

Fachhochschule Aachen



Inhaltsbasierte Bildsuche mit Hilfe von Daubechies Wavelets

Diplomarbeit

zur Erlangung des Grades eines
Diplom-Mathematikers der Physikalischen Technik
Fachrichtung: Technomathematik

Janina Peslak

Betreuender Professor

Prof. Dr. W. Hillen

Band 03, Jahr 2005

ISSN 1860-8906

ISBN 3-9810089-2-8

Aachener Schriften zur Medizinischen Informatik
ISSN 1860-8906
ISBN 3-9810089-?-?

Herausgeber: Institut für Medizinische Informatik der RWTH Aachen
Pauwelsstr. 30
D-52074 Aachen

Geschäftsführender Direktor: Univ.-Prof. Dr. Dr. Klaus Spitzer

Kurzfassung

Oft sind Bilddatenbanken so angelegt, dass zu jedem Bild eine alphanumerische Beschreibung existiert und bei einer Anfrage an die Datenbank ein Textvergleich durchgeführt wird. Es ist von Vorteil statt einer textuellen Beschreibung zur Bildsuche intrinsische Eigenschaften des Bildes in Form von abstrakten Merkmalen zu verwenden, die direkt aus den Bilddaten berechnet werden. Auf Basis dieser Merkmalsvektoren kann eine Datenbank nach Bildern durchsucht werden, die einem vorgegebenen Bild möglichst ähnlich sind. Ziel der Diplomarbeit ist es, einen Algorithmus zu entwerfen, der mit Hilfe der Wavelet-Transformation aus einem vorgegebenen Bild den Merkmalsvektor berechnet und mit den a-priori berechneten Vektoren der Bilder in der Datenbank vergleicht.

Wavelets sind Funktionen, die Signale (hier Bilder) in verschiedene Frequenzkomponenten zerlegen. Im Gegensatz zur Fourier-Reihe, sind sie jedoch nicht periodisch, so dass auch Aussagen über das räumliche Auftreten einer Frequenz gemacht werden können. Durch Dehnung und Stauchung der Wavelet-Funktionen ist es möglich, das Bildsignal in verschiedenen Auflösungsstufen zu betrachten, so dass sowohl tiefe, als auch hohe Frequenzen korrekt erfasst werden können.

Zur Berechnung des Merkmalsvektors wird das eingelesene Bild zunächst auf eine zuvor festgelegte Größe normiert. Dann wird auf das Bild wiederholt eine Wavelet-Transformation mit Hilfe von Daubechies Wavelets angewandt. Durch die Wavelet-Transformation werden hochfrequente Bildanteile von den niederfrequenten getrennt. Auf Basis dieser Trennung findet dann eine Extraktion der für die Bildsuche genutzten Daten in Form eines Merkmalsvektors statt.

Die Merkmalsvektoren werden anhand von Distanzmaßen und einer Korrelationsfunktion verglichen.

Das Verfahren wurde für die Anwendung im medizinischen Bildretrieval optimiert und evaluiert. Die Tests wurden dazu als Anfragen auf einen vorklassifizierten Bildbestand durchgeführt. Auf diese Weise kann ein Retrieval-Test als „Leaving-One-Out-Anfrage“ auf dem bestehenden Datensatz durchgeführt werden.

Bei der Kategorisierung wurde eine Erfolgsrate von ca. 75 Prozent korrekt zugeordneter Bilder erreicht. Dies entspricht einer leichten Verbesserung gegenüber ähnlichen Verfahren, die keine Wavelet-Transformation zur Merkmalsextraktion nutzen.



Inhalt

1	Einleitung	1
2	Grundlagen und Umgebung	2
2.1	Inhaltsbasierter Bildzugriff	2
2.2	Das Projekt „Image Retrieval in Medical Applications“ (IRMA) ..	2
2.3	Der IRMA-Code.....	3
3	Aufgabenstellung	4
4	Verfahren.....	5
4.1	Die Wavelet-Transformation	6
4.2	Vorverarbeitung	11
4.3	Bildung des Merkmalsvektors.....	11
4.4	Klassifikation.....	14
4.4.1	Entscheidungsfindung	14
4.4.2	Klassifikatoren	14
5	Auswertung.....	20
5.1	Evaluationssystem	20
5.2	Durchführung	22
5.3	Ergebnisse.....	23
5.3.1	Visuelle Auswertung	23
5.3.2	Systematik der Auswertung.....	25
5.3.3	Klassifikatorvergleich.....	28
5.3.4	Ergebnisse für die unterschiedlichen Vektorgrößen	29
5.3.5	Laufzeiten	35
5.4	Bewertung	37
6	Diskussion und Fazit.....	39
7	Literaturverzeichnis	40
	Erklärung.....	41

1 Einleitung

In der medizinischen Diagnostik sind zahlreiche bildgebende Verfahren im Einsatz. Hierzu gehören z.B. Röntgenaufnahmen, Radiographien oder Tomographien. Im Zuge der zunehmenden Digitalisierung werden medizinische Bilddaten in Datenbanken verwaltet. Die Pflege dieser Datenbanken erfordert naturgemäß einen großen Aufwand, da sie in klinisch relevante Kategorien eingeteilt und die Bilder in diese eingeordnet werden müssen.

Das Projekt „Image Retrieval in Medical Applications“ (IRMA) [1] des Instituts für medizinische Informatik der RWTH Aachen, befasst sich mit der Kategorisierung medizinischer Bilddaten mit Mitteln der automatischen Bilderkennung. Hierdurch wird insbesondere die inhaltsbasierte Bildsuche ermöglicht, bei der eine Anfrage an die Datenbank nicht mit Hilfe einer Texteingabe durchgeführt wird, sondern durch ein Anfragebild zu welchem möglichst ähnliche Bilder gesucht werden.

Die Bilderkennung wird hierbei mit Verfahren der globalen Bildinhaltsextraktion und Verfahren zum Vergleich der extrahierten Merkmale erreicht.

Das im Paper „Content based image indexing and searching using Daubechies Wavelets“[2] beschriebene Verfahren nutzt die Wavelet-Transformation, um aus einem Bild globale, inhaltsbeschreibende Merkmale zu extrahieren, welche dann für die Bildsuche oder Kategorisierung verwendet werden. Die Implementierung dieses Verfahrens und das Einfügen des Verfahrens in die vorhandene „Image Retrieval in Medical Applications“-Architektur ist Aufgabe dieser Diplomarbeit. Hierbei gilt es zu untersuchen, ob dieses für allgemein Bilddaten erdachte Verfahren für die medizinische Bildklassifikation genutzt werden kann.

Im Projekt sind bereits Verfahren implementiert, die basierend auf einem verkleinerten Originalbild eine Abstandsberechnung mit Hilfe des euklidischen Abstands bzw. einer Korrelationsfunktion durchführen und anhand dieser Abstände zu einer Bildklassifikation gelangen.

Diese Verfahren stellen eine einfachere Version des zu implementierenden Verfahrens dar, da sie zwar dieselben Funktionen zur Distanzberechnung verwenden, allerdings zur Merkmalsextraktion nicht die Wavelet-Transformation nutzen. Eine Gegenüberstellung soll klären, ob sich gegenüber diesen Verfahren eine Verbesserung in der Bildklassifikation erreichen lässt. Hierbei wird vermutet, dass sich eine Zunahme der korrekt zugeordneten Bilder einstellt. In Kapitel 2 werden zunächst die Grundlagen der inhaltsbasierten Bildsuche herläutert, und es wird ein Einblick in das IRMA-Projekt gegeben. Hierbei wird auch der im IRMA-Projekt genutzte Code zur Kategorisierung vorgestellt. In



Kapitel 3 folgt eine detaillierte Darstellung der Aufgabenstellung. Kapitel 4 beschreibt das implementierte Verfahren sowie seine Grundlagen, insbesondere die Wavelet-Transformation. In Kapitel 5 werden die Auswertungen beschrieben, wobei zunächst die Durchführung erläutert wird und dann die Ergebnisse dargestellt und schließlich bewertet werden. Kapitel 6 fasst die Ergebnisse sowie ihre Bedeutung noch einmal zusammen, so dass schließlich ein Fazit geschlossen werden kann.

2 Grundlagen und Umgebung

2.1 Inhaltsbasierter Bildzugriff

Diese Diplomarbeit behandelt ein Verfahren zur inhaltsbasierten Bildsuche, welche die Suche von Bildern in einer Datenbank ermöglicht, ohne dass zu jedem Bild ein beschreibender Text vorhanden sein muss, auf welchen sich die eigentliche Suche stützt. Stattdessen beruht die Bildsuche auf Methoden des Bildvergleichs bzw. der Bildinterpretation mit Hilfe von Informationen, die unmittelbar aus den Bilddaten gewonnen werden. Hierzu werden formale Merkmale eines Bildes extrahiert. Anhand dieser Merkmale kann das Bild interpretiert werden. Die Merkmale werden genutzt, um einen Vergleich mit den Bildern der Datenbank durchzuführen. Der Vergleich zwischen zwei Bildern findet somit auf der abstrakten Ebene zwischen zwei Merkmalsvektoren und nicht zwischen den gesamten Bilddaten statt. Hierzu werden verschiedene Distanz- oder Ähnlichkeitsmaße genutzt. Typische für die inhaltsbasierte Bildsuche verwendete Merkmale sind Farbe, Textur oder Form.

Die inhaltsbasierte Bildsuche bietet den Vorteil, dass die Interpretation und Einordnung der Bilder in die Datenbank automatisiert werden kann und keine aufwändige und fehleranfällige manuelle Beschreibung der Bilder mehr nötig ist. Die textuelle Beschreibung von Bilddaten hat neben dem Zeitaufwand den Nachteil, dass sie nicht eindeutig ist. Für das gleiche Bild können unterschiedliche Beschreibungen existieren.

2.2 Das Projekt „Image Retrieval in Medical Applications“ (IRMA)

Die vorliegende Arbeit wurde im Rahmen des IRMA-Projekts ausgeführt. Das IRMA-Projekt beschäftigt sich mit der Entwicklung eines Verfahrens der inhaltsbasierten Bildsuche speziell für medizinische Bilddaten. Zum Testen des Verfahrens, aber auch zur automatischen Kategorisierung der Daten steht eine Referenzdatenbank mit klassifizierten Bildern aus der klinischen Routine zur Verfügung. Eine genauere Beschreibung des Klassifizierungsschemas ist in 2.3 gegeben.

