden auch manchmal.

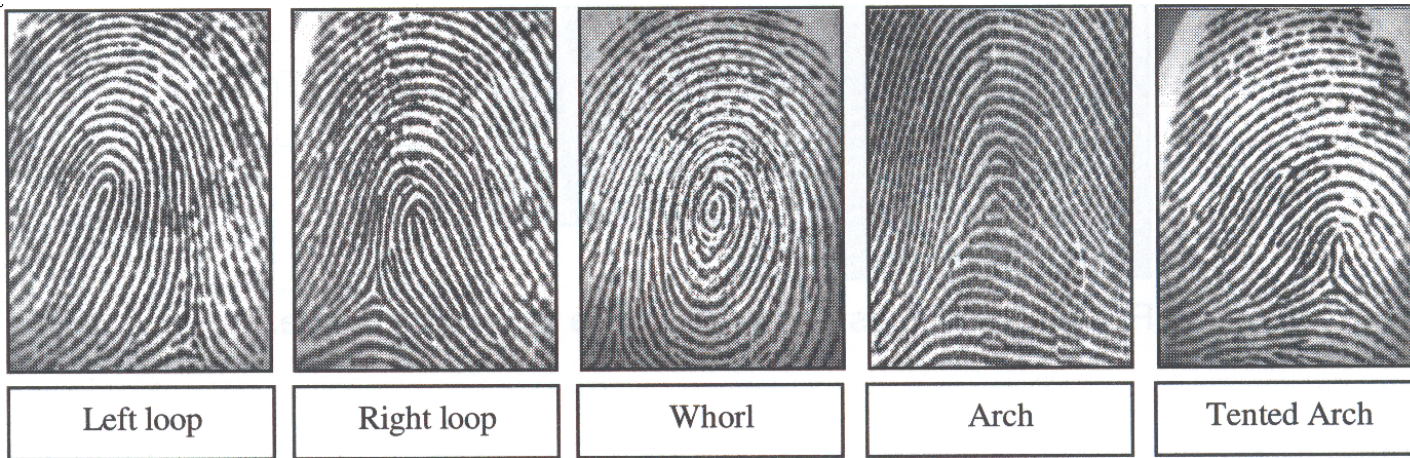
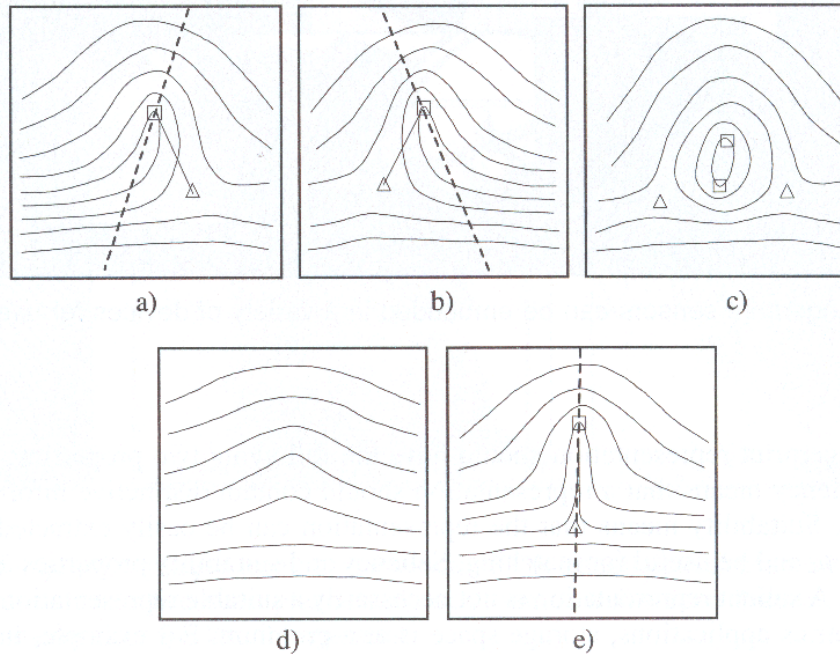
FP Features: Singuläre Punkte

Betrachtet man Ridges und Valley auf globaler Ebene gibt es eine oder mehrere Regionen in denen die Ridges bestimmte Formen annehmen (hohe Krümmung, häufiges Enden etc.) – diese singulären Regionen können in drei Typen klassifiziert werden: loop, delta, and whorl. Eine Sonderstellung nimmt der sog. “core” ein, der in der Klassifikation nach Henry dem Zentrum der “nördlichsten” Singularität entspricht.



Gibt es keine Singularität mit Zentrum, wird der Punkt mit maximaler Ridge Krümmung als core definiert. Der core ist ein wichtiger Punkt für FP Registrierung, wenn für Matching zwei FPs gegeneinander optimal ausgerichtet werden sollen (vgl. Zentrum der Pupille bei Iris und Optical Disk bei Retina). Verschiedene Singularitätstypen werden für die Klassifikation von FPs verwendet (z.B. um eine Suche durch Beschränkung auf eine Klasse beschleunigen zu können).


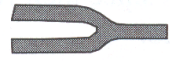


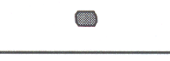

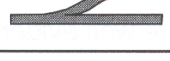
FP Features: Henry-Galton Klassifikation

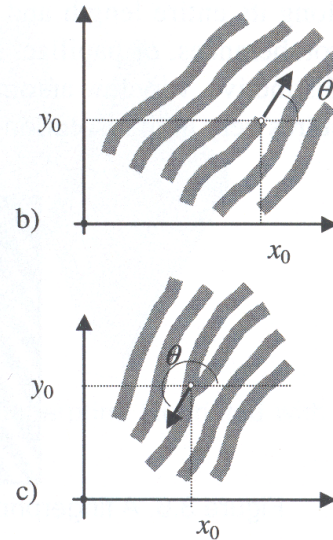


FP Features: Minutien

Minutia bedeutet kleines Detail – Im FP Kontext sind das Unstetigkeitsstellen der Ridges: Endpunkte, Bifurkationspunkte, Trifurkation und unbestimmter Typus (ANSI Typisierung), FBI verwendet nur Endpunkte und Bifurkationen. Galton entdeckte dass die Minutien sich in der Lebenszeit nicht verändern.

a)

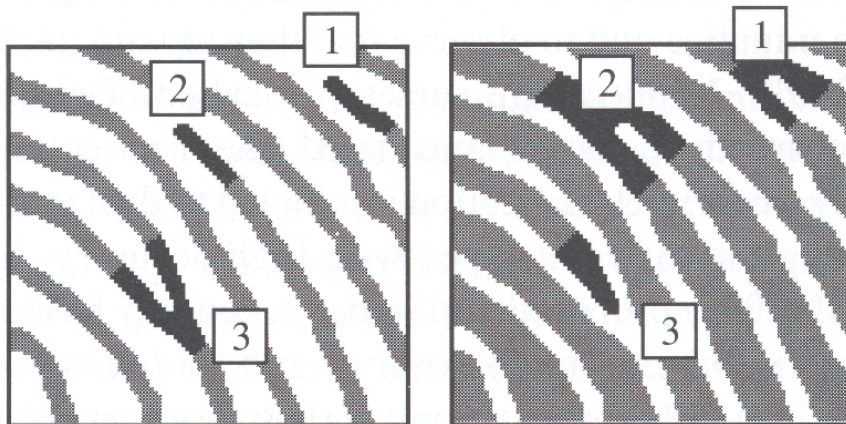
	Termination
	Bifurcation
	Lake
	Independent ridge
	Point or island
	Spur
	Crossover



Zusätzlich zu den Minutienkoordinaten wird auch häufig der Winkel zwischen Ridge-Tangente und der horizontalen Achse bestimmt, im Fall von Bifurkationen wird der analoge Winkel zwischen endendem Valley und der Achse genommen.

