Fusion von biometrischen Verfahren zur Benutzerauthentifikation

Claus Vielhauer · Tobias Scheidat

Otto-von-Guericke Universität Magdeburg
Advanced Multimedia and Security Lab (AMSL)
claus.vielhauer@iti.cs.uni-magdeburg.de · scheidat@cs.uni-magdeburg.de

Zusammenfassung

Der biometrischen Benutzerauthentifikation durch die Handschrift wird heutzutage eine besondere Bedeutung beigemessen. Dies liegt hauptsächlich an der seit Jahrhunderten andauernden und weit verbreiteten Nutzung der Unterschrift zur Identifikation der Unterzeichner wichtiger Papiere. Die aktuelle Hardware bietet nun die Möglichkeit, neben dem (statischen) Schriftbild die dynamischen Merkmale der Handschrift aufzunehmen. Gerade diese zeitveränderlichen Merkmale, wie zum Beispiel der Druckverlauf, bieten eine sehr gute Grundlage für die automatische Benutzerauthentifikation durch Computer. Dadurch ist es möglich, nicht nur den Unterzeicher von Dokumenten einfacher und genauer festzustellen, sondern auch die individuelle Handschrift als Berechtigungsnachweis zu nutzen. Ausgangspunkt dafür ist ein biometrischer Algorithmus, der die Berechtigung einer Zugang verlangenden Person überprüfen kann. Der Ausgabewert des Algorithmus wird vom System genutzt, um den Zugang zum geschützten Bereich zu gewähren oder zu verweigern. Aufbauend auf den Fehlerraten mehrerer einzelner Algorithmen werden unterschiedliche Strategien vorgestellt, mit denen eine vorteilhafte Fusion der Algorithmen erreicht wird. In dieser Arbeit wird gezeigt, dass es durchaus möglich ist eine Verbesserung durch die Fusion zu erzielen. Grundlage ist eine geeignete Gewichtung an angebrachter Stelle innerhalb der Algorithmen, um eine Verbesserung hervorzurufen.

1 Motivation

Die biometrische Benutzerauthentifikation steht kurz vor dem Einsatz im Massenmarkt, angetrieben u.a. durch die Bestrebungen zur Einführung von Biometrie in Ausweisdokumenten [PeSS03]. Darüber hinaus werden biometrische Verfahren künftig im Bereich des ganzheitlichen IT Managements eine wichtige Rolle spielen, da sie als Ergänzung oder Alternative zur Authentifikation basierend auf Besitz und Wissen das Sicherheitsniveau in Infrastrukturen signifikant erhöhen können. Aus der Menge der unterschiedlichen geeigneten biometrischen Modalitäten (z.B. Fingerabdruck, Gesicht, Iris oder Stimme) bietet sich durch ihre individuelle Einzigartigkeit die Handschrift für die Nutzung als biometrisches Merkmal an. Schon seit vielen Jahrhunderten wird die persönliche Unterschrift im gesellschaftlichen und privaten Bereich genutzt, um eine willentliche Einverständniserklärung abzugeben. Dies ist zum Beispiel beim Abschluss von Verträgen üblich. Aus diesem Grund genießt die Handschrift ein hohes Maß an Akzeptanz als Authentifikationsmerkmal [Kais01].

Durch die Nutzung von Computer gestützter Aufnahme und Untersuchung von Handschriften ist es heute möglich, diese nicht nur zur Unterzeichnung wichtiger Unterlagen, sondern auch für den Einsatz im Bereich der digitalen Signaturen [ViSt03] oder zur Benutzerauthentifizie-

P. Horster (Hrsg.) · D•A•CH Security 2005 · syssec (2005) 82-97.

rung im Kontext von Sicherheitsinfrastrukturen mittels biometrischer Systeme zu nutzen. Dabei werden die dynamischen Merkmale der Schrift über Digitalisiertabletts oder auch Druck sensitive Displays von Personal Digital Assistants (PDAs) und Tablet-PCs aufgenommen. Dynamische Merkmale sind dabei zeitveränderliche Werte, die im Verlauf des Schreibens durch den Urheber erzeugt werden. Dazu zählen beispielsweise der zeitliche Verlauf der horizontalen und vertikalen (XY) Koordinaten und des Drucks der Stiftspitze während des Schreibvorgangs. Aus diesen Daten können weitere Größen abgeleitet werden, beispielsweise Schreib-Geschwindigkeit und Schreib-Beschleunigung. Durch die Aufzeichnung der temporären Veränderung der XY-Koordinaten ist auch die spätere Wiederherstellung des grafischen Schriftbildes möglich. Für die Schreiberverifikation gibt es eine große Vielzahl von biometrischen Verfahren, ein umfangreicher Überblick findet sich in der Literatur beispielsweise bei [GuCa97] bzw. [PlLe94].

In dieser Arbeit werden Ansätze untersucht, inwieweit es möglich ist, durch die Kombination mehrerer biometrischer Algorithmen zur handschriftbasierten Benutzerauthentifikation eine Verbesserung in der Erkennung zu erzielen. Durch die Fusion mehrerer Algorithmen soll versucht werden, die Leistungsfähigkeit der Handschriftenerkennung eines biometrischen Systems zu erhöhen. Dazu wurde der Biometric Hash Algorithmus [ViSM02] verwendet.

Da der Biometric Hash Algorithmus, wie auch die große Vielzahl anderer, nicht handschriftbasierter biometrischer Verfahren, mit einem Abstandsmaß arbeitet, lag der Gedanke nahe, alternative Distanzen in den Algorithmus zu integrieren. Bei den Tests des Biometric Hash Verfahrens mit den unterschiedlichen Abstandsmaßen stellte sich heraus, dass das optimale Ergebnis für verschiedene Tabletts bzw. Semantik-Klassen (d.h. alternative Schriftzüge zur Unterschrift wie z.B. Passphrases, siehe [Viel00]) nicht immer auf dem gleichen Abstands-Algorithmus basierte. Dadurch entstand die Idee, die Klassifikation durch eine Kombination mehrerer Algorithmen durchführen zu lassen. Um auch der jeweiligen Eignung für bestimmte Tabletts und Semantiken Rechnung zu tragen, musste eine Strategie zur Bestimmung einer geeigneten Gewichtung der einzelnen Verfahren entwickelt werden.

Der Beitrag ist wie folgt gegliedert: im anschließenden Kapitel wird auf ein allgemeines biometrisches System eingegangen. Das dritte Kapitel befasst sich mit der Kombination verschiedener biometrischer Merkmale zur Benutzerauthentifikation. Dabei wird insbesondere auf drei verschiedenen Ansätze einer biometrischen Fusion eingegangen, welche grundsätzlich Modalitäten übergreifend anwendbar, und somit unabhängig von der von uns im Speziellen untersuchten Modalität der Handschrift sind. Darauf folgt eine Beschreibung der Umsetzung der multialgorithmischen Fusion in der handschriftbasierten Benutzerauthentifikation. Dabei werden auch die verwendeten Abstandsmaße genauer beschrieben. Auf die Testergebnisse basierend auf der Fusion von vier biometrischen Algorithmen wird im fünften Kapitel eingegangen. Im letzten Kapitel wird der Beitrag zusammengefasst und ein Ausblick auf die zukünftige Arbeit in Bezug auf die vorgestellten Ergebnisse gegeben.