Die im IRMA-Projekt angewendeten Verfahren basieren auf einer schrittweisen Bildanalyse, die auf unterschiedlichen Abstraktionsebenen durchgeführt wird. Die Ebenen sind hierarchisch angeordnet, wobei die unterste Schicht das eigentliche Bild repräsentiert und in den nachfolgenden Schichten eine fortwährende Abstraktion stattfindet. Nach einer globalen Bildinhaltsextraktion werden zur Bildinterpretation Vergleichsverfahren genutzt. Hierfür stehen die Bilder der Referenzdatenbank sowie Distanz- und Ähnlichkeitsmaße zur Verfügung.

Der erste Schritt der Bildanalyse ist die Kategorisierung. Hierbei wird das Bild entsprechend dem IRMA-Code (vgl. Kap. 2.3) in eine passende Kategorie eingeordnet. Der IRMA-Code ermöglicht eine eindeutige Einteilung aller medizinischen Bilddaten.

Im zweiten Schritt findet eine geometrische Registrierung der Bilddaten statt. Durch Rotation, Translation oder Skalierung werden hierbei Transformationsparameter generiert. Die Registrierung beruht auf manuell festgelegten Prototypen. Im nächsten Schritt, der Merkmalsextraktion, werden für jedes Pixel Merkmalswerte gewonnen. Diese können z.B. durch Kantenextraktion oder Texturanalyse erzeugt werden. Dann findet eine Auswahl dieser Merkmale statt. Hierbei werden verschiedene Merkmalskombinationen zusammengestellt.

Der letzte Schritt fasst ähnliche Bildregionen zu Ellipsen zusammen. Hierbei können verschiedene Auflösungsstufen gewählt werden. Bei einer feinen Auflösung besteht das Bild aus vielen Ellipsen, während eine gröbere Auflösung mehrere der kleinen Ellipsen zu einer Region zusammenfasst. Die Darstellungen in verschiedenen Auflösungen ergeben somit eine Baumstruktur [1]. Das in dieser Arbeit vorgestellte Verfahren kann auf der Ebene der Kategorisierung eingeordnet werden.

Das IRMA Projekt bietet ein komplettes Anfragesystem mit Datenbank, Computerverbund und Webserver. Neben den Bildern sind in der Datenbank zu jedem Bild seine Merkmale sowie Informationen über den hierarchischen Aufbau dieser Merkmale, welcher sich aus den verschiedenen Stufen der Bildinterpretation ergibt, gespeichert. Außerdem enthält sie die verschiedenen Programme zur Merkmalsextraktion und administrative Informationen über den Computerverbund. Der Webserver bietet eine graphische Benutzeroberfläche, mit welcher Anfragen an die Datenbank gestellt, Ergebnisse betrachtet oder neue Einträge eingefügt werden können.

2.3 Der IRMA-Code

Der IRMA-Code ist ein Codierungssystem für medizinische Bilder, welches eine eindeutige Kategorisierung ermöglicht [3]. In dieser



Arbeit werden Bilder anhand des IRMA-Codes in festgelegte Klassen eingeordnet.

Der IRMA-Code besteht aus vier Teilen.

Im ersten Teil wird die Aufnahmetechnik beschrieben, der zweite Teil beschreibt die Orientierung des Körpers, der dritte Teil zeigt die aufgenommene Körperregion und der vierte, welches biologische System betrachtet wurde. Der gesamte Code besteht aus höchstens 13 Zeichen: TTTT-DDD-AAA-BBB

(T = technical, D = directional, A = anatomical, B = biological).

Je mehr Zeichen in einem Teil verwendet werden, desto genauer ist die Beschreibung. Bei der aufgenommenen Körperregion bedeutet z.B. eine 3 „Wirbelsäule“ eine 31 „Halswirbelsäule“ und eine 311 „dens“ (Fortsatz des zweiten Halswirbels). Eine null bedeutet „unspezifiziert“ und zeigt das Ende der Beschreibungstiefe im entsprechenden Teil. Durch diesen streng hierarchischen Aufbau wird es möglich, eine eindeutige Kategorisierung medizinischer Bilder automatisch durchzuführen bzw. eine solche auf Korrektheit zu überprüfen. Die entsprechende Zuordnung aller im IRMA Datenbestand enthaltenen Bilder wurde von erfahrenen Radiologen durchgeführt.

3 Aufgabenstellung

Ziel dieser Arbeit ist es, das in [2] dargestellte Verfahren (vgl. Kap. 4) zu implementieren und speziell für die Anwendung im IRMA-Projekt also für medizinische Bilddaten zu evaluieren. Die Grundlage des Verfahrens ist die Wavelet-Transformation. Aus den transformierten Bilddaten wird dann ein Merkmalsvektor gebildet. Auf Basis dieser Merkmalsvektoren findet ein Bildvergleich statt, welcher mit Hilfe von Distanzmaßen realisiert wird.

Das Verfahren muss in die Architektur des IRMA-Projekts eingegliedert werden. Es fällt in die Kategorisierungsschicht. Mit seiner Hilfe soll es also möglich sein, Anfragebilder in eine dem IRMA-Code entsprechende Kategorie einzuordnen. Dies wird mit Hilfe von Referenzbildern ermöglicht, welche bereits einen gültigen IRMA-Code aufweisen. Während sich das in [2] beschriebene Verfahren auf allgemeine Bilddaten bezieht, bei welchen auch die Farbe ein leicht zu verwendendes Unterscheidungsmerkmal darstellt, wird in dieser Arbeit allein die Anwendbarkeit des Verfahrens auf medizinische Schwarz-Weiß Bilder untersucht.

Die Schwierigkeiten bei der medizinischen Kategorisierung liegen insbesondere darin, dass zwei identisch codierte Bilder mitunter optisch voneinander abweichen. So können zwei Röntgenthoraxe unterschiedlich aussehen, wenn es sich um verschiedene Patienten handelt, z.B. wenn Organe in der Größe krankhaft verändert sind.

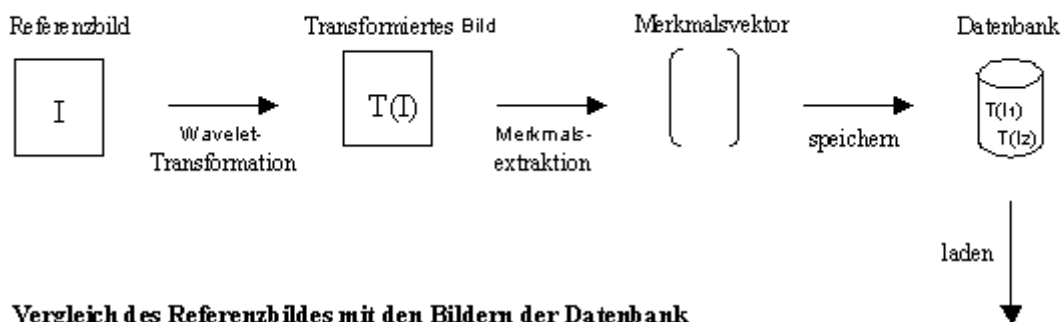
Genauso kann es passieren, dass völlig unterschiedlich codierte Bilder für das klassifizierende System sehr ähnlich sind, z.B. wenn bei Röntgenaufnahmen der größte Teil des Bildes schwarz bleibt.

Es ist also zu prüfen, ob das dargestellte Verfahren für die Anwendung auf medizinische Bilddaten und speziell für die hier am Institut geforderte Klassifizierung medizinischer Bilddaten geeignet ist. Desweiteren soll versucht werden, das in [2] vorgestellte Verfahren im Hinblick auf die Klassifizierung weiter zu verbessern. Hierzu werden verschiedene Normierungen der wavelet-transformierten Bilddaten untersucht. Außerdem wird der in [2] verwendete gewichtete euklidische Abstand durch andere Distanzfunktionen ersetzt.

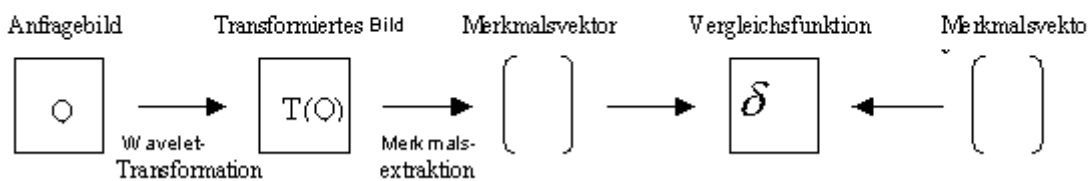
Abbildung 3-1 verdeutlicht das zu implementierende Verfahren zur Merkmalsextraktion und zum Bildvergleich.

Abb. 3-1: Bildung und Vergleich der Merkmalsvektoren

Bildung der Merkmalsvektoren der Referenzdaten



Vergleich des Referenzbildes mit den Bildern der Datenbank



4 Verfahren

Im nun vorgestellten Ansatz wird das bestehende Wavelet-Verfahren [2] folgendermaßen erweitert

Die Bildgröße, auf die ein Bild zu Anfang skaliert wird und die Anzahl der Iterationsschritte bei der Wavelet-Transformation wurde variabel gestaltet. Außerdem wurde das in [2] beschriebene Distanzmaß durch entsprechende Normierungen erweitert, und es wurden zwei weitere Distanzmaße definiert.



Da die hier zum Einsatz kommende Merkmalsextraktion auf dem Verfahren der Wavelet-Transformation beruht, beschäftigt sich das nächste Kapitel ausschließlich mit der Theorie der Wavelet-Transformation und versucht, dieses grundlegende Verfahren genauer zu erläutern. Im folgenden wird dann beschrieben, wie mit Hilfe der Wavelet-Transformation die gewünschten Daten aus einem Bild extrahiert werden können.

4.1 Die Wavelet-Transformation

Wie bei der Fourier-Transformation wird durch die Wavelet-Transformation ein Signal in seine Frequenzkomponenten zerlegt, indem es als gewichtete Summe spezieller Basisfunktionen dargestellt wird.

Bei der Fourier-Transformation sind das die Sinus- und Kosinusfunktionen. Da das gesamte Signal für die Berechnung der Fourier-Transformierten herangezogen wird, erhält man zwar eine exakte Frequenzinformation, jedoch keine Information über das ortsabhängige Auftreten der einzelnen Frequenzen. Um dieses Problem zu lösen, kann das Signal zuvor mit einer Fensterfunktion multipliziert werden. Die Fourier-Transformation wird dann immer nur auf einen durch das Fenster bestimmten Bereich des Signals angewandt. So erhält man für jedes Fenster ein eigenes Frequenzspektrum und damit auch Informationen über die Orte des Auftretens von einzelnen Frequenzen. Allerdings gestaltet es sich als schwierig, eine passende Größe für die Fensterfunktion zu bestimmen, so dass sowohl hohe als auch tiefe Frequenzen richtig erkannt werden können.