FP Features: Minutien und Sweat Pores

In der Praxis kann je nach Fingerdruck auf dem Sensor ein Endpunkt in Wahrheit Bifurkation sein und umgekehrt. Zusätzlich gibt es die sog. termination/bifurcation duality: betrachtet man Originalbild und dessen Negativ, sind Endpunkte im Original Bifurkationspunkte im Negativbild und umgekehrt.



Werden FPs mit hoher Auflösung abgenommen, können auf den Ridges die Schweissporen identifiziert werden (60-250 Micrometer); Obwohl deren Position, Anzahl und Form ein zur Unterscheidung geeignetes Merkmal wäre wird das kaum angewendet, da besonders hohe Auflösung und gute FP Qualität notwendig sind.

FP Features: Lokale Ridge Richtung I

Der Winkel $\theta_{x,y}$ an der Position (x, y) im FP Bild ist gewöhnlich definiert als Winkel zwischen Ridgetangente und horizontaler Achse, er wird im Bereich $[0, 180]$ bestimmt, da es keine Flussrichtung gibt. Die Ridge Orientierung wird nicht für jedes FP-pixel bestimmt sondern nur an diskreten Punkten. Das sog. "Orientierungsbild" zeigt pro Bildblock eine durchschnittliche Ridge Orientierung, manchmal wird auch ein zusätzlicher Zuverlässigkeitswert mitbestimmt (siehe lokale Qualität !).

Der Gradient $\Delta(x_i, y_i)$ im Punkt (x_i, y_i) ist ein zweidimensionaler Vektor mit den Komponenten Δ_x und Δ_y . Der Winkel θ_{x_i, y_j} ist orthogonal zum Gradienten ($\theta_{x_i, y_j} = \arctan \frac{\Delta_y}{\Delta_x}$).

Diese Methode hat aber Probleme bei 90 Grad (durch den Nenner im Bruch) und Schwierigkeiten bei der Berechnung von durchschnittlichen Winkeln: Durchschnitt von 5 und 175 Grad ist nicht 90 sondern 0 Grad, der Durchschnitt von 0 und 90 Grad könnte 45 oder 135 Grad sein.

FP Features: Lokale Ridge Richtung II

Basierend auf der Idee alle involvierten Winkel zu verdoppeln, wird zur Berechnung von θ_{x_i, y_j} in einem 17×17 Fenster folgender Ausdruck vorgeschlagen:

$$\theta_{x_i, y_j} = 90 + 1/2 \arctan \left(\frac{2G_{xy}}{G_{xx} - G_{yy}} \right)$$

wobei

$$G_{xy} = \sum_{h=-8}^8 \sum_{k=-8}^8 \Delta_x(x_i + h, y_j + k) \Delta_y(x_i + h, y_j + k)$$

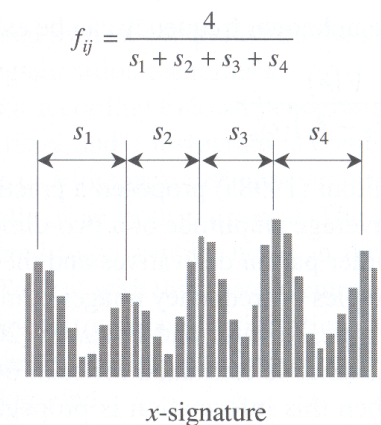
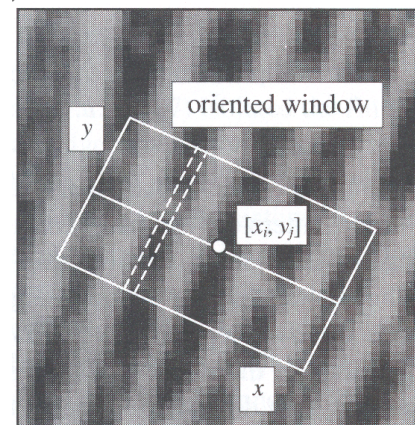
$$G_{xx} = \sum_{h=-8}^8 \sum_{k=-8}^8 \Delta_x(x_i + h, y_j + k)^2$$