2 Biometrische Systeme

Im Folgenden soll die Funktionsweise eines biometrischen Systems zur Authentifikation beschrieben werden. Bei der Authentifikation wird überprüft, ob die Daten, die am Sensor aufgenommen wurden mit den in der Datenbank hinterlegten in ausreichendem Maß übereinstimmen. Sie kann unterteilt werden in die Identifikation und die Verifikation. Wird überprüft, ob eine Person diejenige ist, für die sie sich ausgibt, spricht man von einer Verifikation.

Dabei werden die aktuell aufgezeichneten biometrischen Daten mit den in der Datenbank hinterlegten Referenzdaten verglichen, die sich auf einen bestimmten Identifikator (d.h. der behaupteten Identität, z.B. Nutzername) beziehen. Bei der Identifikation werden die biometrischen Merkmale einer vorerst unbekannten Person mit den Daten aller dem System bekannten Personen aus der Referenzdatenbank verglichen. Die Person gilt als identifiziert, wenn ein Datensatz gefunden wird, in dessen Toleranzbereich die Eingabemerkmale liegen.

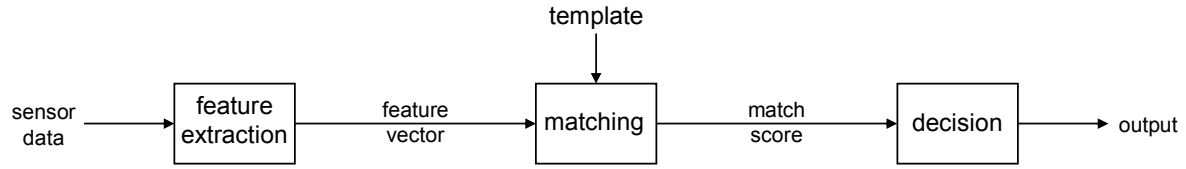

Abb. 1: Struktur eines biometrischen Systems

In Abbildung 1 ist die Funktionsweise eines biometrischen Systems dargestellt. Am Sensor werden die Daten des Nutzers aufgenommen und nach einer eventuellen Vorverarbeitung an den Feature-Extraktor weitergegeben. Dieser wertet die biometrischen Eigenschaften des Merkmals aus und wandelt die Daten mittels mathematischer und statistischer Operationen in einen n-dimensionalen Feature-Vektor um, welcher innerhalb des Systems das biometrische Merkmal beschreibt. Im Matching-Modul wird der aktuelle Feature-Vektor mit dem ebenfalls n-dimensionalen Referenz-Feature-Vektor des Nutzers verglichen. An dieser Stelle wird der Grad der Ähnlichkeit zwischen beiden Vektoren bestimmt. Als Ergebnis des Vergleichs wird ein Wert (Match Score) berechnet, der den Grad der Übereinstimmung von Test- und Referenz-Vektor beschreibt. Dieser wird im Decision-Prozess zum Herbeiführen der Entscheidung genutzt, ob die Person diejenige ist, die sie vorgibt zu sein bzw. um welche Person es sich handelt. Diese Entscheidung stellt gleichzeitig den Ausgabewert des biometrischen Systems dar.

Fehlerraten sind Werkzeuge, um die Erkennungsgenauigkeit biometrischer Systeme zu bestimmen. Auch der Vergleich biometrischer Systeme und Algorithmen ist mit ihnen möglich. In der Biometrie können die Fehlerraten nicht gemessen werden, stattdessen müssen sie empirisch ermittelt werden. Dazu muss, um möglichst kleine Fehlerraten zu erreichen, ein sehr großer Testaufwand betrieben werden. Weiter führende Informationen zur Evaluationsproblematik sind u.a. in der Literatur unter [Laßm02] zu finden.

Um einen Algorithmus beurteilen zu können, wird er auf mehrere unterschiedliche Test-Szenarien angewandt. Die Ergebnisse aller Algorithmen werden dabei miteinander verglichen, um festzustellen, welcher der vorteilhafteste ist. Hierbei wird häufig die so genannte die Equal Error Rate (EER) herangezogen. Das ist der Punkt, an dem die FNMR (False Non Match Rate, d.h. prozentuale Häufigkeit von Abweisungen authentischer Benutzer) und die FMR (False Match Rate, d.h. prozentuale Häufigkeit der Akzeptanz nicht-authentischer Personen) identisch sind. Dieser Punkt muss nicht notwendigerweise der optimale Arbeitspunkt für den jeweiligen Algorithmus bzw. die Anwendung sein, ist aber ein guter Anhaltspunkt zur Einschätzung der Leistung eines Algorithmus. Die FNMR wird über die Verifikationen eines Users und dessen Enrollments bestimmt. Aus den zufälligen bzw. Brute Force Angriffen kann die FMR berechnet werden. Als zufällige Angriffe werden in unserem System die Verifikationen der jeweils anderen Nutzer bezeichnet, die mit den Enrollments einer Person verglichen werden. Die willentlichen Angriffszenarios basieren auf einer Evaluierungsmethodologie aus [ZoVi03] und umfassen Fälschungen mit unterschiedlichem Angriffsaufwand wie Blind At-

tack, Low Force Attack und Brute Force Attack. Sie unterscheiden sich im Wissenstand des Fälschers über die jeweilige Schriftprobe. Dieser beinhaltet beim blinden Angriff nur die Kenntnis, welcher textuelle Inhalt zu schreiben ist. Bei der Low Force Attacke ist zudem noch das Schriftbild bekannt. Die meiste Information steht beim Brute Force Angriff zur Verfügung, hier sind zusätzlich noch physische Gegebenheiten, wie Schreibgeschwindigkeit oder –Schreibsequenz bekannt. Bei den Angriffstests findet jeweils ein Vergleich zwischen einem Enrollment und den damit verbundenen Fälschungen statt.

3 Multibiometrische Systeme

Wir bezeichnen multibiometrische Systeme als solche, die aus mehreren biometrischen Sub-Systemen für unterschiedliche Modalitäten (z.B. Fingerabdruck und Iris) bestehen, dergestalt, dass sie sich gegenseitig ergänzen. Die wechselseitige Beeinflussung der Systeme bezieht sich dabei auf die Erkennungsgenauigkeit, die Sicherheit und die Variabilität. In der Literatur ist ein solches System beispielsweise bei [JaRo04] beschrieben, das auf den Merkmalen Gesicht, Fingerabdruck und Handgeometrie basiert.

Für die Fusion von biometrischen Systemen gibt es nach [JaRo04] drei unterschiedliche Ansatzpunkte, an denen die Ergebnisse bzw. Zwischenergebnisse der beteiligten biometrischen Systeme vereinigt werden. Diese drei möglichen Levels der Fusion werden in den nachfolgenden Abschnitten genauer beschrieben. Wichtig zu erwähnen ist, dass sich die verwendeten Eingangs- und Referenzdaten sehr stark unterscheiden. Dies ist bedingt durch die unterschiedlichen Sensoren und Merkmale auf denen die einzelnen Systeme basieren.

3.1 Feature Extraction Level

Beim Feature Extraction Level werden nach der Merkmals-Extraktion die Feature-Vektoren der einzelnen Systeme zu einem gemeinsamen Vektor vereinigt. Dies kann beispielsweise durch einfaches Aneinanderhängen (Konkatenation) der Vektoren geschehen. Der neue, erweiterte Feature-Vektor ist dann die Eingabemenge für den Matching-Prozess. Ab hier verhält sich das System wie ein einfaches biometrisches System.