Bei der Wavelet-Transformation ist dagegen eine signalangepasste Auflösung möglich. Als Basisfunktionen werden die Elemente einer Wavelet-Funktionenfamilie verwendet, die aus einem Basiswavelet durch Skalierung und Verschiebung gebildet werden. Durch die Skalierung wird die Ausgangsfunktion gedehnt bzw. gestaucht, was einer Vergrößerung bzw. Verkleinerung der Fensterfunktion entspricht. Mit den gedehnten Funktionen können tiefe Frequenzen analysiert werden, die gestauchten Funktionen eignen sich für die hochfrequenten Bildanteile. Durch die Ortsverschiebung des Basiswavelets wird eine Abdeckung des gesamten Bildbereichs erreicht.

Jeder Skalierungsschritt entspricht einer bestimmten Auflösungsstufe. In jeder Auflösungsstufe muss das Wavelet angepasst an seine Breite so verschoben werden, dass es einerseits nicht zu Überschneidungen kommt und andererseits das gesamte Signal erfasst wird. Für die vollständige Darstellung eines digital aufgenommenen Signals, werden die Wavelets so lange gestaucht, bis die Auflösungstiefe des Signals erreicht ist.

Der Anteil jeder Basisfunktion an der darzustellenden Funktion wird durch den Wavelet-Koeffizienten angegeben, welcher eine Gewichtung für jedes skalierte und verschobene Wavelet bestimmt.

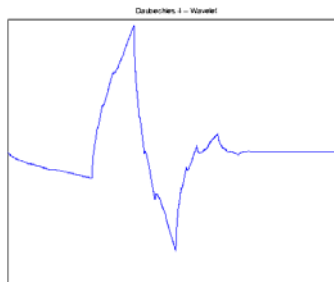
Die Vorschrift für die Wavelet-Transformation einer Funktion $f(x)$ lautet:

$$F^W(m,n) = \int_{-\infty}^{+\infty} f(x) 2^{-m/2} \psi(2^{-m}x - n) dx = \langle f(x), \psi_{m,n}(x) \rangle$$

Gl. 4-1

wobei m die Skalierung und n die Translation des Wavelets $\psi(x)$ beschreibt und $F^W(m,n)$ den entsprechenden Wavelet-Koeffizienten angibt. In der vorliegenden Arbeit wurde das Daubechies-Wavelet [4] verwendet. Dieses ist durch seine Form (Abb. 4-1) mit einem Kantendetektor wie z.B. dem Sobelfilter vergleichbar und somit gut geeignet, um Helligkeitsübergänge in einem Bild darzustellen.

Abb. 4-1: Daubechies Basis-Wavelet [4], die Kantenform ist deutlich zu erkennen



4.1.1.1.1 Anwendung für die Bildanalyse

Mit Hilfe der Multi-Resolutions-Analyse [5] [6] können aus einer Wavelet-Basis Filterkoeffizienten für einen Hochpass- und einen Tiefpassfilter berechnet werden. Der Tiefpassfilter ermöglicht hierbei den Übergang einer Auflösungsstufe zur nächstgrößeren, während der Hochpassfilter die Unterschiede zwischen beiden Auflösungen erfasst.

Die Filter stehen in einer orthogonalen Beziehung zueinander:

$$\int HP \circ TP = 0$$

Gl. 4-2

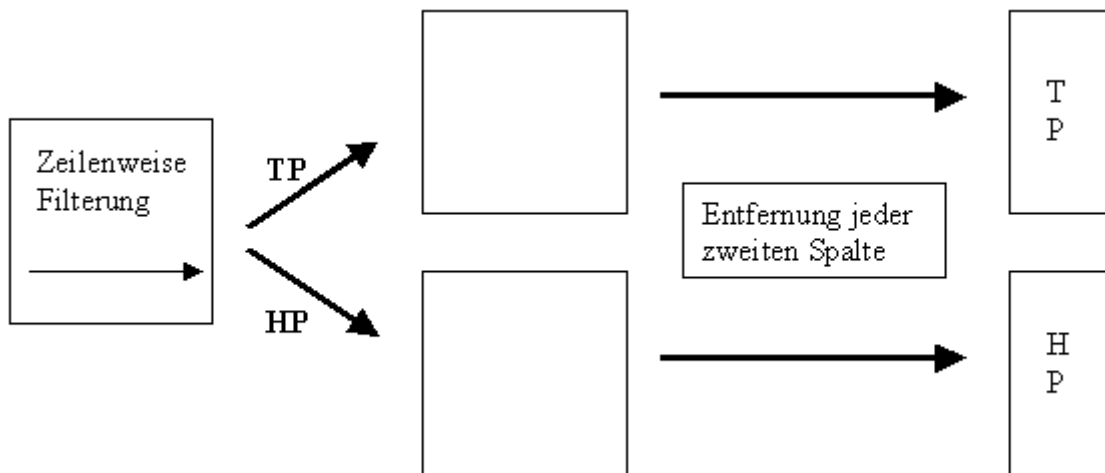
Dies bedeutet, dass keine Bildinformation von beiden Filtern gleichzeitig erfasst wird. Durch den Hochpassfilter werden genau die Bildinformationen erfasst, die durch die Tiefpassfilterung verloren gehen. Um nun eine Wavelet-Transformation durchzuführen, wird ein Bild zeilen- und spalten-



weise sowohl mit dem Tiefpass als auch mit dem Hochpass gefiltert.

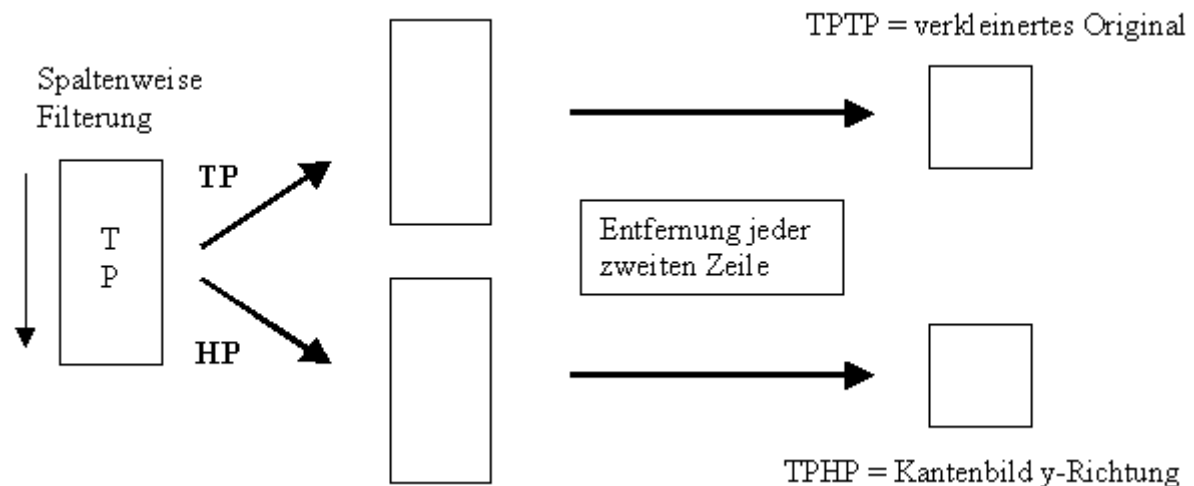
Als erstes wird eine zeilenweise Filterung durchgeführt. Durch den Tiefpass gehen hierbei die hochfrequenten Bildinformationen verloren, während der Hochpass genau diese herausfiltert. Im hochpassgefilterten Bild sind also die Detailinformationen enthalten, die im Tiefpassbild fehlen. Da die Bandbreite in horizontaler Richtung nun nur noch halb so groß ist, kann jede zweite Spalte verlustfrei entfernt werden (down-sampling), so dass Hochpass- und Tiefpassbild zusammengenommen in ihrer Größe wieder dem Originalbild entsprechen.

Abb. 4-2: zeilenweise Wavelet-Transformation

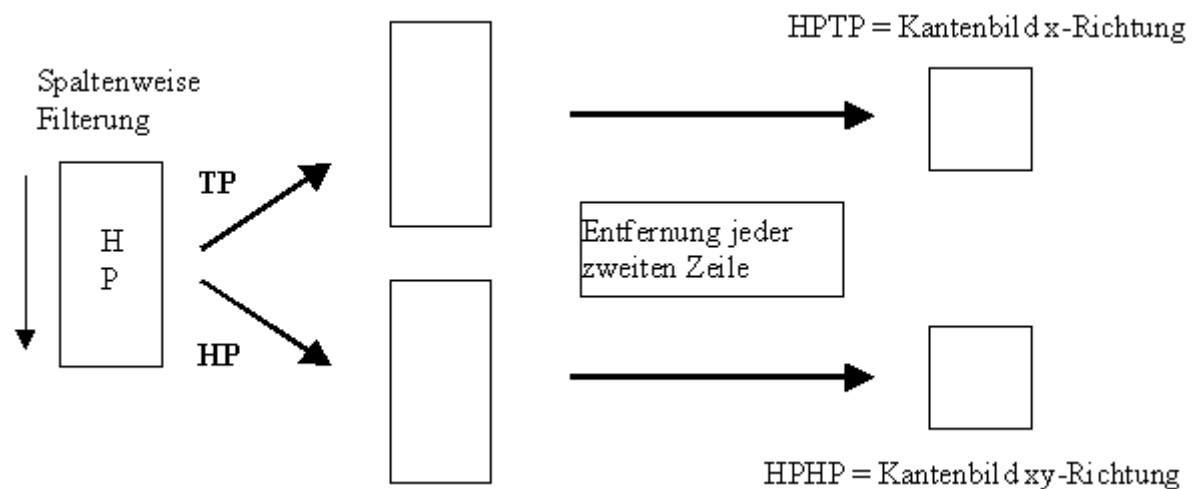


Im zweiten Schritt findet eine spaltenweise Filterung statt. Anschließend kann jede zweite Zeile verlustfrei entfernt werden. Durch die spaltenweise Filterung entstehen vier verschiedene Teilbilder.