$$G_{yy} = \sum_{h=-8}^8 \sum_{k=-8}^8 \Delta_y(x_i + h, y_j + k)^2$$

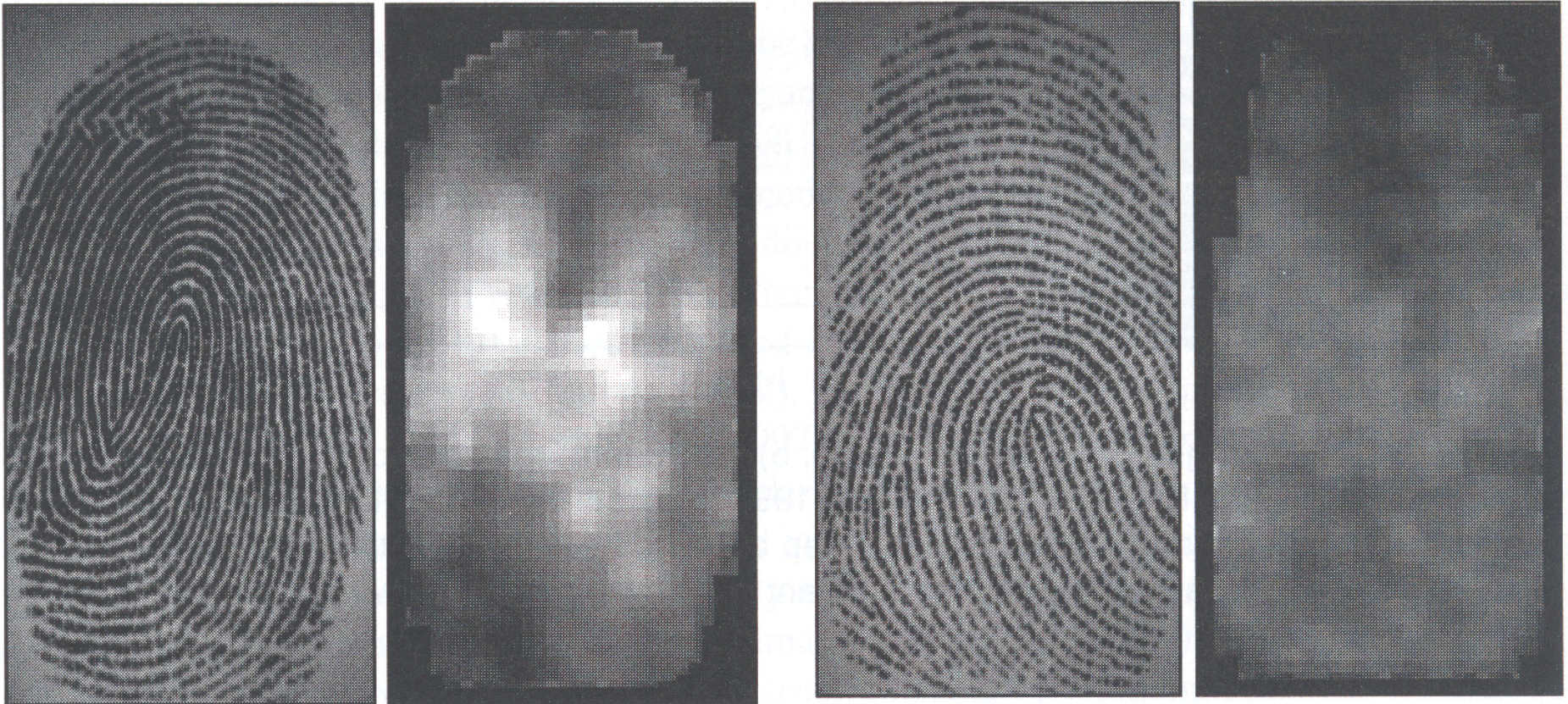
FP Features: Lokale Ridge Frequenz I

Die lokale Ridge Frequenz $f_{x,y}$ im Punkt (x,y) ist das Inverse der Anzahl der Ridges per Einheitslänge entlang einem Segment zentriert in (x,y) und orthogonal zur lokalen Ridge Richtung. Analog zum Orientierungsbild kann auch ein Frequenzbild definiert werden, indem die Ridge Frequenz nur an diskreten Positionen durch einen Durchschnittswert angegeben ist. Diese Frequenz variiert zwischen verschiedenen Fingern und variiert zwischen verschiedenen Regionen eines Fingers. $f_{x,y}$ wird wie folgt berechnet:

1. Ein 32×16 Fenster zentriert in (x,y) wird rotiert dass die y-Achse parallel zur Ridge Richtung liegt.
2. Die x-Signatur der Grauwerte wird erstellt durch Aufsummieren der Werte in jeder Spalte x des Fensters.
3. $f_{x,y}$ ist das inverse des durchschnittlichen Abstands zwischen zwei lokalen Maxima der x-Signatur.

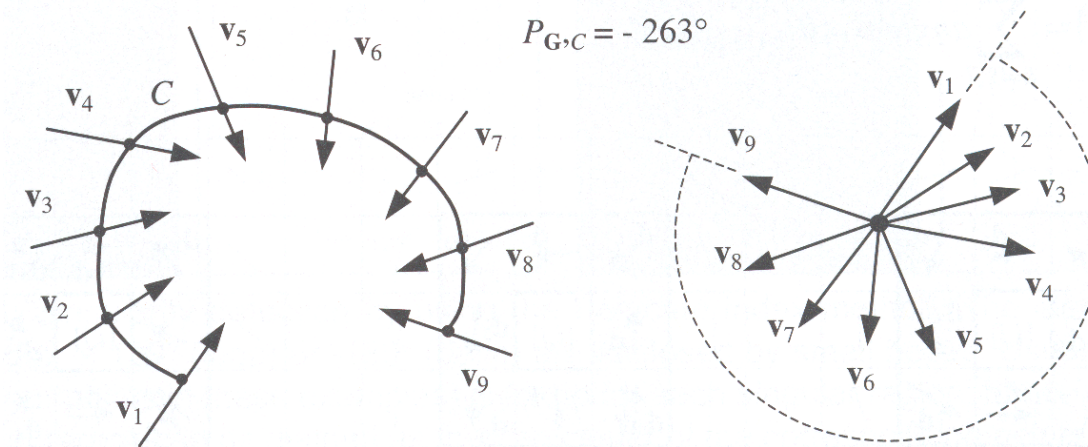


FP Features: Lokale Ridge Frequenz II



FP Features: Erkennung von Singularitäten und Core I

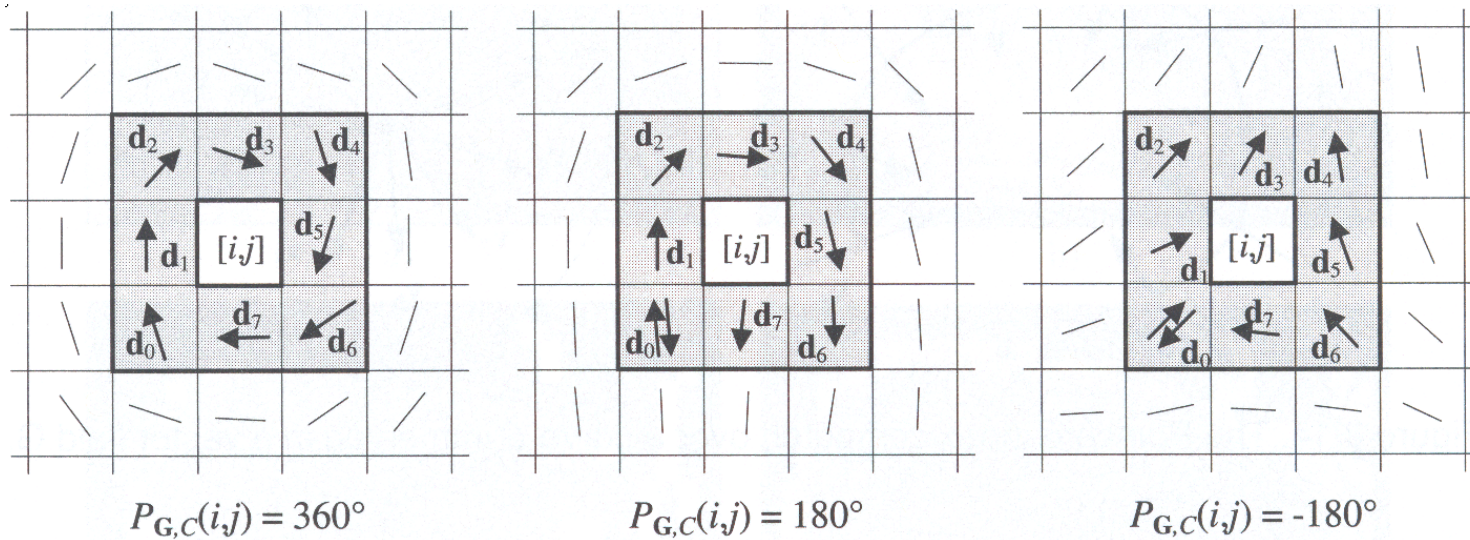
Die eleganteste Methode benutzt den sog. "Poincare Index". Wenn G ein Vektorfeld und C eine Kurve in diesem Vektorfeld ist, ist der Poincare Index $P_{G,C}$ definiert als die gesamte Rotation der Vektoren von G entlang der Kurve C :



Im FP Fall ist G das Orientierungsbild und C ein geschlossener Pfad von Elementen von G sodass (x, y) ein innerer Punkt ist. $P_{G,C}(x, y)$ wird berechnet durch Aufsummieren der Richtungsunterschiede benachbarter Elemente von C (hier braucht man eine gerichtete Richtung die bei ersten Vektor zufällig gewählt wird). Es ist bekannt, dass $P_{G,C}(x, y)$ für geschlossene Kurven nur Werte von 0 , $+180$ und $+360$ Grad annimmt. Bezogen auf FP-Singularitäten bedeutet das: $P_{G,C}(x, y) = 0$ keine Singularität, $P_{G,C}(x, y) = 360$ whorl, $P_{G,C}(x, y) = 180$ loop, $P_{G,C}(x, y) = -180$ delta Singularität.