Die Fusion auf Feature Extraction Level wird nur sehr selten genutzt, weil zum einen bestimmte biometrische Merkmale nicht bei allen teilnehmenden Personen vorhanden oder gleich gut ausgebildet sind. Auf der anderen Seite können auch rechnerische Probleme auftreten, da durch die Vereinigung mehrerer hoch dimensionaler Vektoren ein entsprechend großer und unhandlicher Feature-Vektor entstehen kann.

3.2 Matching Score Level

Wird auf dem Matching Score Level fusioniert (Abbildung 2), durchlaufen die Testdaten das jeweilige biometrische System bis zum Matching Prozess. Dort werden sie zunächst separat mit den entsprechenden Referenzdaten verglichen und nach einem bestimmten Algorithmus ein fusionierter Matching Score berechnet.

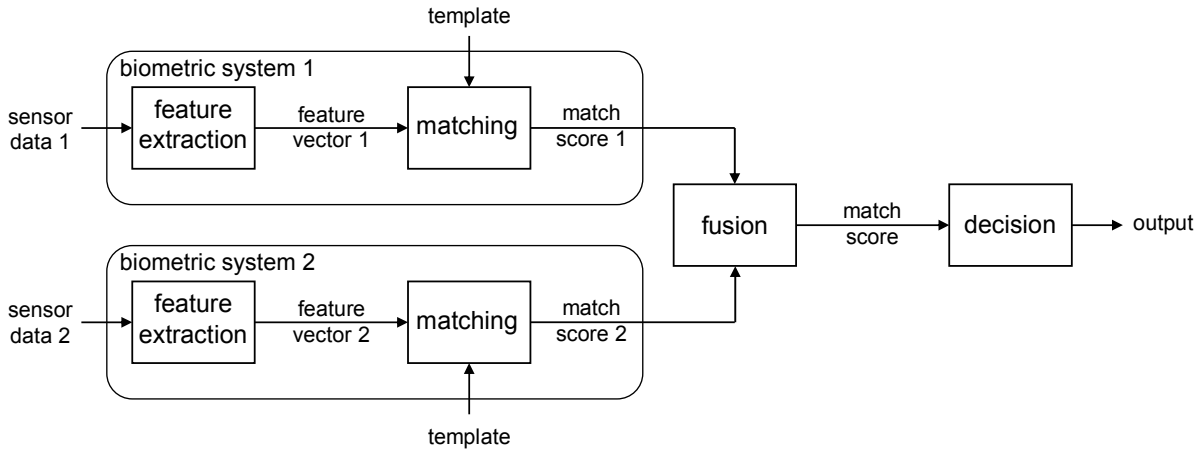

Abb. 2: Fusion auf Matching Score Level

Die Ergebnisse der einzelnen Systeme müssen dabei zunächst normalisiert werden, dann können sie zu einem neuen Match Score kombiniert werden. Dieser bildet dann die Grundlage für den Entscheidungsprozess. Die Fusion auf Matching Score Level wird gegenüber dem Feature Extraction Level häufig bevorzugt, weil es hier relativ einfach ist, auf die von den einzelnen Verfahren gelieferten Matching Score Werte zuzugreifen und zu kombinieren. Dabei kann auf der Fusionsebene von der Frage der Systemparameter und Feature Repräsentation abstrahiert werden, lediglich die Normalisierung der Matching Scores auf definierte Wertebereiche ist notwendig.

3.3 Decision Level

Die Grundlage der Fusion auf Decision Level ist der komplette Durchlauf aller beteiligten biometrischen Systeme bis zur Entscheidung in Bezug auf die Verifikation bzw. Identifikation des Nutzers. Im Anschluss werden die Entscheidungswerte zu einer finalen Entscheidung kombiniert. Die Fusion kann beispielsweise durch boolesche Operationen (z.B. AND, OR), durch eine Mehrheitsentscheidung oder durch eine Zufallsentscheidung erfolgen.

Vorteilhaft bei der Fusion auf Decision Level ist die Trivialität der Entscheidungsstrategie: der Fusions-Prozess benötigt keinerlei Information über die einzelnen Verfahrensparameter. Ein Nachteil der Fusion auf Decision Level ist ihre geringe Leistungsfähigkeit. Grund hierfür ist, dass die Fusion zu einem sehr späten Zeitpunkt im gesamten Prozess erfolgt und somit eine Gewichtung der Einzelentscheidungen aufgrund des binären Charakters der Einzelentscheidungen nicht möglich ist.

4 Multialgorithmische biometrische Systeme

Multialgorithmische Systeme beruhen auf biometrischen Merkmalen einer einzigen Modalität (z.B. nur Handschrift), welche anhand verschiedener unabhängiger Vergleichsverfahren zur Authentifikation herangezogen werden. Die Authentifikationsentscheidung basiert dabei auf einer Fusionsstrategie der jeweiligen Einzelergebnisse. Ein Beispiel für ein multialgorithmisches System findet sich in [CzKV04].

4.1 Multialgorithmische Fusion auf Matching Score Level

Aufgrund der im vorigen Kapitel erwähnten Nachteile der Fusion auf Feature Extraction bzw. Decision Level, basieren unsere ersten Untersuchungen auf dem Matching Score Level. Von großem Vorteil ist hier die Möglichkeit der Parametrisierung der einzelnen Match Scores vor der Fusion. Dadurch ist eine bessere Anpassung an Nutzer und Hardware möglich. Auch die notwendige Normalisierung, die zur Vergleichbarkeit bzw. zur Kombination der einzelnen Werte erforderlich ist, lässt sich an dieser Stelle relativ einfach durchführen. Im Folgenden wird das Vorgehen bei der Fusion auf Matching Score Level genauer beschrieben.

Zur Bewertung des Ausmaßes der Übereinstimmung zwischen einem Enrollment und einem Sample sind entsprechende Maßzahlen notwendig, welche anhand des folgenden Beispiels erläutert werden. In Abbildung 3 ist die Variationsbreite (grau) eines horizontalen Schreibsignals dargestellt, welche aus einem Enrollment, basierend auf fünf Unterschriften, berechnet wurde. Die schwarze Kurve stellt eine weitere Unterschrift derselben Person zu einem anderen Zeitpunkt dar. Dabei ist deutlich zu erkennen, dass sich diese in den meisten Bereichen innerhalb der Variationsbreite des Enrollments befindet. An einigen Stellen verläuft sie jedoch außerhalb. Um nun zu ermitteln, ob die Unterschrift innerhalb eines festgelegten Toleranzbereiches mit den Referenzdaten übereinstimmt, muss die Ähnlichkeit zwischen beiden ermittelt werden. Die Ähnlichkeiten werden hier durch Abstandsmaße beschrieben. Sie geben in einem einzelnen Wert an, wie weit voneinander entfernt die beschreibenden Merkmale von Enrollment und Testsample liegen. Idealerweise sollten bei der Berechnung alle zur Verfügung stehenden statistischen Merkmale berücksichtigt werden.