Die spaltenweise Tiefpassfilterung des zeilenweise tiefpassgefilterten Bildes, ergibt ein dem verkleinerten Original entsprechenden Tiefpassbild. Die Hochpassfilterung desselben Bildes bringt dagegen die Kanten in y-Richtung zum Vorschein, da ja in diese Richtung gefiltert wird.

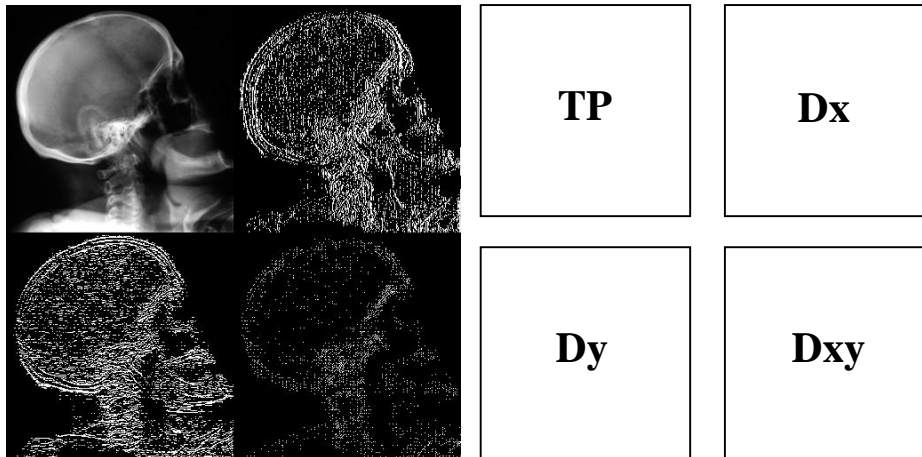
Abb. 4-3: spaltenweise Wavelet-Transformation des TP-Bildes

Das zeilenweise gefilterte Hochpassbild enthält bereits die Kanten in x-Richtung. Bei der spaltenweisen Hochpassfilterung werden außerdem die Kanten in y-Richtung sichtbar. Durch die spaltenweise Tiefpassfilterung erhält man dagegen ein verkleinertes Bild, das nur die Kanten in x-Richtung enthält

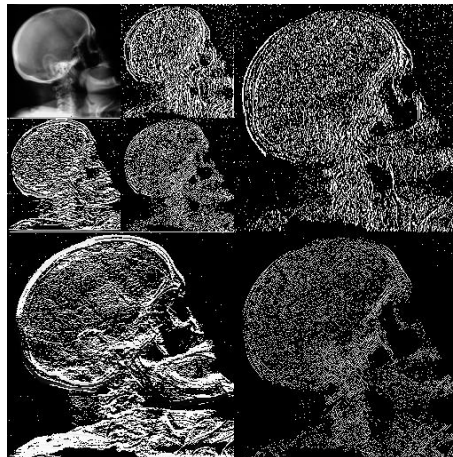
Abb. 4-4: spaltenweise Wavelet-Transformation des HP-Bildes

Der gesamte Filterungsschritt zerlegt das Bild also in vier Teilbilder. Links oben (TP) erhält man ein zweimal tiefpassgefiltertes Bild, das dem verkleinerten Original entspricht. Außerdem erhält man drei Differenzbilder (Dx, Dy, Dxy), welche als Kantendetektoren in horizontaler (Dx), vertikaler (Dy) und diagonaler (Dxy) Richtung wirken.

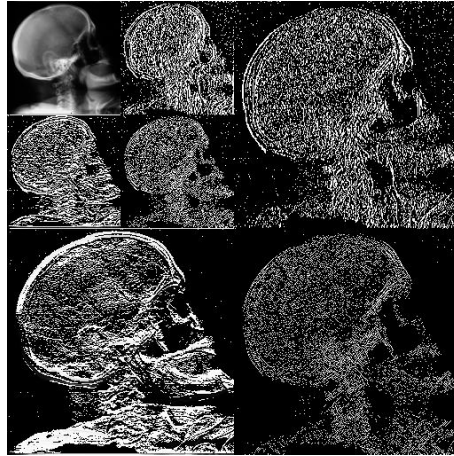


Abb. 4- 5: Wavelet-Transformation, erster Iterationsschritt

Auf das verkleinerte Originalbild kann nun ein weiterer Transformationsschritt angewandt werden.

Abb. 4-6: Wavelet-Transformation, zweiter Iterationsschritt

Mit weiteren Iterationsschritten wird das herunterskalierte Originalbild immer kleiner. Seine Bandbreite wird immer geringer, da die für die jeweilige Bandbreite hohen Frequenzen gefiltert und in den Differenzbildern gespeichert werden.

Abb. 4-7: Wavelet-Transformation, dritter Iterationsschritt

4.2 Vorverarbeitung

Bilder in einer Datenbank liegen in verschiedenen Größen vor. Bevor aus einem Bild also bestimmte Merkmale für einen Vergleich extrahiert werden können, muss zunächst eine Größennormalisierung des Bildes durchgeführt werden. Alle zu vergleichenden Bilder werden hierzu auf eine vorher definierte Größe skaliert. Hierbei muss darauf geachtet werden, dass die Seitenlängen der Bilder nur Zweierpotenzen sein dürfen. Weiterhin ist darauf zu achten, dass die Bilder auf eine quadratische Größe skaliert werden. Bei der Skalierung nicht quadratischer Bilder entsteht hierdurch zwar eine Verzerrung der Dimensionen, für den Bildvergleich ist dies jedoch weitgehend unerheblich, da alle Bilder auf dieselbe Weise verändert werden.

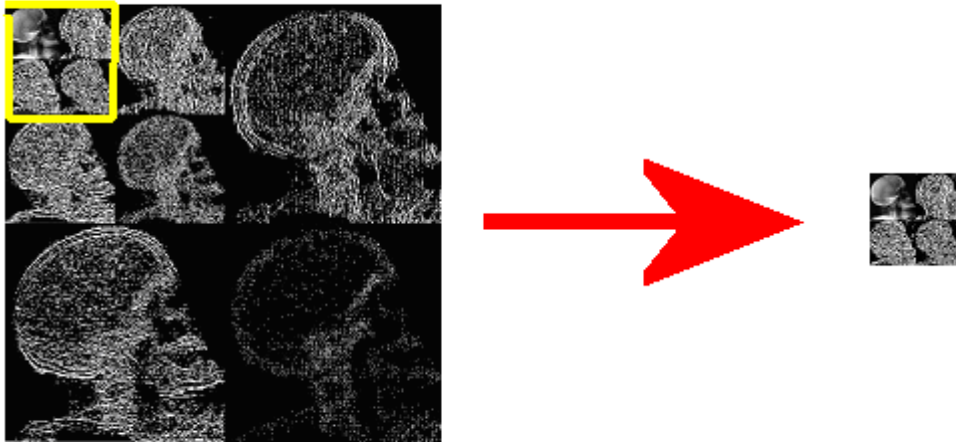
Der Algorithmus für die Wavelet-Transformation wurde aus [7] übernommen. Für die Umskalierung der Bilder auf eine frei wählbare Bilddimension $2^n \times 2^n$ wurde die lineare Interpolation verwendet.

4.3 Bildung des Merkmalsvektors

Zur Bildung des Merkmals-Vektors wird die Wavelet-Transformation je nach Testreihe mit fünf, sechs oder sieben Iterationsschritten durchgeführt. Alle herausgefilterten Detailinformationen des Bildes werden verworfen. Nur das tiefpassgefilterte, verkleinerte Original und die direkt zugehörigen Kantenbilder der letzten Iterationsstufe bleiben erhalten (Abb. 4-8).



Abb. 4-8: extrahierte Daten (64x64 Pixel), aus dem 512x512 Pixel großen Originalbild, zur Illustration wurden drei Iterationsschritte durchgeführt, da das verkleinerte Tiefpassbild sonst nur schwer zu erkennen ist.

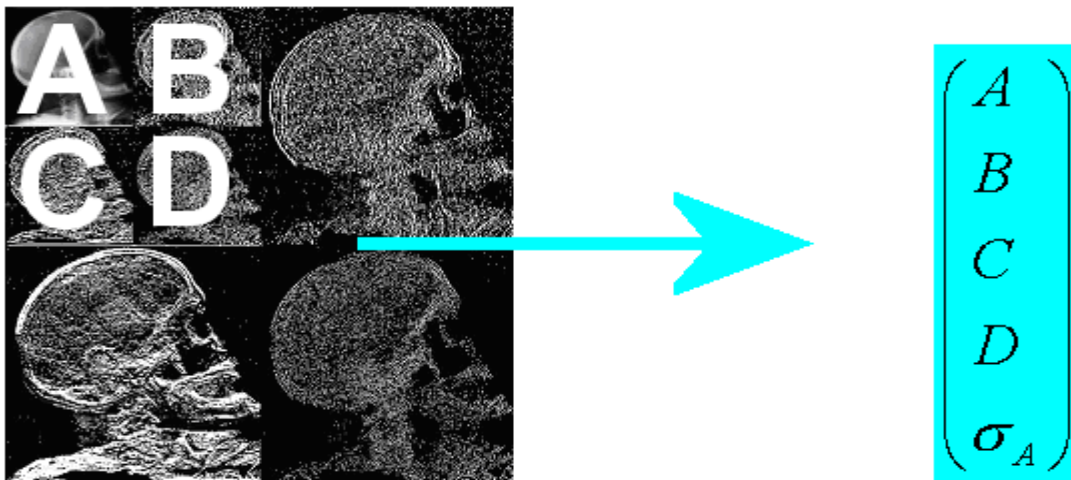


Da die Seitenlänge des Tiefpassbildes mit jedem Iterationsschritt halbiert wird, kann die Anzahl der verbleibenden Bildpunkte leicht berechnet werden. Beginnt man z.B. mit einem auf 512x512 Bildpunkte skalierten Bild, so bleiben nach sechs Iterationsschritten für das Tiefpassbild noch 8x8 also 64 Pixel übrig. Hinzu kommen die drei Kantenbilder, so dass das Bild insgesamt von 262144 Pixel auf 256 Bildinformationen reduziert wird.

Je häufiger die Wavelet-Transformation ausgeführt wird, desto mehr Details gehen verloren. Andererseits bedeutet ein stark detailreduziertes also kleines Bild, dass Bildvergleiche schneller ausgeführt werden können.