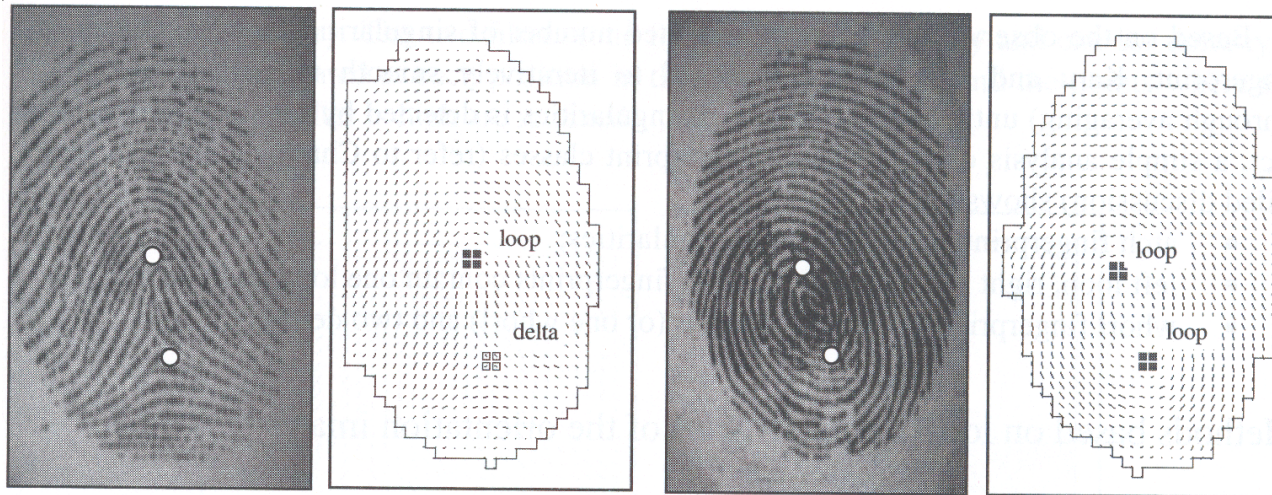
FP Features: Erkennung von Singularitäten und Core II

In der Graphik sind drei Ausschnitte aus Orientierungsbildern zu sehen. C ist die geordnete Folge der 8-Nachbarn d_k von (i, j) . Die Richtung der d_k ist definiert sodass d_0 nach oben zeigt; d_k ist gerichtet sodass der Absolutwert der Winkel zwischen d_k und d_{k+1} ist ≤ 90 Grad.

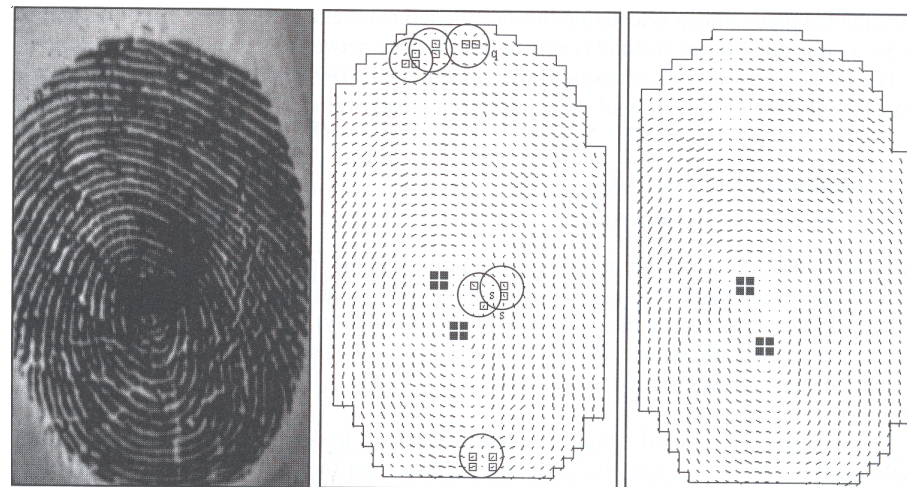


$$P_{G,C}(i, j) = \sum_{k=0}^7 \text{Winkel}(d_k, d_{(k+1) \bmod 8})$$

FP Features: Erkennung von Singularitäten und Core III

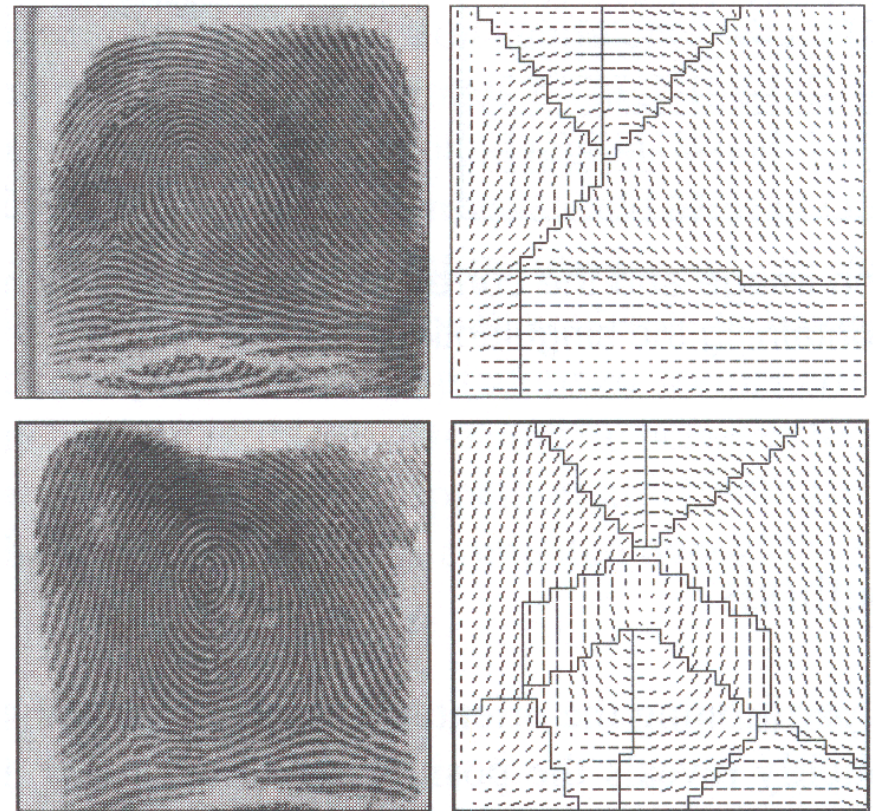


Glättung der Orientierungsbildes verhindert falsche Singularitäten !



FP Features: Erkennung von Singularitäten und Core IV

- Singularitäten sind Regionen wo das Orientierungsbild Irregularitäten aufweist, d.h. schnell wechselnde Richtung, hohe Krümmung
- Beim Partitionieren des Orientierungsbildes in homogene Regionen sind die Schnittpunkte der Grenzlinien Singularitäten
- Lokale Templates werden verwendet um den Core zu finden ("Sextet-Methode" des FBI)
- Focal Point: Schnittpunkt von Normalen auf die Ridge-Richtung



FP Enhancement

Klassische Verfahren wie Kontrastverbesserung, Histogrammequalisierung, Wiener Filterung und Normalisierung sind die ersten Schritte bei FP Enhancementverfahren. Diese Verfahren lösen allerdings das Problem von unterbrochenen Ridges u.s.w. nicht.