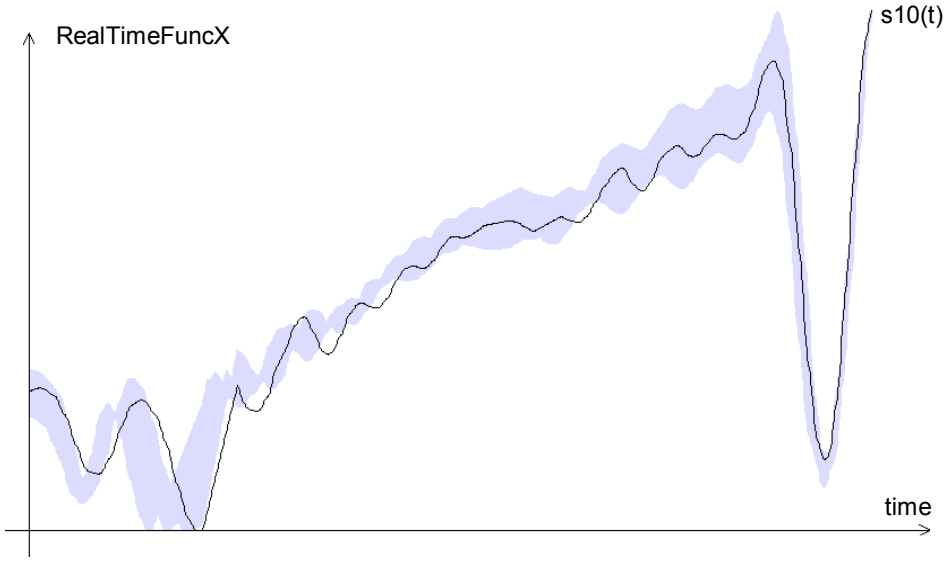

Abb. 3: Vergleich zwischen Variationsbreite eines Enrollments und einer einzelnen Unterschrift

Für die Untersuchungen der diesem Artikel zugrunde liegenden Arbeiten wurden die folgenden Abstandsmaße verwendet: die Canberra-Distanz, die City-Block-Distanz, die Euklidische Distanz und die Hamming-Distanz.

Als Algorithmus zur Extraktion statistischer Merkmale wird der Biometric Hash Algorithmus [ViSM02] verwendet. In diesen wurden die angegebenen Abstandsmaße zur Bestimmung der Ähnlichkeit eingebettet.

Die Funktionsweise des multialgorithmischen Systems ähnelt sehr dem Ablauf innerhalb eines multibiometrischen Systems. Dabei ist jedoch zu beachten, dass im Unterschied zur Fusion von biometrischen Systemen hier jeweils die gleichen Test- und Referenzdaten als Berechnungsgrundlagen für die einzelnen Algorithmen dienen. Bei einem multibiometrischen System unterscheiden sich die aktuellen Eingabedaten bzw. Templates entsprechend der verwendeten biometrischen Merkmale zum Teil sehr stark von einander. In Abbildung 4 ist die Struktur eines multialgorithmischen biometrischen Systems dargestellt. Es ist zu sehen, dass beide Algorithmen bis nach der Ermittlung des Matching Score autonom arbeiten. Danach werden beide Match Scores im Fusionsprozess normalisiert und miteinander kombiniert. Der entstehende neue Match Score wird dann wiederum dem Decision Modul zur Entscheidungsermittlung zur Verfügung gestellt.

Zur Untersuchung der vier Algorithmen und ihrer Fusion mussten für jedes einzelne Verfahren und deren Kombination Tests durchgeführt werden. Dazu wurden im ersten Schritt Testmengen definiert, die auf identischen Semantik-Klassen und einzelnen Tabletts bzw. Tablett-Gruppen mit ähnlichen physikalischen Eigenschaften basieren. In dieser Arbeit wird hauptsächlich auf die Testergebnisse eingegangen, die auf einem Wacom Cintiq15 Digitalisiertablett aufgezeichnet wurden und in einer Evaluierungsdatenbank persistent hinterlegt wurden. Die Tests wurden dann in einem Evaluierungssystem unter Zugriff auf ebendiese Daten durchgeführt. In den nächsten Abschnitten folgt eine kurze Erläuterung des Biometric Hash Algorithmus und der damit verwendeten Abstandsmaße. Die hier angegeben Abstandsmaße und Abstandsdefinitionen basieren darauf, dass x und y zwei k-dimensionale Vektoren als Instanzen zweier Biometric Hash Berechnungen sind. Dabei besteht ein Vektor aus k Merkmalen (x_1, x_2, …, x_k). Je geringer der Abstand zweier Vektoren, umso ähnlicher sind sie sich.

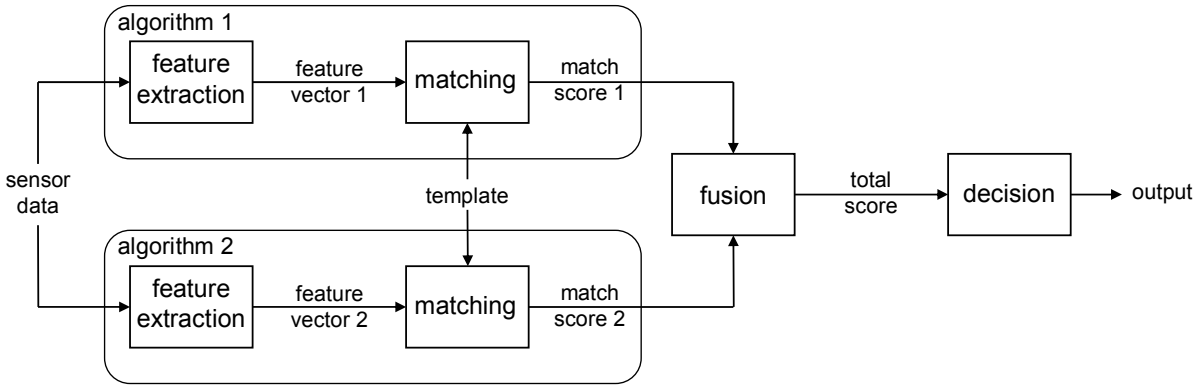

Abb. 4: Fusion zweier biometrischer Algorithmen auf Matching Score Level

4.2 Der Biometric Hash Algorithmus

Der Ansatz des Biometric Hash Algorithmus beruht auf der ursprünglichen Zielsetzung, aus den verschiedenen Signalen, die durch ein Digitalisiertablett während des Schreibprozesses erfasst werden, eine statistische Vektorrepräsentation dergestalt zu gewinnen, dass diese für einen gegebenen Schriftzug stabil durch alle Nutzer eines biometrischen Systems reproduziert werden können. Ein solcher Biometric Hash eignet sich dann beispielsweise zur Generierung individueller kryptographischer Schlüssel. Abbildung 5 zeigt ein vereinfachtes Modell zu diesem Verfahren, wobei ausgehend von fünf verschiedenen Schreibsignalen auf der linken Seite (basierend auf physikalischen Messgrößen zu horizontaler und vertikaler Stiftposition x(t)

und y(t), Stiftdruck p(t) sowie Stiftwinkel Θ(t) und ϕ(t)) zunächst eine Menge von k statistischen Parametern bestimmt wird. Die Größe von k ergibt sich dabei aus der Anzahl der unterschiedlichen statistischen Größen in der Merkmalsmenge $\{n_1, \ldots, n_k\}$, in unseren Untersuchungen finden derzeit k=68 unterschiedliche Merkmale Anwendung.