Die für die Bildsuche relevanten Daten werden in einem Merkmalsvektor gespeichert, welcher von nun an das Bild repräsentiert. Diese Daten bestehen aus allen nach der Datenreduktion verbleibenden Bildpunkten, wobei die vier Einzelmatrizen, A, B, C, D (Abb. 4-9) getrennt abgespeichert werden. Zu den Bildpunkten der Matrix A, die das verkleinerte Originalbild enthält, wird außerdem die Standardabweichung berechnet. Diese wird später beim Bildvergleich für eine Vorauswahl benötigt.

Abb. 4-9: Bildung des Merkmalsvektors aus der Wavelet-Pyramide bestehend aus Tiefpass (A), x-Ableitung (B), y-Ableitung (C), xy-Ableitung (D) und der Standardabweichung von A



Aus dem Merkmals-Vektor lässt sich das Bild leicht rekonstruieren. Hierbei erkennt man, welche Bilddaten bei der Datenreduktion verworfen wurden bzw. auf welchen

Abb. 4-10: Aus dem Merkmalsvektor (4 Iterationsschritte, 64x64 Pixel gespeichert) rekonstruierte Bilddaten auf Originalgröße von 512x512 Pixeln skaliert



Bilddaten der eigentliche Vergleich durchgeführt wird.

4.3.1.1.1 Detailreduktion

Abb. 4-10 zeigt ein rekonstruiertes 512x512 Bild nach einer vierfachen Wavelet-Transformation. Nach dieser Transformation bleibt ein Tiefpassbild von 32x32 Pixeln, sowie drei Kantenbilder derselben Größe. Aus diesen vier Teilbildern kann mit Hilfe einer Rücktransformation ver-



lustfrei das 64x64 Pixel große Tiefpassbild dargestellt werden. Da für die Speicherung im Merkmalsvektor alle anderen Kantenbilder verworfen wurden, können die noch fehlenden Informationen nicht wiederhergestellt werden. Alle Strukturen, die kleiner als 4x4 Pixel waren sind weggefallen. Das dargestellte Bild zeigt somit alle im Merkmalsvektor gespeicherten Informationen.

4.4 Klassifikation

4.4.1 Entscheidungsfindung

Nachdem die für die Bildsuche relevanten Daten in einem Merkmalsvektor gespeichert wurden, soll anhand dieses Merkmalsvektors ein Vergleich zweier Bilder zur Klassifikation stattfinden. Die im Vektor gespeicherten d Bildpunkte kann man sich hierbei als d Merkmale in einem d -dimensionalen Merkmalsraum vorstellen. Jede Dimension entspricht dabei einer Koordinate. Unterschiedliche Bildinhalte oder allgemein Muster sollen zu weit auseinanderliegenden Vektoren im Merkmalsraum führen. Umgekehrt sollen ähnliche Muster auch im Merkmalsraum benachbart sein. Ist diese Clusterung gegeben, so können Grenzen bestimmt werden, die die Muster in bestimmte Klassen einteilen. Dazu stehen in der Literatur zahlreiche Verfahren zur Verfügung.

Zur Zuordnung eines Vektors in einen bestimmten Cluster werden entsprechende Klassifikatoren in Form von Distanzen oder Vektornormen benötigt. Anhand von Trainingsdaten werden die Grenzen im Merkmalsraum bestimmt. Die Klassifikatoren können dann genutzt werden, um neue Muster in die entsprechende Klasse einzuordnen.

4.4.2 Klassifikatoren

In dieser Arbeit wurde die nearest-neighbor-Klassifikation mit verschiedenen Abstandsfunktionen untersucht. Diese bestimmen für ein jeweiliges Anfragebild die Abstände zu allen in der Referenzmenge vorhandenen Vektoren. Die vorklassifizierten Bilder der Datenbank bilden somit die Referenzdaten, wobei im Zuge des Leaving-One-Out-Tests jedes Bild der Datenbank auch einmal als Anfragebild auftritt. Die Klassifikation findet anhand der Referenzdaten statt, zu denen das Anfragebild den geringsten Abstand aufweist.

Um eine Klassifikation durchzuführen, muss zum Anfragebild ein Merkmalsvektor berechnet werden, der in seiner Dimensionierung denen der Referenzbilder entspricht. Dazu muss das Bild auf die Standardgröße der Datenbankbilder skaliert wer-

den, und die Wavelet-Transformation muss mit der gleichen Anzahl Iterationsschritte durchgeführt werden. Da in dieser Arbeit jedes Anfragebild auch als Referenzbild genutzt wird, ist der Vektor eigentlich bereits vorhanden. Die jeweilige Neuberechnung des Vektors des Anfragebildes ist allerdings nicht sehr zeitaufwändig und ermöglicht es, das Verfahren auch für beliebige Anfragebilder außerhalb der Datenbank zu nutzen.

4.4.2.1.1 Distanzmaß

Zu jedem Vektor der Datenbank wird die Distanz zum Vektor des Anfragebildes berechnet.

Hierbei wird zunächst eine Vorauswahl anhand der Standardabweichung des Tiefpassbildes durchgeführt. Die Standardabweichungen der zwei zu vergleichenden Bilder sollen möglichst dicht beieinander liegen, wobei das zulässige Abweichungsintervall durch einen Parameter festgelegt werden kann. Das Akzeptanz-Kriterium ist gegeben durch:

$$\sigma\beta < \sigma' < \frac{\sigma}{\beta}$$

Gl. 4-3

Hierbei gilt $\beta = 1 - \frac{\text{percent}}{100}$. Percent ist die einzustellende Variable, mit der festgelegt werden kann, wie viele Bilder die Vorauswahl bestehen. σ bezeichnet die Standardabweichung des Anfragebildes, während σ' die Standardabweichung des Referenzbildes darstellt.

Besteht ein Bild diese Vorauswahl nicht, so wird der Abstand zum Anfragebild auf einen Maximalwert gesetzt und es wird zum nächsten Bild übergegangen.

Ansonsten wird nun eine gewichtete Abstandsberechnung auf dem in 4.3 beschriebenen Merkmalsvektor der beiden Bilder durchgeführt. Neben dem in [2] beschriebenen euklidischen Abstand (vgl. Kap. 4.4.2.1) wurde auch der Mahalanobis-Abstand (vgl. Kap. 4.4.2.2) sowie eine Korrelationsfunktion (vgl. Kap. 4.4.2.3) implementiert. Diese können nun als Vergleichsfunktionen (vgl. Abb. 1-2) verwendet werden.

1.1.1.1. Euklidischer Abstand

Der in [2] beschriebene euklidische Abstand wird nach der Formel

$$\text{dist}(\text{image}, \text{image}') = \alpha \|A - A'\| + \beta \|B - B'\| + \gamma \|C - C'\| + \delta \|D - D'\|$$

Gl. 4-4



berechnet. Die A bis D entsprechen dabei den vier Teilmatrizen (Abb. 4-9), also dem Tiefpass-Bild und den drei zugehörigen Kantenbildern. Die α bis δ stellen die einzelnen Gewichte dar, die für die vier Teilmatrizen zusätzlich eingestellt werden können.

Anstelle des tatsächlichen euklidischen Abstands wird sein Quadrat berechnet, um die Rechenzeit der Wurzeloperation zu sparen.

Erhalten alle Gewichte den Wert eins, so entspricht die Formel tatsächlich dem euklidischen Abstand bzw. dessen Quadrat. Möchte man allerdings besonderen Wert auf eine bestimmte Kantenrichtung legen, so kann man das Gewicht für die entsprechende Matrix erhöhen.

Um die Rechengeschwindigkeit beim Bildvergleich über die gesamte Datenbank zu erhöhen, wird außerdem ein weiterer Schwellenwert gesetzt, der festlegt, wie groß die Distanz zwischen den einzelnen Komponenten der beiden Merkmalsvektoren maximal sein darf. Wird dieser Wert überschritten, so kann die Berechnung sofort abgebrochen werden. Die Distanz der beiden Bilder wird dann, wie bei der Betrachtung der Standardabweichung, auf den Maximalwert gesetzt.

4.4.2.1.2 Normierung der Grauwerte pro Komponente

Ein Problem bei der Bestimmung des euklidischen Abstands zwischen zwei Bildern besteht darin, dass die einzelnen Bildpunkte (= Komponenten des Merkmalsvektors) nicht unbedingt einen gleichwertigen Einfluss auf das Endergebnis haben. Wenn zwei Komponenten in unterschiedlichen Wertebereichen liegen, so ist auch der Einfluss auf das Gesamtergebnis verschieden groß. Ein geläufiger Vorverarbeitungsschritt ist die Normierung. Sie wird nach der Wavelet-Transformation auf die Wavelet-Koeffizienten des Bildes angewandt. Aus den normierten Koeffizienten wird dann der Merkmalsvektor gewonnen, welcher für den Bildvergleich genutzt wird. Die Parameter für die Normierung werden aus der Referenzmenge berechnet, dieselbe Normierung kann dann für das Anfragebild genutzt werden.

Um allen Bildpunkten einen gleichwertigen Einfluss auf das Endergebnis zu verleihen, kann eine Normierung über alle Komponenten durchgeführt werden. Die Parameter der Normierung werden hierbei nicht für jedes Bild neu berechnet, sondern für jeden Bildpunkt, dafür aber über alle Bilder der Referenzmenge.

Für jeden Bildpunkt k wird dabei der in der Referenzmenge angenommene minimale $\min[k, R]$ und maximal $\max[k, R]$ Grauwert gesucht. Aus deren Differenz

$scale[k] = \max[k, R] - \min[k, R]$ wird für jeden Bildpunkt ein Skalierungsfaktor berechnet. Dann werden die Grauwerte aller Bilder mit Hilfe der Formel

$$(bildpunkt[k] - \min[k]) \cdot \frac{1}{scale[k]}$$

Gl. 4-5

skaliert. Durch die Skalierung wird der minimale Wert eines jeden Bildpunktes auf Null abgebildet, der maximale Wert auf Eins. Hierdurch kommt es zu einem gleichwertigen Einfluss aller Bildpunkte auf das Endergebnis.

Problematisch bei dieser Art der Normierung ist die Behandlung von Ausreißern, da das Intervall für die Normierung allein durch den minimalen und maximalen Wert eines Bildpunktes bestimmt wird. Sind diese Werte Ausreißer, liegen sie also weit von den anderen Werten entfernt, so wird das Skalierungsintervall zu groß gewählt und die einzelnen Bildwerte liegen bis auf den Ausreißer zu dicht beieinander. Während der Auswertung konnte allerdings festgestellt werden, dass dieses Problem beim zur Verfügung stehenden Bildmaterial kaum auftrat und somit eine untergeordnete Rolle spielte.