Ein klassischer Normalisierungsansatz bei gegebenen $I(x, y)$, m und v und dem gewünschten MW m_0 und Varianz v_0 ist

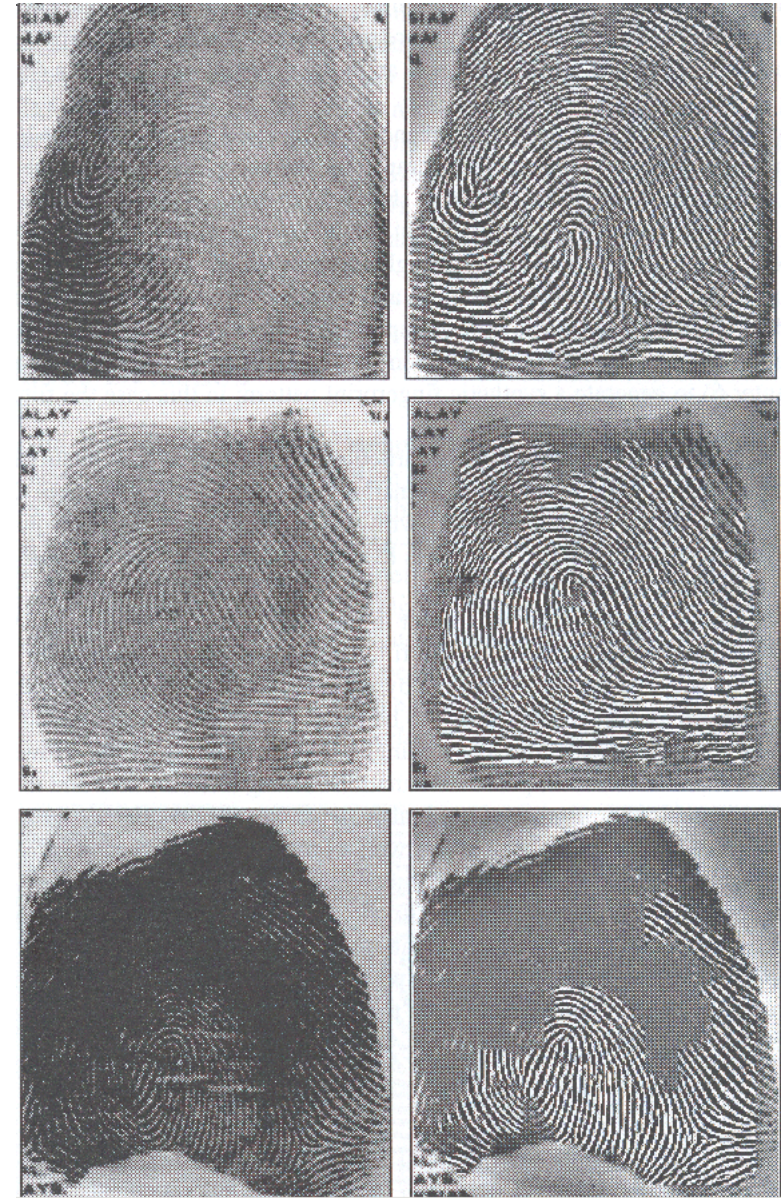
$$I'(x, y) = m_0 + ((I(x, y) - m)^2 v_0 / v)^{1/2} \text{ wenn } I(x, y) > m$$

ansonsten $m_0 - \dots$. Da diese Operation nur von einem Pixel abhängt, wird die grundlegende Struktur von Ridges und Valleys nicht verändert.

Die verbreitetste Technik ist kontextbasierte Filterung, wo sich die Filtercharakteristik entsprechend dem lokalen Kontext ändert. Dieser Kontext ist im Fall von FPs definiert durch die Ridgeorientierung und Ridgefrequenz. Die Sinus-ähnliche Form des Wechslens von Ridges und Valleys ist durch diese Parameter bestimmt und ändert sich eher langsam über die FP Fläche. Ein entsprechend optimierter Filter kann Noise und andere Effekte gut entfernen und die tatsächliche Struktur betonen.

FP Enhancement mit Gabor Filterung

Ein Set von Gabor Filtern wird vorherberechnet mit Frequenz und Orientierung wie sie in den ebenfalls vorberechneten Orientierungs- und Frequenzbildern vorkommen. Dann wird jedes Pixel im Bildbereich mit dem Filter gefaltet, der der Ridgeorientierung und Ridgefrequenz an dieser Position am ähnlichsten ist (ansonsten ist das zu aufwendig wenn für jedes Pixel der optimale Filter berechnet wird). Es gibt auch analoge Methoden mit direktonaler DFT Bandpassfilterung, die aber wegen der schlechteren Lokalisierung problematische Artefakte gibt.



Minutien Erkennung

Klassische Vorgehensweise ist FP Enhancement und nachfolgende Binarisierung. Anschliessend wird ein Thinning auf 1-Pixel breite Ridges durchgeführt. Das Ergebnisbild kann dann mit Templates auf entsprechende Minutienmuster abgesehen werden.



a)



b)

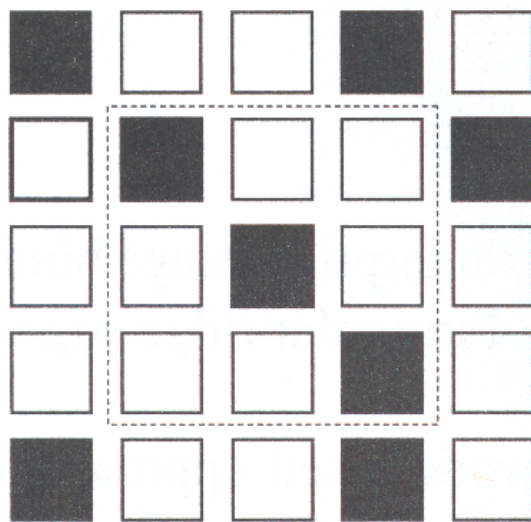


c)

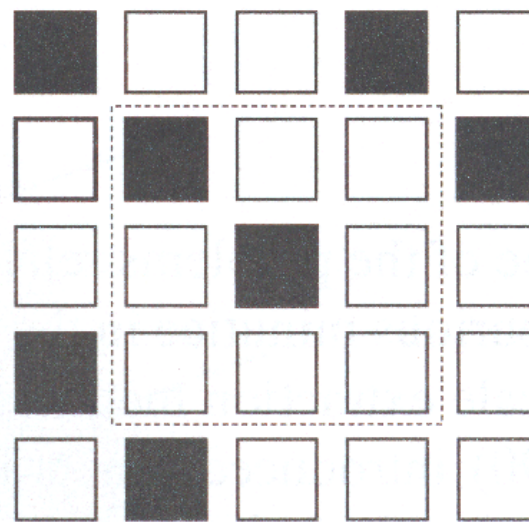
Minutien Erkennung: Binarisierungsmethoden

Einfachste Methoden verwenden globale und lokale Thresholds. Verbesserungen werden erzielt z.B. durch Verwendung von x-Signaturen (lokales Averaging) und die Verwendung der Peaks und danebenliegenden Pixeln als Vordergrund. Typische Methoden folgen den Ridges und eliminieren "Löcher", auch morphologische Operatoren mit speziell auf Ridge Form abgestimmten strukturierenden Element kommen zum Einsatz.