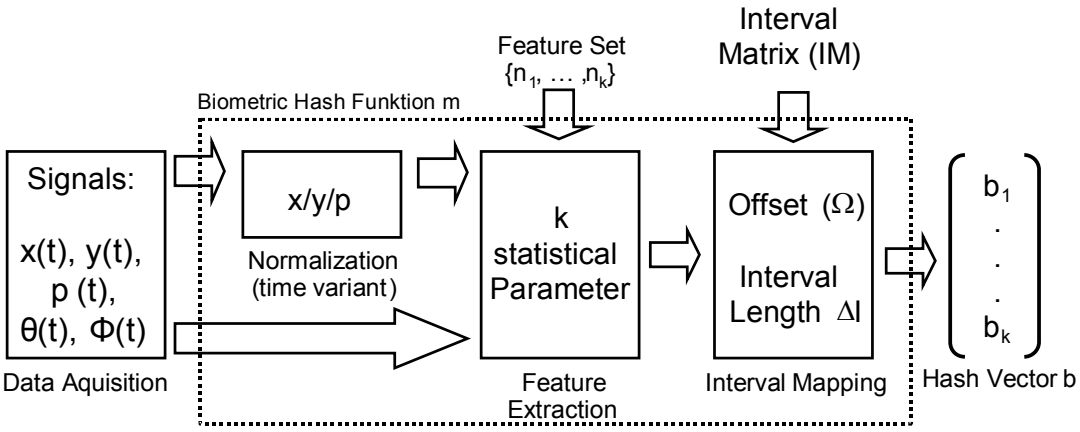

Abb. 5: Modell der Biometric Hash Generierung

Aus den so gewonnenen Merkmalen wird dann über eine so genannte Interval Mapping Funktion der Biometric Hash Vektor b gebildet, wobei dies unter Anwendung eines Verfahrensparameters, der Interval Matrix IM, erfolgt. Weitere Details zum Biometric Hash Verfahren finden sich in [ViSM02], sowie vertiefende Betrachtungen, u.a. zur Eignung des Biometric Hash nicht nur zur Schlüsselgenerierung, sondern auch als Merkmalsvektor zur biometrischen Benutzerauthentifikation in [Viel04]. In den folgenden Abschnitten werden die Bezeichnungen x und y für jeweils zwei Instanzen eines Biometric Hash Vektors b verwendet.

4.3 Hammingdistanz

Bei der Hammingdistanz werden die Elemente x_i und y_i der beiden Vektoren x und y, die sich jeweils an der i-ten Position befinden, mit einander verglichen. Sind sie identisch, ist das Ergebnis des Vergleichs 0, im anderen Fall 1. Im Anschluss werden die Einzelergebnisse summiert. Der Vorteil dieses Verfahrens liegt darin, dass der minimal bzw. maximal mögliche Abstandswert bekannt ist. Das Ergebnis kann Werte von 0 bis zur Anzahl k der untersuchten Features (hier k = 68) reichen. In diesem Fall gilt also:

$$0 \leq hd(x, y) \leq k \, .$$

4.4 City-Block-Distanz

Die City-Block-Distanz (auch Manhattan-Distanz) [Vanl04] orientiert sich an der Vorstellung einer Fahrt durch eine Stadt mit parallelen bzw. rechtwinkligen Straßenzügen. Dabei ist eine Richtungsänderung nur an bestimmten Stellen und im rechten Winkel möglich. Beispielsweise ist es denkbar, die Länge der einzelnen Wegstrecken zu maximieren, während man die Anzahl der Richtungsänderungen minimiert. Im zweidimensionalen Raum würde in diesem Fall die City-Block-Distanz aus einer einmaligen Änderung der Richtung und damit aus einer Waagerechten und einer Senkrechten parallel zur jeweiligen Koordinatenachse ermittelt werden. Berechnet wird die City-Block-Distanz wie folgt:

$$cbd(x,y)=\sum_{i=1}^{k}\left|x_i-y_i\right|.$$

Da nicht im Voraus bekannt ist, in welchem Intervall sich der Abstandswert bewegen wird, ist hier eine Normalisierung notwendig, um die Werte der unterschiedlichen Distanzfunktionen vergleichen zu können.

4.5 Euklidische Distanz

Die Euklidische Distanz [Vanl04] ist die mathematische Beschreibung des direkten Abstandes zwischen zwei Vektoren. Der Euklidische Abstand wird definiert durch:

$$ed(x,y)=\sqrt{\sum_{i=1}^{k}(x_i-y_i)^2}\ .$$

Auch hier ist, wie bei der City-Block-Distanz, vorher nicht bekannt, in welcher Größenordnung sich das Ergebnis bewegen wird. Daher ist auch bei der Euklidischen Distanz eine Normalisierung notwendig, wenn ein Vergleich mit anderen Abstandsfunktionen erfolgen soll.

4.6 Canberra-Distanz

Die Canberra-Distanz [Vanl04] beschreibt nicht nur den Abstand zweier Punkte, sondern auch deren Lage zum Koordinatenursprung. Auch wenn je zwei Vektoren geometrisch den gleichen Abstand haben, ist der Canberra-Abstand unterschiedlich. In einem solchen Fall ist die Canberra-Distanz der beiden Vektoren, die näher am Ursprung liegen, kleiner als die zwischen den beiden anderen Vektoren. Die Canberra-Distanz ist definiert durch:

$$cd(x,y)=\sum_{i=1}^{k}\frac{\left|x_i-y_i\right|}{\left|x_i\right|+\left|y_i\right|}.$$

Ein großer Vorteil der Canberra-Distanz liegt darin, dass der Wert im Bereich von 0 bis k liegt. k entspricht dabei der Anzahl der zu vergleichenden Features. Das heißt, auch hier gilt:

$$0\le cd(x,y)\le k\ .$$

4.7 Wichtungsstrategien

Für die Fusion sind die Abstandsmaße der einzelnen Algorithmen auf das Intervall [0, …, k] normalisiert worden, wobei sich k = 68 aufgrund des aktuellen Entwicklungsstandes des Biometric Hash Algorithmus ergibt [ViSt04]. Danach wurden die einzelnen Werte nach unterschiedlichen Regeln gewichtet.

An dieser Stelle werden verschiedene Wichtungsstrategien vorgestellt, welche die Werte der einzelnen Algorithmen entsprechend bestimmter Gesichtspunkte (z.B. Größenverhältnis zueinander) vor der Fusion zu bewerten. Dabei werden folgende Variablen verwendet:

Match Scores: $s_1, s_2, \ldots, s_n$

Gewichte: $w_1, w_2, \ldots, w_n$.

In den nächsten Abschnitten werden die insgesamt fünf Strategien zur Wichtung beschrieben. Die hier vorgestellten Taktiken sind nur einige mögliche Verfahren. Es wird nicht der Anspruch erhoben, dass die Auswahl vollständig oder die best mögliche ist.

Binär gewichtete Fusion

Auch eine Fusion auf Decision Level wurde durchgeführt. Diese wurde jedoch durch eine entsprechende Wichtung auf dem Matching Score Level simuliert. Die Entscheidung des Systems beruht in diesem Fall auf der Verwendung des besten einzelnen Abstandswertes. Das Gewicht kann entweder 1 für den besten Algorithmus oder 0 für die anderen sein. Das bedeutet für die Fusion und die Gewichte:

Bedingungen: $w_1 + w_2 + \ldots + w_n = 1$

$w_i = 1$, *wenn* $s_i = max(s_1, s_2, \ldots, s_n)$

$w_i = 0$, *sonst*

Fusion: $s_{fus} = w_1 s_1 + w_2 s_2 + \ldots + w_n s_n = s_{max}$.

Das heißt also, dass nur der Match Score des Algorithmus weitergegeben wird, der in der Testphase die geringste Equal Error Rate hatte.