4.4.2.1.3 Helligkeitsnormierung pro Teilmatrix

Bei dieser Art der Bildnormierung wird das Vorwissen genutzt, dass jedes Bild aus vier Teilmatrizen besteht. Eine für jede Teilmatrix getrennte Normierung erscheint sinnvoll, da die Wavelet-Koeffizienten des Tiefpassbildes in einem anderen Wertebereich liegen als die der Hochpassbilder, welche nur Kanten enthalten. Die Normierungen werden hierbei nicht für jede Komponente über die gesamte Referenzmenge ausgeführt, sondern für jedes Bild der Referenzmenge getrennt. Somit wird auch bei der Normierung des Anfragebildes kein Vorwissen über die Referenzmenge benötigt. Die Helligkeitsnormierung entspricht einer Vektornormierung $x/||x||$. Jeder Intensitätswert wird hierbei durch die Wurzel der Summe aller quadrierten Intensitätswerte in seiner jeweiligen Teilmatrix dividiert. Nach der Normierung ist die Summe aller Intensitätswerte einer Teilmatrix gleich Eins.

$$\frac{\bar{x}}{||\bar{x}||} = \frac{x_i}{\sqrt{\sum_{i=1}^t x_i^2}}, i = 1, \dots, t$$

Gl. 4-6



Durch diese Normierung wird ein gleichwertiger Einfluss aller vier Teilmatrizen auf das Endergebnis gewährleistet. Durch die verschiedenartige Beschaffenheit der Teilbilder besteht in unnormiertem Zustand nämlich die Möglichkeit, dass ein Teilbild in der Abstandsberechnung bevorzugt beachtet wird, da in diesem generell größere Abstände zwischen den einzelnen Bildpunkten vorliegen.

4.4.2.1.4 Histogrammspreizung pro Teilmatrix

Auch diese Normierung wird für jedes Bild getrennt ausgeführt, wobei auch hier wieder die einzelnen Teilmatrizen getrennt betrachtet werden. Zuerst wird für jede Teilmatrix das angenommene Minimum und Maximum der Bildwerte bestimmt. Die Berechnung entspricht dann der Normierung, die bei der Skalierung über alle Komponenten verwendet wurde, nur dass sie diesmal für die einzelnen Teilmatrizen jedes Bildes ausgeführt wird.

Alle Bildpunkte werden also in einen Bereich zwischen Null und Eins skaliert, wobei der maximale Bildwert jeder Teilmatrix den Wert Eins erhält und der minimale den Wert null. Dieses Verfahren kombiniert also die Idee der Gleichberechtigung aller Punkte am Endergebnis mit der der Gleichberechtigung der vier Teilmatrizen.

1.1.1.2. Mahalanobis Abstand

Im Gegensatz zur euklidischen Distanz wird bei der Mahalanobis Distanz die statistische Verteilung der Punkte im Vektorraum in die Distanzberechnung integriert. Damit werden Ausreißer nicht mehr so stark bewertet wie bei der Normierung der Grauwerte pro Komponente (vgl. Kap. 4.4.2.1).

Der Abstand zweier in dieser Verteilung enthaltenen Punkte x und y berechnet sich dann mit Hilfe der Kovarianzmatrix S , welche die Kovarianz der einzelnen Bildpunkte über die gesamte Referenzmenge enthält.

$$d(x, y) = (x - y)^T S^{-1} (x - y)$$

Gl. 4-7

Graphisch bilden die Punkte gleicher Mahalanobis Distanz von einem Zentrum im Zweidimensionalen eine gedrehte Ellipse, während es bei der euklidischen Distanz ein Kreis ist.

Bei der Distanzberechnung zwischen zwei Bildern müssen die Einträge der Kovarianzmatrix aus den Referenzbildern geschätzt werden. Hierbei wird die vereinfachende Annahme gemacht, dass die Komponenten im Merkmalsvektor, also die einzel-

nen Bildpunkte, unabhängig voneinander sind. Hierdurch enthält nur die Hauptdiagonale der Kovarianzmatrix Werte ungleich Null. Diese stellen die Varianz des jeweiligen Bildpunktes in der Referenzmenge dar. Der Abstand zweier Bilder berechnet sich dann einfach aus:

$$d(x, y) = \sqrt{\sum \frac{(x_k - y_k)^2}{\sigma_k^2}}$$

Gl. 4-8

wobei k die Laufvariable über die einzelnen Bildpunkte ist. Das Bild x besteht also aus den Bildpunkten x_k , das Bild y aus allen y_k . σ_k^2 ist die Varianz des jeweiligen Bildpunktes bezogen auf die gesamte Referenzmenge

1.1.1.3. Korrelation

Zwischen zwei Bildern, die eigentlich das selbe Objekt zeigen, kommt es mit herkömmlichen Distanzmaßen zu großen Abständen, wenn die Objekte gegeneinander verschoben sind, da dann Distanzen zwischen den „falschen“ Koordinatenpaaren gebildet werden. Um dieses Problem zu lösen, werden in der hier genutzten Korrelationsfunktion die Bilder explizit gegeneinander verschoben, bis die größte Übereinstimmung gefunden wird. Hierzu kann sowohl in x - als auch in y -Richtung ein Wert angegeben werden, der die maximale Verschiebung angibt. Dann werden alle Kombinationen aus möglichen x - und y -Verschiebungen der Reihe nach getestet. Die maximale Verschiebung wird dabei sowohl in positiver als auch in negativer Richtung betrachtet. Der Korrelationskoeffizient, der die Güte der Übereinstimmung beschreibt, wird aus allen Punkten bestimmt, die im jeweiligen Verschiebungsschritt übereinanderliegen. Er berechnet sich aus der Kovarianz und der Varianz nach der Formel

$$\frac{COV(X, Y)}{\sqrt{V(X)}\sqrt{V(Y)}} = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2} \sqrt{\sum_i (y_i - \bar{y})^2}}$$

Gl. 4-9

wobei die x_i alle Bildpunkte des Bildes X darstellen, während Y aus allen y_i besteht.

\bar{x} und \bar{y} stellen den jeweiligen Mittelwert $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ aller Bildpunkte des Bildes X

bzw. Y dar.



Die Korrelationsberechnung wird für alle vier Teilmatrizen getrennt ausgeführt, wobei jedoch auf eine synchrone Verschiebung geachtet wird. Durch die Trennung der vier Teilmatrizen wird verhindert, dass die nicht überlappenden Randpunkte einer Teilmatrix mit der daneben liegenden Teilmatrix verglichen werden, anstatt aus der Berechnung ausgeschlossen zu werden. Die Gesamtkorrelation ist dann bestimmt durch die maximale Korrelation der Summe aller vier Teilkorrelationen über alle Verschiebungsschritte. Hierbei können wie bei der euklidischen Abstandsberechnung entsprechende Gewichtungen für die einzelnen Teilmatrizen angegeben werden.

5 Auswertung

In diesem Kapitel wird zunächst das verwendete Evaluationssystem vorgestellt (5.1), dann wird ein Überblick über alle durchgeführten Auswertungen gegeben (5.2). Die Ergebnisse der Auswertungen werden erläutert (5.3), und schließlich bewertet (5.4).

5.1 Evaluationssystem

Die Bewertung des Verfahrens und der verschiedenen Klassifikatoren wurde mit Hilfe des IRMA-Codes (Kap. 2.3) durchgeführt. Hierbei wurde die in [8] beschriebene Methode verwendet, welche untersucht, inwieweit ein Klassifikationsverfahren in der Lage ist, medizinische Bilder selbstständig in eine passende Kategorie einzuordnen. Die Auswertungen wurden mit Hilfe der in der IRMA-Datenbank vorhandenen Radiographien durchgeführt, welche 92% des gesamten IRMA-Datenbestandes ausmachen. Somit standen 5839 Bilder zur Verfügung. Zur Einteilung in die verschiedenen Kategorien wurde der IRMA-Code mit der folgenden Codemaske genutzt: $11??-x??-xx?-x??$. Die 11 zu Beginn identifiziert die Radiographien, die mit x gekennzeichneten Stellen markieren die betrachteten Kategorien und die $?$ werden bei der Betrachtung ignoriert. Eine Kategorie wurde nur dann für gültig erklärt, wenn sie mindestens 10 Bilder enthielt, was vor Durchführung der Tests anhand der Codemaske überprüft wurde. Mit dieser Vorauswahl ergeben sich 5756 gültige Bilder aus 54 Kategorien.

5.1.1.1 Testmethode

Die Tests wurden mit der Leaving-One-Out Methode durchgeführt. Hierbei wird jedes Bild der Datenbank nacheinander als Anfragebild verwendet, die restlichen Bilder dienen als Referenz. Für jedes Anfragebild werden mit Hilfe des zu testenden Klassifikators die fünf nächsten Nachbarn bestimmt. Das Bild wird dann in die Kategorie eingeordnet, der die meisten dieser fünf Nachbarn angehören. Die Kategorien der Nachbarn sind am IRMA-Code zu erkennen. Da auch das Anfragebild bereits codiert ist, kann die Korrektheit der Zuordnung sofort überprüft werden. Nachdem mit der

Leaving-One-Out Methode alle Bilder der Datenbank kategorisiert wurden, kann der Prozentsatz korrekt eingeordneter Bilder angegeben werden.

Hierdurch lässt sich der Klassifikator objektiv bewerten, die Auswirkungen der verschiedenen Distanzmaße und der einzustellenden Parameter werden sofort sichtbar.

5.1.1.1.2 Technische Umsetzung

Technisch wird die Evaluierung mit Hilfe verschiedener Matrizen durchgeführt. Jedes Bild ist hierbei mit einer eindeutigen Identifikationsnummer abgespeichert. Zu jedem Bild gehört der bei der Merkmalsextraktion mit den Daubechies Wavelets gewonnene Vektor (vgl. Kap. 5.2), welcher einmal berechnet und dann in die Datenbank eingebunden wurde. Zusätzlich wurden die Klassifikatoren implementiert und mit einer Datenbankschnittstelle versehen.

Mit den Klassifikatoren wird als erstes eine Distanzmatrix berechnet. Jede Spalte dieser Matrix zeigt das Ergebnis eines Leaving-One-Out Durchlaufs für das angegebene Bild, welches hierbei durch seine ID gekennzeichnet ist. Somit enthält die Matrix alle Abstände der einzelnen Bilder zueinander, wobei jedes Bild einmal als Anfragebild verwendet wird.