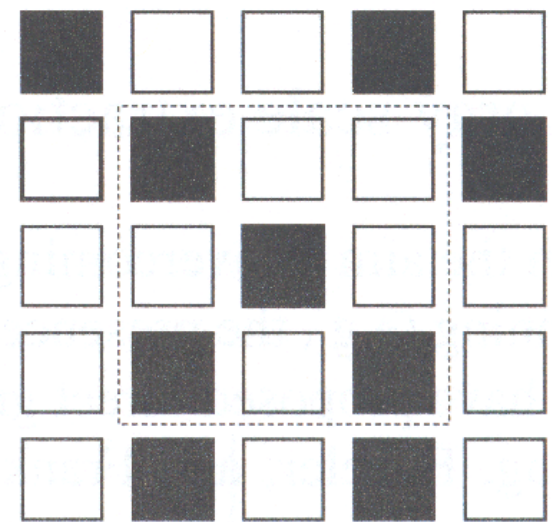
Ist das binäre Skeleton berechnet, kann mit einem einfachen Bildscan jedes Minutienpixel identifiziert werden: es hat eine Kreuzungszahl ungleich 2: Kreuzungszahl ist die Anzahl von 8-Nachbarn im binären Bild.



a) $cn(\mathbf{p}) = 2$



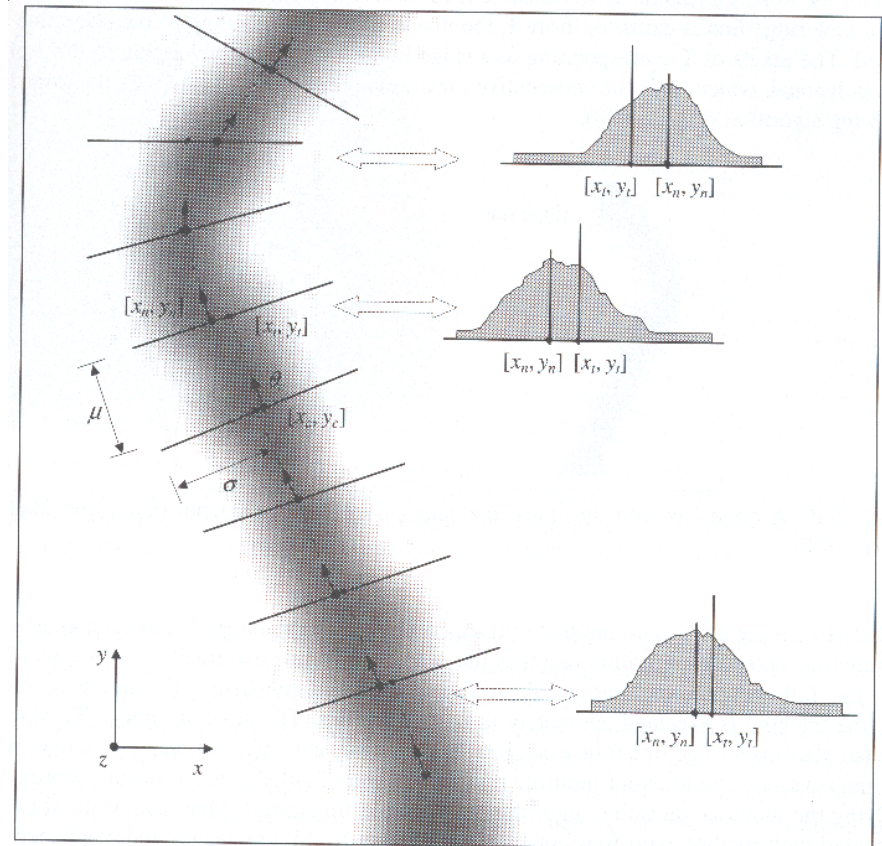
b) $cn(\mathbf{p}) = 1$



c) $cn(\mathbf{p}) = 3$

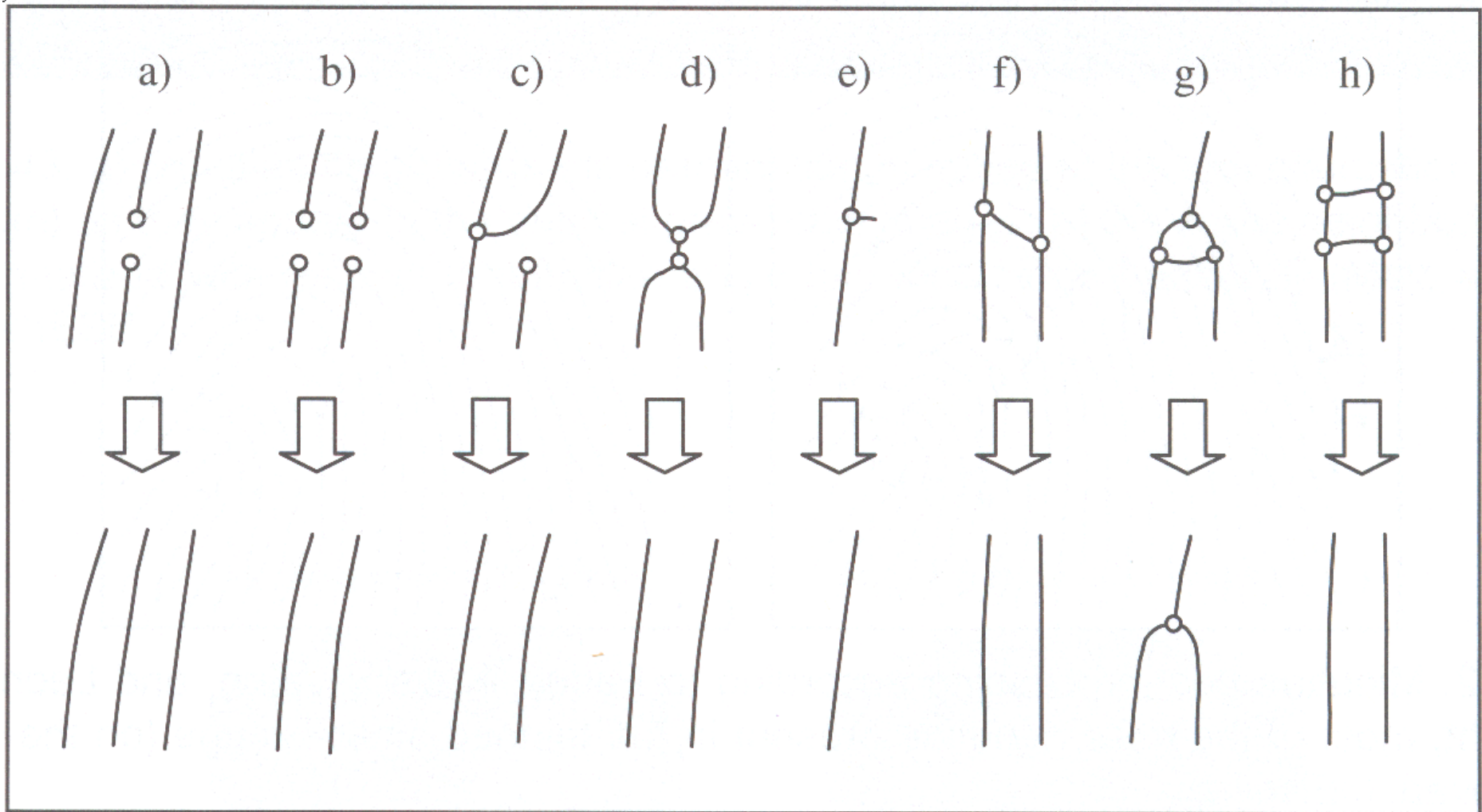
Minutien Erkennung im Grauwertbild

Um Artefakte durch Binarisierung und Thinning zu vermeiden, werden auch Methoden direkt im Grauwertbild vorgeschlagen: hier werden lokale Maxima eines Segments das orthogonal zur Ridgeorientierung liegt berechnet und als Ridge Zentrum definiert. In einem Hilfsbild werden dann diese verbesserten Hilfsridges mit einem konstanten Saum eingetragen und in diesem Hilfsbild die Minutiendetektionsverfahren angewendet.



Minutien Filterung

Für gute Matchingergebnisse ist es wesentlich, falsch erkannte Minutien zu entfernen.
Hier einige Beispiele:



FP Features: Ridge Count

In forensischen Systemen wird gern die Anzahl von Ridges zwischen bestimmten Punkten bestimmt, hier werden v.a. Singularitäten als Bezugspunkte gewählt. Eine mögliche Methode ist die Bestimmung der Anzahl der lokalen Maxima über die Methode der x-Signatur.

Problematisch ist hier wieder die zuverlässige Bestimmung der Referenzpunkte.

FP Matching: Grundlagen

Gegeben sind das Template T (vom Enrollmentvorgang) und das Inputbild I . Das Matching ist für FPs eher schwierig durch die hohe intra-personal Variability, die entsteht durch:

- Translation und Rotation beim Abnahmevorgang
- Nicht-lineare Störungen: Das Sensing bildet den 3D Finger auf den 2D FP ab – durch die Elastizität der Haut entstehen bei zur Sensorfläche nicht-orthogonalen Bewegungen entsprechende Störungen. Es gibt Methoden und Ideen, diese Störungen zu filtern.
- Druck und Hautzustand (Nässe, Fett, Schmutz)
- Noise des Sensors
- Fehler im Bereich der Merkmalsextraktion

FP Matching: Grundtypen

1. Korrelations-basiertes Matching: Zwei FPs werden “übereinandergelegt” und die Korrelation zwischen entsprechenden Pixeln ist berechnet für verschiedene Translationen und Rotationen.
2. Minutien-basiertes Matching: hat seinen Ursprung in manuellen FP-Matching Verfahren; die meisten FP Matching Verfahren arbeiten so. Grundidee ist es, das Alignment zwischen T und I zu finden das in der maximalen Anzahl von übereinstimmenden Paaren resultiert.