Gleich gewichtete Fusion

Die zweite Wichtungsstrategie ist eine lineare Gleichgewichtung der Match Scores s_i der einzelnen Algorithmen. Das bedeutet, dass die einzelnen Gewichte w_i als gleich groß festgelegt wurden und deren Summe 1 betrug. Danach wurden die einzelnen Match Scores mit den entsprechenden Gewichten multipliziert und dann aufsummiert. Nachfolgend sind die einzelnen Voraussetzungen und Regeln für die gleich gewichtete Fusion noch einmal aufgeführt:

Bedingungen: $w_1 + w_2 + \ldots + w_n = 1$

$w_1 = w_2 = \ldots = w_n = n^{-1}$

Fusion: $s_{fus} = w_1 s_1 + w_2 s_2 + \ldots + w_n s_n$.

Die Ergebnisse für die einzelnen Algorithmen über die unterschiedlichen Tabletts bzw. Semantik-Klassen wurden zur Bestimmung einer nicht linearen Gewichtung der Algorithmen im nächsten Schritt herangezogen. Dazu wurde für jedes Tablett und jede Semantik-Klasse eine Rangfolge der Algorithmen bestimmt, die sich durch die Größe der Equal Error Rate ergab.

Linear gewichtete Fusion 1

Bei der ersten linearen Wichtungsstrategie wird der beste Algorithmus in Abhängigkeit zum schlechtesten Algorithmus gewichtet. Das bedeutet, je größer die Equal Error Rate des schlechtesten Algorithmus, umso größer ist das Gewicht für den besten Algorithmus.

Im ersten Schritt werden die Equal Error Rates der Algorithmen der Größe nach sortiert. Dann werden die einzelnen Gewichte nach folgender Formel berechnet:

$$w_i = \frac{eer_i}{\sum_{m=1}^{n} eer_m} .$$

Im letzten Schritt werden die ermittelten Gewichte in absteigender Reihenfolge den aufsteigend sortierten EERs zugeordnet. Für die Fusion ergeben sich damit die folgenden Bedingungen:

Bedingungen: $w_1 + w_2 + \ldots + w_n = 1$

$$eer_{w1} < eer_{w2} < \ldots < eer_{wa}$$
$$eer_{s1} > eer_{s2} > \ldots > eer_{sb}$$

Fusion: $s_{fus} = w_m s_1 + w_{m-1} s_2 + \ldots + w_2 s_{b-1} + w_1 s_b$.

Damit ist die Bewertung eines Algorithmus nicht direkt von seiner eigenen EER abhängig, sondern von seinem Rang und der EER des korrespondierenden Algorithmus.

Linear gewichtete Fusion 2

Die Gewichte der zweiten linearen Fusion berechnen sich aus dem Verhältnis der Equal Error Rates zueinander. Die einzelnen Gewichte werden so normiert, dass ihre Summe 1 ergibt. Das heißt, das Gewicht für Algorithmus i berechnet sich aus der Summe der EERs der übrigen Algorithmen dividiert durch die Summe aller EERs. Das Ergebnis wird dann durch die Anzahl der übrigen Algorithmen dividiert. Daraus ergibt sich für die zweite linear gewichtete Fusion:

Bedingungen: $w_1 + w_2 + \ldots + w_n = 1$

$$w_i = \frac{\left(\sum_{j=1}^{n} eer_j\right) - eer_i}{\sum_{j=1}^{n} eer_j} \cdot \frac{1}{(n-1)}$$

Fusion: $s_{fus} = w_1 s_1 + w_2 s_2 + \ldots + w_n s_n$.

In diesem Fall stehen also die einzelnen Gewichte in einem direkten Verhältnis zur Größe der entsprechenden Equal Error Rates.

Quadratisch gewichtete Fusion

Bei den ersten Fusions-Tests stellte sich heraus, dass die meisten Ergebnisse der ersten linearen Wichtungsstrategie besser waren als die der zweiten. Andererseits fiel jedoch auch auf, dass der Unterschied zwischen zwei Gewichten sehr groß war, auch wenn die dazu gehörenden Equal Error Rates dicht beieinander lagen. Dies war immer der Fall, wenn die EERs der korrespondierenden Algorithmen sehr viel größer waren. Um diesem Umstand Rechnung zu tragen, wurde eine weitere Wichtungstaktik eingeführt. Diese quadriert die Gewichte der linearen Wichtung 1 ($w_{linear1x}$) und normiert sie so, dass sich als Summe aller Gewichte wieder 1 ergibt.

Bedingungen: $w_1 + w_2 + \ldots + w_n = 1$

$$w_i = \frac{w_{li}^{\,2}}{\sum_{j=1}^{n} w_{lj}^{\,2}} \quad mit \quad w_{lx} = w_{linear1x}$$

Fusion: $s_{fus} = w_1 s_1 + w_2 s_2 + \ldots + w_n s_n$.

5 Testergebnisse

Bei der Vorstellung der Testergebnisse beschränken wir uns auf die Präsentation der Resultate von Schriftproben die alle auf einem einzigen Typ von Digitalisiertabletts aufgezeichnet wurden, dem Wacom Cintiq15. Dabei wird auf den Vergleich der Verifikation mit einem zufälligen bzw. mit einem Brute Force Angriff eingegangen. Diese stellen zwei extreme Einstellungen eines biometrischen Systems dar. Zum einen sind die Verifikation und der zufällige Angriff (Test A/B) eine Annahme, bei der davon ausgegangen wird, dass das System nur mit echten Authentifizierungsversuchen berechtigter Benutzter konfrontiert wird. Auf der anderen Seite handelt es sich bei den Brute Force Angriffen (Test A/C3) ausschließlich um Fälschungen, die mit dem höchstmöglichen Wissen über das Original angefertigt wurden. Die hier beschriebenen beiden Fälle sind aber nur theoretische Szenarien, die bei realen biometrischen Systemen nicht vorkommen. Ein reales System wird sich in der Regel dazwischen befinden, idealer Weise möglichst dicht am besten Fall.

Neben Evaluierungen basierend auf Unterschriften als Schreibsemantik basiert unsere Testmethodologie zusätzlich auf alternativen Schreibsemantiken Passphrase (pas), „Sauerstoffgefäß“ (sau), PIN (871) und Symbol (sym), siehe [Viel00] und [Viel04]. Die Darstellung der Ergebnisse erfolgt in zwei Schritten. Zunächst soll in einer globalen Betrachtung für die betrachteten Abstandsfunktionen die Erkennungsgenauigkeit anhand der EER und separat je Semantikklasse untersucht werden. Aus dieser Beobachtung werden dann für die zweite Betrachtung verschiedene Parametrisierungen der im vorigen Abschnitt diskutierten Wichtungsverfahren abgeleitet. Dabei beschränken wir uns in diesem Beitrag auf die Semantikklasse, für die die größte Zahl an Schriftsätzen in unserer Datenbank zur Verfügung steht, der Signatur. Insgesamt basieren unsere Testdaten auf einer Erhebung im studentischen Umfeld (22 Studenten, insgesamt 1761 Enrollments, 1101 Verifikationen und 431 Brute Force Fälschungen) und bilden durch Beschränkung auf eine Tablettkategorie eine Teilmenge der in [Viel2004] verwendeten Datenbank.