Bei der Korrelation wird stattdessen eine Korrelationsmatrix erstellt, welche dann genau entgegengesetzt bewertet werden muss, da ja nicht die Bilder mit dem geringsten Abstand, sondern die mit der größten Korrelation gesucht werden.

Aus dieser Matrix können nun für jedes Anfragebild die Ergebnisse mit dem kleinsten Abstand gefunden werden. Das Anfragebild, wird damit in die Kategorie eingeordnet, der die meisten seiner fünf nächsten Nachbarn angehören. Gehören alle Nachbarn unterschiedlichen Kategorien an, so wird die Kategorie des nächsten Nachbarn gewählt. Sind alle fünf nächsten Nachbarn in ungültigen Kategorien, so werden die nächsten fünf Nachbarn bestimmt. Gehören auch diese nur ungültigen Kategorien an, so wird das Anfragebild übersprungen und nicht in die Bewertung aufgenommen.

Während der Kategorisierung wird eine sogenannte Confusion-Matrix gebildet. Die Zeilen der Matrix stellen alle bei den Anfragebildern vorkommenden Kategorien dar, wie sie aus dem IRMA-Code abgelesen werden können. Die Spalten enthalten dieselben Kategorien in derselben Reihenfolge. In den einzelnen Matrixeinträgen werden die tatsächlichen Zuordnungen gezählt. Die korrekte Zuordnung eines Anfragebildes führt somit zu einem Eintrag auf der Hauptdiagonalen. Aus der Confusion-Matrix kann der Prozentsatz aller korrekt zugeordneten Bilder bestimmt werden. Außerdem lässt sich leicht feststellen, wo die Probleme des angewendeten Zuordnungsmechanismus liegen, da

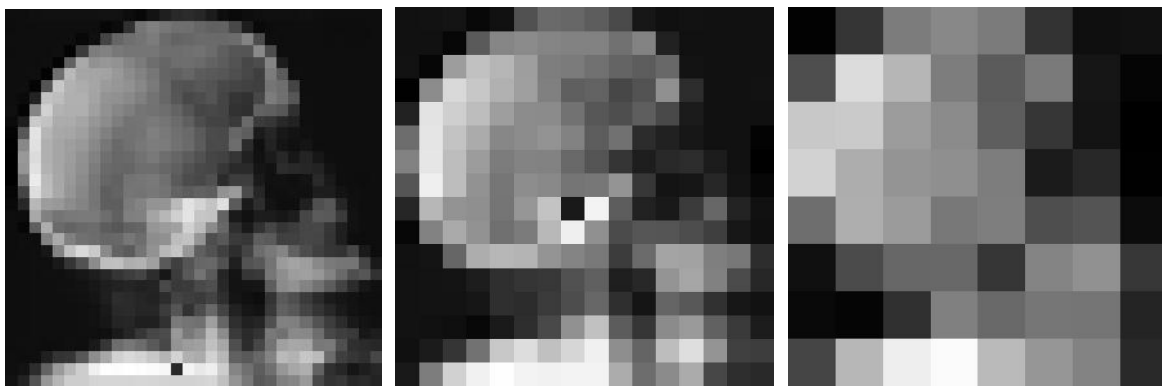


leicht zu ersehen ist, welche Kategorien häufig verwechselt werden

5.2 Durchführung

Für die Auswertungen wurden alle verwendeten Radiographien auf 512x512 Pixel skaliert. Die Tests wurden mit fünf, sechs und sieben Iterationen der Wavelet-Transformation durchgeführt. Somit standen für den Bildvergleich anhand des Merkmalsvektors 32x32, 16x16 und 8x8 Pixel zur Verfügung. Die dargestellten Bilder (Abb.5-1) zeigen in vergrößerter Darstellung die verwendeten Pixel. Wie man sieht, wird nur sehr wenig Bildinformation für den Vergleich benötigt.

Abb. 5-1: Daten für den Bildvergleich, aus Merkmalsvektor rekonstruiert. Die 512x512 Ausgangspixel wurden auf 32x32, 16x16 und 8x8 Pixel reduziert



Zuerst wurde der in [2] beschriebene gewichtete euklidische Abstand getestet. Die vier Teilmatrizen wurden jeweils mit 0, 0.5 oder 1 gewichtet, wobei alle möglichen Kombinationen untersucht wurden.

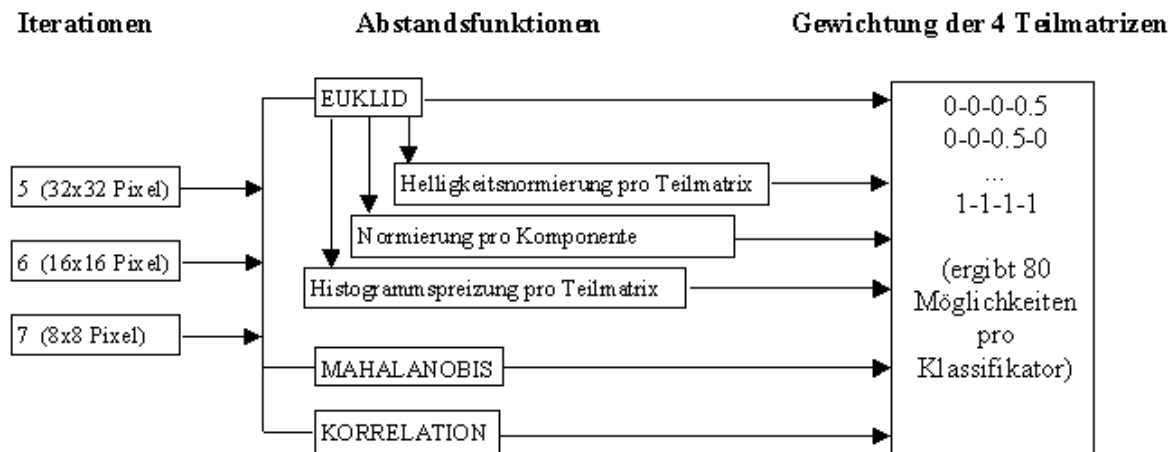
Dieselben Berechnungen wurden danach nochmals ausgeführt, wobei zuvor jeweils eine der drei Normierungen verwendet wurde.

Danach wurden die Tests mit der Mahalanobis-Distanz wiederholt.

Zum Schluss wurden die Berechnungen mit der Korrelationsfunktion durchgeführt.

Eine Übersicht aller durchgeführten Tests ist in Abb. 5-2 gegeben.

Abb. 5-2: Durchgeführte Tests, insgesamt 1440 (Merkmalsvektor in 3 verschiedenen Größen x 6 Klassifikatoren mit je 80 Gewichtungen)



Der Schwellenwert für die Vorauswahl (vgl. Kap. 4.4.2) wurde in allen Fällen auf 50% festgelegt.

Zum Vergleich wurde außerdem der einfache euklidische Abstand für die auf die entsprechende Größe skalierten Bilder berechnet. Hierbei wurde also nicht das Verfahren der Wavelet-Transformation genutzt und es konnten auch keine Gewichte eingestellt werden, da der Vergleich einfach auf dem verkleinerten Originalbild durchgeführt wurde.

Auch eine Korrelationsfunktion wurde ohne vorherige Wavelet-Transformation auf die verkleinerten Originalbilder angewandt

5.3 Ergebnisse

5.3.1 Visuelle Auswertung

Zunächst kann man sich einen optischen Eindruck verschaffen, ob die vom jeweiligen Verfahren ausgewählten nächsten Nachbarn dem Anfragebild tatsächlich ähnlich sehen. Hierbei wird das Anfragebild oben links dargestellt. Danach folgt eine Auflistung der nächsten Nachbarn. Unter jedem Bild ist hierbei der berechnete Abstand zum Anfragebild angegeben.

Abb. 5-3 und 5-4 zeigen, dass diese Methode gut geeignet ist, um sich einen ersten Eindruck über die Effizienz des Verfahrens zu verschaffen. Allerdings fällt eine weitere Bewertung des Verfahrens mit dieser Methode schwer, da kaum objektive Merkmale vorhanden sind, die einen Vergleich mit anderen Verfahren ermöglichen.



sample 491

491

(sample)

11_-1_-41_-7_



524

(1/20)

5.193582e-02

11_-1_-41_-7_



484

(2/20)

5.321559e-02

11_-1_-41_-7_



540

(3/20)

6.339452e-02

11_-1_-41_-7_



493

(4/20)

6.546012e-02

11_-1_-41_-7_



531

(5/20)

6.553286e-02

11_-1_-41_-7_

Abb. 5-3: Die 5 nächsten Nachbarn des Anfragebildes 491. Die Abstandsberechnung wurde mit dem gewichteten Euklid bei vorheriger Helligkeitsnormierung durchgeführt. Es wurden 5 Wavelet-Passes berechnet. Neben dem Abstand zum Anfragebild ist für jedes Bild sein IRMA-Code angegeben. Die Abbildung zeigt eine erfolgreiche Klassifikation, bei der alle 5 Nachbarn dieselbe Codierung aufweisen, welche dann auch dem Anfragebild zugewiesen wird.

sample 722



722

(sample)

11_-1_-80_-7_



724

(1/20)

7.151663e-02

11_-2_-94_-7_



10

(2/20)

8.604936e-02

11_-1_-70_-4_



621

(3/20)

8.741453e-02

11_-1_-80_-7_



492

(4/20)

8.864062e-02

11_-1_-41_-7_



619

(5/20)

8.889676e-02

11_-1_-80_-7_

5.3.2 Systematik der Auswertung

Für eine objektivere Bewertung wurde also das in 5.1 beschriebene Verfahren genutzt. Der folgende Ausdruck (Abb. 5-5) zeigt anhand des gewichteten, euklidischen Abstands mit Helligkeitsnormierung die Tests, die für jeden Klassifikator durchgeführt wurden. *Spie2004_512_5* bezeichnet den verwendeten Datensatz der Radiographien, welche hier auf 512x512 Pixel skaliert wurden und fünf Wavelet-Passes durchliefen. Die nächsten vier Zahlen zeigen die Gewichtungen der vier Teilmatrizen in der Reihenfolge links oben, rechts oben, links unten und rechts unten (Abb. 4-9, Teilmatrizen A, B, C, D). Dann folgt der Parameter *-norm*, welcher angibt, dass vor der euklidischen Abstandsberechnung die Helligkeitsnormierung durchgeführt wurde. *10_radiographs* bedeutet, dass eine Kategorie nur dann für gültig erklärt wurde, wenn sie mindestens zehn Bilder enthielt. In der zweiten Zeile ist zu sehen, wie viele der 5756 Bilder in eine gültige Kategorie eingeordnet werden konnten und somit zum Ergebnis beitragen. Der letzte Wert zeigt den Prozentsatz aller richtig eingeordneten Bilder.