3. Ridge-feature basiertes Matching: In FPs schlechter Qualität können Minutien schlecht extrahiert werden. Andere Features wie lokale Orientierung und Frequenz, Textur Information u.s.w. können mit besserer Zuverlässigkeit extrahiert werden, sind aber auch weniger diskriminativ.

FP Matching: Korrelationsmethoden

Wenn $CC(T, I) = T^T I$ die Kreuzkorrelation zwischen I und T ist, so wird die Ähnlichkeit zwischen zwei FPs bestimmt durch

$$S(T, I) = \max_{\delta x, \delta y, \Theta} CC(T, I^{(\delta x, \delta y, \Theta)})$$

Θ bezeichnet eine Rotation um des Zentrum des Bilder (Schwerpunkt ?), $(\delta x, \delta y)$ eine Verschiebung.

Eine direkte Anwendung bringt kaum Erfolge:

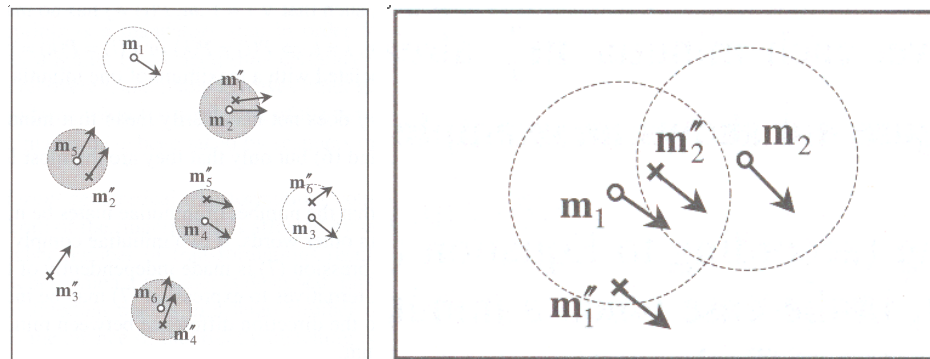
- Nicht-lineare Distortion macht globales Matching kaum durchführbar
- Hautzustand und Fingerdruck verursachen eine Veränderung von Helligkeit, Kontrast und Ridgedicke. Dies muss jedenfalls bei der Berechnung korrigiert werden (normalisierte Kreuzkorrelation, zero mean normalisierte Kreuzkorrelation u.s.w.)
- Der rechnerische Aufwand ist enorm

Lösungsmöglichkeiten bringen lokale Berechnungen von Korrelation in Fenstern (z.B. solchen mit interessanten Punkten wie Minutien oder Singularitäten).

Komplexitätsreduktion bringen Multiresolution verfahren und Korrelationsberechnung in der DFT Domain (Faltungssatz!).

FP Matching: Minutienmethoden I

Das Matching von zwei Minutienmengen wird in der Literatur als “Point Pattern Matching” und es gibt ein enormes Ausmass an Literatur dazu. Eine Minutia in I und eine in T werden als Matching betrachtet wenn der Abstand zwischen ihnen kleiner als eine gegebene Schranke ist und wenn die Richtungsabweichung kleiner als eine Schranke ist (und wenn der Typ übereinstimmt im Fall von klassifizierten Minutien). Zusätzlich müssen die FPs gegeneinander ausgerichtet werden, um ein maximales Ausmass an Minutienpaaren zu finden: Translation und Rotation werden üblicherweise zugelassen um die Anzahl zu maximieren. Es ist zu beachten dass es klassischerweise um die Maximierung der Anzahl von Paaren geht, auch wenn in einer anderen Konfiguration der Fehler bei den Paaren kleiner wäre. Aber auch die Abstände können mitverwendet werden. Wenn bekannt wäre, welche Minutien Paare bilden sollen, wäre die Lösung einfach, so ist sie aber exponentiell in der Anzahl der Minutien. Beispiele:



FP Matching: Minutienmethoden II

- Relaxation: iterativ wird die Zusammengehörigkeit von Minutienpaaren bewertet unter Miteinbeziehung von benachbarten Minutienpaaren (vgl. Edge Relaxation)
- Hough Transformation: ein Akkumulations Array wird initialisiert, das die Parameter von möglichen Alignments zwischen den FPs enthält. Dann wird eine Doppelschleife über alle Minutien in T und I gestartet, in der für alle Winkel und Scales der Abstand in x- und y-Richtung zwischen den Minutien bestimmt wird und dieser quantisiert wird. Das Akkumulatorarray an der Position von Winkel, Scale und den beiden Abständen wird inkrementiert. Das Maximum im 4D Array bestimmt das optimale Alignment bzgl. Translation, Rotation und Scale. Das Matching ist dann trivial. Hier werden FPGA Implementierungen vorgeschlagen (Komplexität !)
- Pre-alignment
 - ★ Absolut: Translation in bestimmte Position des Core Points
 - ★ Relativ (immer ein FP Paar): Bestimmung eines "Principal Pairs" durch das dann die FPs ausgerichtet werden. Dieses Minutienpaar wird durch die verbindenden Linien charakterisiert, deren Winkel und Längen untersucht werden. Alternativ ist Ausrichtung nach Singularitäten, Korrelation des Orientierungsbildes, Korrelation von Ridge Features.

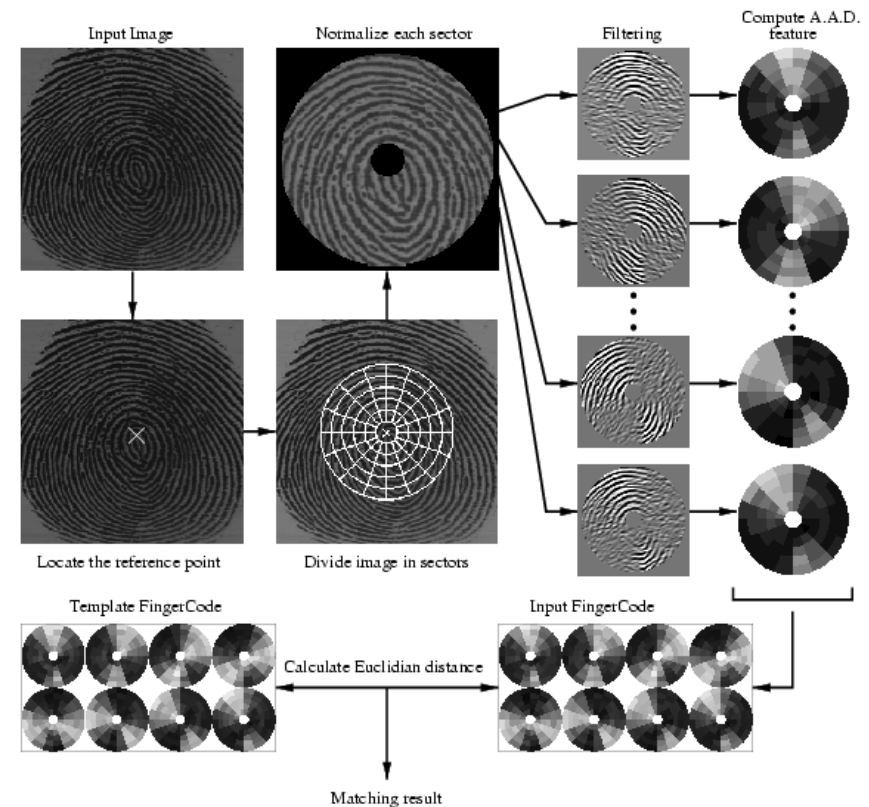
FP Matching: Lokale Minutienmethoden

Die Idee ist hier lokale Minutienstrukturen zu verwenden, die weniger anfällig gegenüber globalen Transformationen sind. Der Aufwand ist geringer, aber auch die Distinctiveness geringen. Eine mögliche Anwendung verwendet lokales Matching für ein Alignment mit einem finalen globalen Matchingschritt. Varianten:

- Featurevektor enthält die Anzahl der Minutien eines bestimmten Typs in einer Umgebung.
- Relative Orientierung zur Orientierung der zentralen Minutia, Entfernung und Ridge-count zu den Minutien in der Umgebung wird zusätzlich aufgenommen
- Graphen Notation: Ein Stern der zu einer Minutie gehört besteht aus den Knoten (die Minutien in der Umgebung) und den Edges (Verbindungsgeraden zu diesen Minutien) der Umgebung. Jeder Stern in T wird gegen jeden Stern in I gematched, unter verschiedenen Rotationswinkeln.

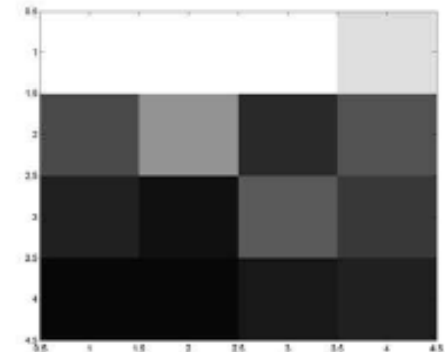
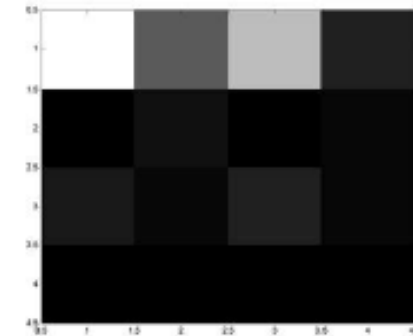
FP Matching: Ridge Feature Methoden - Gabor

Idee ist hier einen FingerCode zu entwickeln: nach der Identifikation des Referenzpunkts (Core – kritisch !!) wird ein normiertes Kreissegmentmuster über den FP gelegt (80 Segmente). Die Segmente werden normalisiert und anschliessend mit Gabor Filtern einer Frequenz (entspricht Ridge Frequenz) und 8 Orientierungen gefiltert. Für jedes dieser 640 Segmente wird der durchschnittliche absolute Unterschied zum Sektormittelwert berechnet und als Feature verwendet. Matching geschieht durch Ermittlung der Euklidischen Distanz zwischen den Featurevektoren. Rotationsinvarianz durch wiederholtes Matching von rotierten FingerCodes (und dem Speichern feiner rotierter FingerCodes).



FP Matching: Ridge Feature Methoden - DWT & WPT

Idee ist hier Texturdeskriptoren zu verwenden, wie schon in einem Verfahren für Iriserkennung. Dabei wird der FP in eine bestimmte Subbandstruktur zerlegt (die DWT basierend sein kann oder eine für Matching optimierte WP Version ist). Diese wird in Blöcke bestimmter Größe eingeteilt und für diese Blöcke wird dann ein Koeffizientenbasierter Featurewert berechnet (z.B. Energie als Summe der quadrierten Koeffizienten). Der Featurevektor wird durch diese Features gebildet, Matching wieder durch Euklidische Distanz zwischen diesen Vektoren, hier kann eine Gewichtung je nach Wichtigkeit vorgenommen werden. Auch hier ist eine Registrierung der FPs notwendig, Rotationsrobustheit besteht zumindest gegenüber kleinen Rotationen.



Finale

Ich hoffe ich habe mit dieser LVA ihr Interesse an der Thematik wecken können. Ich wünsche ihnen schöne Ferien und alles Gute bei der Klausur.

Literaturhinweise