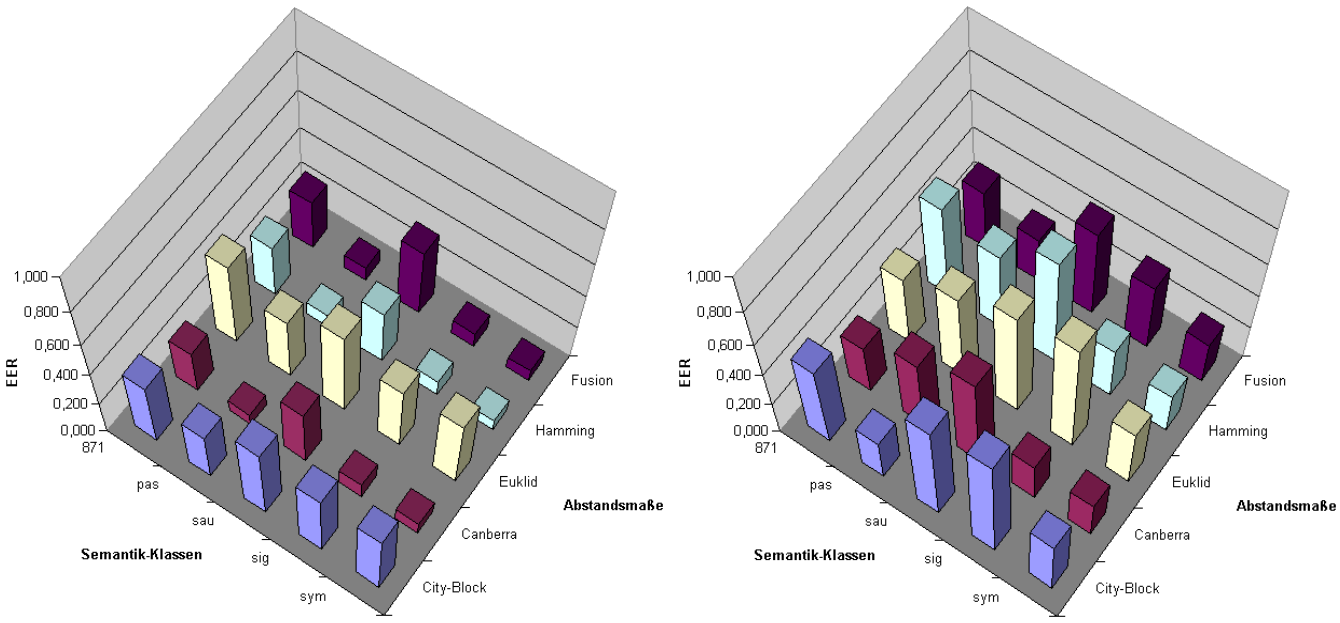

Abb. 6: Darstellung der EERs für Verifikation (links) und Brute Force Angriff (rechts)

Die Abbildung 6 stellt für die erste globale Betrachtung die Größenverhältnisse zwischen den ermittelten EERs der Algorithmen bzw. Semantiken jeweils für die Verifikation und den Brute Force Angriff dar. Diese grafische Darstellung gibt auch einen Überblick über die unterschiedliche Eignung der einzelnen Algorithmen für bestimmte Semantik-Klassen. Beispielsweise sieht es so aus, als ob die Canberra- und die Hamming-Distanz sehr gut für die Semantiken Signatur und Symbol auf dem Wacom Cintiq15 geeignet sind.

Für die zweite Betrachtung sind in Tabelle 1 und Tabelle 2 die aus dem ersten Test ermittelten Gewichte für die untersuchten Wichtungsstrategien und die im zweiten Test ermittelten Equal Error Rates für die Verifikation (Test A/B, Tabelle 1) bzw. den Brute Force Angriff (Test A/C3, Tabelle 2) für die Semantik Signatur aufgeführt. Hier ist deutlich zu sehen, dass durch die quadratische Fusion bei der Verifikation eine Verbesserung gegenüber dem besten Einzelalgorithmus erreicht werden konnte. Dabei konnte im ersten Fall die EER der Canberra-Distanz von 0,091 (erste Zeile, Spalte „Canberra“) durch die quadratische Fusion auf 0,080 verbessert werden (rechts unten in Tabelle 1). Bei den Brute Force Angriffen ist durch die Fusion eine Verschlechterung um ca. 65% gegenüber dem besten Einzelergebnis eingetreten.

Tab. 1: Cintiq15 – Signatur – Test A/B: Gewichte und EERs für Trainingsdaten und Testdaten

Algorithmus		City-Block	Canberra	Euklid	Hamming	EER
EER		0,388	0,091	0,376	0,092	Fusion
Gewichte	binär	0,000	1,000	0,000	0,000	0,091
	gleich	0,250	0,250	0,250	0,250	0,276
	linear1	0,096	0,410	0,097	0,397	0,122
	linear2	0,197	0,301	0,201	0,301	0,235
	quadratisch	0,027	0,488	0,027	0,458	0,080

Tab. 2: Cintiq15 – Signatur – Test A/C3: Gewichte und EERs für Trainingsdaten und Testdaten

Algorithmus		City-Block	Canberra	Euklid	Hamming	EER
EER		0,642	0,230	0,655	0,301	Fusion
Gewichte	binär	0,000	1,000	0,000	0,000	0,230
	gleich	0,250	0,250	0,250	0,250	0,512
	linear1	0,165	0,358	0,126	0,351	0,465
	linear2	0,216	0,291	0,214	0,278	0,525
	quadratisch	0,092	0,436	0,054	0,419	0,379

Weitere Tests über andere Semantik-Klassen haben ergeben, dass der beste EER Wert auch hier nicht immer von ein und demselben Algorithmus ermittelt wird.

Tab. 3: Verteilung der Algorithmen mit der besten EER entsprechend der Semantik-Klassen

Semantik	Test A/B		Test A/C3	
	bester Algorithmus	EER	bester Algorithmus	EER
8710 (PIN)	Canberra	0,253	Canberra	0,289
Passphrase	Canberra	0,073	City-Block/ Fusion	0,250
Sauerstoffgefäß	Hamming	0,312	Fusion	0,494
Signature	Fusion	0,080	Canberra	0,230
Symbol	Hamming / Fusion	0,064	Canberra	0,228

In Tabelle 3 ist eine Aufstellung der jeweils besten Abstandsfunktionen für alle fünf untersuchten Semantiken auf dem Wacom Cintiq15 gegeben. Die Fusion basiert dabei auf der quadratischen Fusionsstrategie.

Dabei fällt auf, dass in vier der zehn untersuchten Szenarien die Fusion das beste Ergebnis ermittelte. Besser ist hier die Canberra-Distanz mit fünf besten Ergebnissen. Dies legt nahe, dass der Algorithmus, der auf dem Canberra-Abstand basiert, für ein zu realisierendes System auf diesen Daten die beste Wahl wäre. Zieht man die binäre Fusion in Betracht, so erhält man in acht Fällen den besten Wert. Dabei ist aber auch zu beachten, dass sich diese Erkenntnisse auf die bereits vorhandenen Daten in der Datenbank stützen. Sollten neue Daten hinzukommen müssen die Gewichte neu bestimmt oder untersucht werden, ob die Ergebnisse weiterhin zufrieden stellend sind. Im zweiten Fall kann es natürlich auch vorkommen, dass ein anderer Algorithmus auf der neuen Datenzusammenstellung besser wäre. Gewählt wird jedoch der Algorithmus, der vorher das beste Ergebnis bestimmt hat, da er durch die binäre Fusion bevorzugt wurde. Als günstigere Lösung könnte sich hier die quadratische Fusion herausstellen. Dabei könnte es zwar im Durchschnitt aller Nutzer zu etwas schlechteren EERs kommen, aber durch die entsprechenden Gewichte trägt der sonst gewählte einzelne Algorithmus nicht mehr zu 100% zum Endergebnis bei. Vielmehr wird das Ergebnis durch die Fusion aller Algorithmen herbeigeführt.

6 Zusammenfassung und Ausblick

Die Tests der einzelnen Biometric Hash Algorithmen zur handschriftbasierten Authentifizierung haben ergeben, dass die Algorithmen, die auf der Hamming- bzw. Canberra-Distanz basieren, vor allem für die Verifikation gut geeignet sind. Der Vorsprung zur Euklidischen und zur City-Block-Distanz ist hier in den meisten Fällen sehr groß. Dieser verringert sich bei den Brute Force Angriffen etwas. Bei der Untersuchung der Fusion von vier Algorithmen zur Authentifikation stellte sich heraus, dass eine Verbesserung gegenüber den einzelnen Algorithmen möglich ist. Die Verbesserung ist dabei abhängig von der verwendeten Wichtungsstrategie. Durch die quadratisch gewichtete Fusion ist bei vier von zehn Testszenarien auf dem Wacom Cintiq15 eine Verbesserung zu beobachten gewesen.

In zukünftige Untersuchungen können weitere Abstandsmaße, wie zum Beispiel die Mahalanobis- oder Korrelationsdistanz, einbezogen werden. Ebenfalls denkbar ist die Anpassung der verwendeten statistischen Merkmale für jeden einzelnen Algorithmus. So kann man beispielsweise diejenigen Merkmale unterdrücken, die das Ergebnis eines Algorithmus zu stark verfälschen. Natürlich liegt es ebenfalls nahe, auch andere Verifikationsalgorithmen, die nicht auf Abstandsmaßen basieren, zu kombinieren bzw. in die Kombination mit einzubeziehen. Gegebenenfalls muss dann über die Änderung der Fusionsstrategie, zum Beispiel durch Wahl eines anderen Fusionslevels, nachgedacht werden.

Weitere Verbesserungen durch die Fusion sind denkbar, wenn man die Gewichte nicht global für eine ganze Semantik-Klasse bzw. Tablettkategorie ermittelt. Stattdessen erscheint eine Herangehensweise plausibel, bei der die Enrollments und Verifikationen einzelner Personen beobachtet werden und daraus automatisch individuelle Gewichte für jede einzelne Person bestimmt werden (adaptive Wichtung). Für bereits vorhandene Nutzer wäre dann ein einmaliger Durchlauf der Datenbank zur Bestimmung persönlicher Gewichte notwendig. Bei erneuter Verifikation können dann die Wichtungen automatisch weiter angepasst werden. Diese Vorgehensweise wäre für den Fall erfolgreich, wenn sich herausstellen sollte, dass die Schreibweise einzelner Personen bzw. Personengruppen die Erkennungsgenauigkeit der Algorithmen beeinflusst. Dies muss jedoch zunächst durch weitere Tests untersucht und belegt werden.

Darüber hinaus eröffnen die Arbeitsergebnisse das Potenzial für multimodale biometrische Verfahren, da die untersuchten Abstandsmaße auch künftig für andere biometrische Merkmale, wie z.B. den Iris-Code ([Daug03]) eingesetzt und kombiniert werden können.

Literatur

[CzKV04] J. Czyza, J. Kittler, L. Vandendorpe: Multiple classifier combination for face-based identity verification, Pattern Recognition 37, Elsevier (2004) 1459-1469.

[Daug03] J. Daugman: The importance of being random: Statistical principles of iris recognition, Pattern Recognition, Vol. 36, No. 2 (2003) 279-291.

[GuCa97] J. Gupta, A. McCabe: A Review of Dynamic Handwritten Signature Verification. James Cook University, Australia, 1997, cite-seer.nj.nec.com/ gupta97 review.html

[JaRo04] A.K. Jain, A. Ross: Multibiometric Systems, Communications of the ACM; January 2004, Vol. 47, No. 1.

[Kais01] J. Kaiser: Vertrauensmerkmal Unterschrift – Gestaltungskriterien für sichere Signierwerkzeuge. In: Informatik 2001 – Tagungsband der GI/OCC-Jahrestagung, 25. - 28. September 2001, 2001,. ISBN 3-85403-157-2, S 500-504.

[Laßm02] G. Laßmann (Ed.): Bewertungskriterien zur Vergleichbarkeit biometrischer Verfahren (Kriterienkatalog), Version 2.0, TeleTrusT Deutschland e.V., 07/2002.

[PeSS03] T. Petermann, C. Scherz, A. Sauter: Biometrie und Ausweisdokumente, TAB Arbeitsbericht 93, http://www.tab.fzk.de/de/projekt/zusammenfassung/ ab93.pdf, 2003.

[PlLe94] R. Plamondon, F. Leclerc: Automatic Verification and Writer Identification: The State of the Art 1989-1993. International Journal of Pattern Recognition and Artificial Intelligence, 1994, S. 8:643-660

[VanL04] K. van Laerhoven: Basic Statistics and Metrics for Sensor Analysis; http://www.comp.lancs.ac.uk/~kristof/research/notes/basicstats/ (26.08.2004).

[Viel00] C. Vielhauer: Handschriftliche Authentifikation für digitale Wasserzeichenverfahren. In: M. Schumacher, R. Steinmetz: Sicherheit in Netzen und Medienströmen, Berlin: Springer Verlag, 2000, S. 134-148

[Viel04] C. Vielhauer: Handwriting Biometrics for User Authentication: Security Advances in Context of Digitizer Characteristics, Dissertationsschrift, Fachbereich Elektrotechnik und Informationstechnik, Technische Universität Darmstadt, 2004

[ViSM02] C. Vielhauer, R. Steinmetz, A. Mayerhöfer: Biometric Hash based on Statistical Features of Online Signature Proceedings of the International Conference on Pattern Recognition (ICPR); Conference on Pattern Recognition (ICPR), August, Quebec City, Canada, ISBN 0-7695-1696-3, 2002 Vol. 1, S. 123-126.

[ViSt03] C. Vielhauer; R. Steinmetz: Handschriftliche biometrische Signaturen. In P. Horster (Hrsg.): DACH Security IT Security & IT Management, DACH Security, 25.-26.03., Erfurt, Germany, ISBN 3-00-010941-2, 2003, S. 344-353

[ViSt04] C. Vielhauer; R. Steinmetz: Handwriting Feature Correlation Analysis for Biometric Hashes, In Bourlard, H. ; Pitas, I.; Lam, K.; Wang, Y.: EURASIP Journal on Applied Signal Processing, Special Issue on Biometric Signal Processing, Hindawi Publishing Corporation, Sylvania, OH, U.S.A., ISSN 1110-8657, 2004, S. 542-558.

[ZoVi03] F. Zöbisch; C. Vielhauer: A Test Tool to support Brut-Force Online and Offline Signature Forgery Tests on Mobile Devices, In: Proceedings of the IEEE International Conference on Multimedia and Expo 2003 (ICME), Baltimore, MD, U.S.A., Vol. 3, ISBN 0-7695-1062-0, 2003, S. 225-228.