Abb. 5-5: Auflistung der durchgeführten Tests für die verschiedenen Teilmatrizengewichtungen am Beispiel des gewichteten Euklids mit Helligkeitsnormierung bei 5 Wavelet-Passes. Für jede Gewichtung wurde ein kompletter Leaving-One-Out Test durchgeführt.

```

all results:
spie2004_512_5-0.00-0.00-0.00-0.50-norm.10_radiographs
1969/5755 = 34.21
spie2004_512_5-0.00-0.00-0.00-1.00-norm.10_radiographs
1969/5755 = 34.21
spie2004_512_5-0.00-0.00-0.50-0.00-norm.10_radiographs
2713/5756 = 47.13
spie2004_512_5-0.00-0.00-0.50-0.50-norm.10_radiographs
2749/5756 = 47.76
spie2004_512_5-0.00-0.00-0.50-1.00-norm.10_radiographs
2645/5756 = 45.95
spie2004_512_5-0.00-0.00-1.00-0.00-norm.10_radiographs
2713/5756 = 47.13
spie2004_512_5-0.00-0.00-1.00-0.50-norm.10_radiographs
2764/5756 = 48.02
spie2004_512_5-0.00-0.00-1.00-1.00-norm.10_radiographs
2749/5756 = 47.76
spie2004_512_5-0.00-0.50-0.00-0.00-norm.10_radiographs
2749/5755 = 47.77
spie2004_512_5-0.00-0.50-0.00-0.50-norm.10_radiographs
2702/5756 = 46.94
spie2004_512_5-0.00-0.50-0.00-1.00-norm.10_radiographs
2605/5756 = 45.26
spie2004_512_5-0.00-0.50-0.50-0.00-norm.10_radiographs
3104/5756 = 53.93
spie2004_512_5-0.00-0.50-0.50-0.50-norm.10_radiographs
3108/5756 = 54.00
spie2004_512_5-0.00-0.50-0.50-1.00-norm.10_radiographs
2991/5756 = 51.96
spie2004_512_5-0.00-0.50-1.00-0.00-norm.10_radiographs
3063/5756 = 53.21
spie2004_512_5-0.00-0.50-1.00-0.50-norm.10_radiographs
3048/5756 = 52.95
spie2004_512_5-0.00-0.50-1.00-1.00-norm.10_radiographs
3050/5756 = 52.99
spie2004_512_5-0.00-1.00-0.00-0.00-norm.10_radiographs
2749/5755 = 47.77
spie2004_512_5-0.00-1.00-0.00-0.50-norm.10_radiographs
2768/5755 = 48.10
spie2004_512_5-0.00-1.00-0.00-1.00-norm.10_radiographs
2702/5756 = 46.94
spie2004_512_5-0.00-1.00-0.50-0.00-norm.10_radiographs
3094/5756 = 53.75
spie2004_512_5-0.00-1.00-0.50-0.50-norm.10_radiographs
3098/5756 = 53.82
spie2004_512_5-0.00-1.00-0.50-1.00-norm.10_radiographs
3036/5756 = 52.74
spie2004_512_5-0.00-1.00-1.00-0.00-norm.10_radiographs
3104/5756 = 53.93
spie2004_512_5-0.00-1.00-1.00-0.50-norm.10_radiographs
3103/5756 = 53.91
spie2004_512_5-0.00-1.00-1.00-1.00-norm.10_radiographs
3108/5756 = 54.00
spie2004_512_5-0.50-0.00-0.00-0.00-norm.10_radiographs
4169/5755 = 72.44
spie2004_512_5-0.50-0.00-0.00-0.50-norm.10_radiographs
3354/5756 = 58.27

```

spie2004_512_5-0.50-0.00-0.00-1.00-norm.10_radiographs
2948/5756 = 51.22
spie2004_512_5-0.50-0.00-0.50-0.00-norm.10_radiographs
3517/5756 = 61.10
spie2004_512_5-0.50-0.00-0.50-0.50-norm.10_radiographs
3314/5756 = 57.57
spie2004_512_5-0.50-0.00-0.50-1.00-norm.10_radiographs
3156/5756 = 54.83
spie2004_512_5-0.50-0.00-1.00-0.00-norm.10_radiographs
3262/5756 = 56.67
spie2004_512_5-0.50-0.00-1.00-0.50-norm.10_radiographs
3202/5756 = 55.63
spie2004_512_5-0.50-0.00-1.00-1.00-norm.10_radiographs
3116/5756 = 54.13
spie2004_512_5-0.50-0.50-0.00-0.00-norm.10_radiographs
3603/5756 = 62.60
spie2004_512_5-0.50-0.50-0.00-0.50-norm.10_radiographs
3429/5756 = 59.57
spie2004_512_5-0.50-0.50-0.00-1.00-norm.10_radiographs
3213/5756 = 55.82
spie2004_512_5-0.50-0.50-0.50-0.00-norm.10_radiographs
3506/5756 = 60.91
spie2004_512_5-0.50-0.50-0.50-0.50-norm.10_radiographs
3449/5756 = 59.92
spie2004_512_5-0.50-0.50-0.50-1.00-norm.10_radiographs
3314/5756 = 57.57
spie2004_512_5-0.50-0.50-1.00-0.00-norm.10_radiographs
3395/5756 = 58.98
spie2004_512_5-0.50-0.50-1.00-0.50-norm.10_radiographs
3377/5756 = 58.67
spie2004_512_5-0.50-0.50-1.00-1.00-norm.10_radiographs
3307/5756 = 57.45
spie2004_512_5-0.50-1.00-0.00-0.00-norm.10_radiographs
3391/5756 = 58.91
spie2004_512_5-0.50-1.00-0.00-0.50-norm.10_radiographs
3273/5755 = 56.87
spie2004_512_5-0.50-1.00-0.00-1.00-norm.10_radiographs
3166/5755 = 55.01
spie2004_512_5-0.50-1.00-0.50-0.00-norm.10_radiographs
3417/5756 = 59.36
spie2004_512_5-0.50-1.00-0.50-0.50-norm.10_radiographs
3379/5756 = 58.70
spie2004_512_5-0.50-1.00-0.50-1.00-norm.10_radiographs
3324/5755 = 57.76
spie2004_512_5-0.50-1.00-1.00-0.00-norm.10_radiographs
3369/5756 = 58.53
spie2004_512_5-0.50-1.00-1.00-0.50-norm.10_radiographs
3350/5756 = 58.20
spie2004_512_5-0.50-1.00-1.00-1.00-norm.10_radiographs
3303/5756 = 57.38
spie2004_512_5-1.00-0.00-0.00-0.00-norm.10_radiographs
4169/5755 = 72.44
spie2004_512_5-1.00-0.00-0.00-0.50-norm.10_radiographs
3686/5756 = 64.04
spie2004_512_5-1.00-0.00-0.00-1.00-norm.10_radiographs
3354/5756 = 58.27
spie2004_512_5-1.00-0.00-0.50-0.00-norm.10_radiographs
3729/5756 = 64.78
spie2004_512_5-1.00-0.00-0.50-0.50-norm.10_radiographs
3591/5756 = 62.39
spie2004_512_5-1.00-0.00-0.50-1.00-norm.10_radiographs



```

3429/5756 = 59.57
spie2004_512_5-1.00-0.00-1.00-0.00-norm.10_radiographs
3517/5756 = 61.10
spie2004_512_5-1.00-0.00-1.00-0.50-norm.10_radiographs
3436/5756 = 59.69
spie2004_512_5-1.00-0.00-1.00-1.00-norm.10_radiographs
3314/5756 = 57.57
spie2004_512_5-1.00-0.50-0.00-0.00-norm.10_radiographs
3807/5756 = 66.14
spie2004_512_5-1.00-0.50-0.00-0.50-norm.10_radiographs
3675/5756 = 63.85
spie2004_512_5-1.00-0.50-0.00-1.00-norm.10_radiographs
3489/5756 = 60.62
spie2004_512_5-1.00-0.50-0.50-0.00-norm.10_radiographs
3732/5756 = 64.84
spie2004_512_5-1.00-0.50-0.50-0.50-norm.10_radiographs
3651/5756 = 63.43
spie2004_512_5-1.00-0.50-0.50-1.00-norm.10_radiographs
3524/5756 = 61.22
spie2004_512_5-1.00-0.50-1.00-0.00-norm.10_radiographs
3577/5756 = 62.14
spie2004_512_5-1.00-0.50-1.00-0.50-norm.10_radiographs
3549/5756 = 61.66
spie2004_512_5-1.00-0.50-1.00-1.00-norm.10_radiographs
3449/5756 = 59.92
spie2004_512_5-1.00-1.00-0.00-0.00-norm.10_radiographs
3603/5756 = 62.60
spie2004_512_5-1.00-1.00-0.00-0.50-norm.10_radiographs
3549/5756 = 61.66
spie2004_512_5-1.00-1.00-0.00-1.00-norm.10_radiographs
3429/5756 = 59.57
spie2004_512_5-1.00-1.00-0.50-0.00-norm.10_radiographs
3580/5756 = 62.20
spie2004_512_5-1.00-1.00-0.50-0.50-norm.10_radiographs
3592/5756 = 62.40
spie2004_512_5-1.00-1.00-0.50-1.00-norm.10_radiographs
3503/5756 = 60.86
spie2004_512_5-1.00-1.00-1.00-0.00-norm.10_radiographs
3506/5756 = 60.91
spie2004_512_5-1.00-1.00-1.00-0.50-norm.10_radiographs
3506/5756 = 60.91
spie2004_512_5-1.00-1.00-1.00-1.00-norm.10_radiographs
3449/5756 = 59.92

```

5.3.3 Klassifikatorvergleich

Im folgenden werden die Ergebnisse der verschiedenen Klassifikatoren graphisch gegenübergestellt. Hierbei wird die Haupteinteilung anhand der berechneten Wavelet-Passes durchgeführt, da sie die Dimension des Vektorraums für die Klassifikation bestimmen.

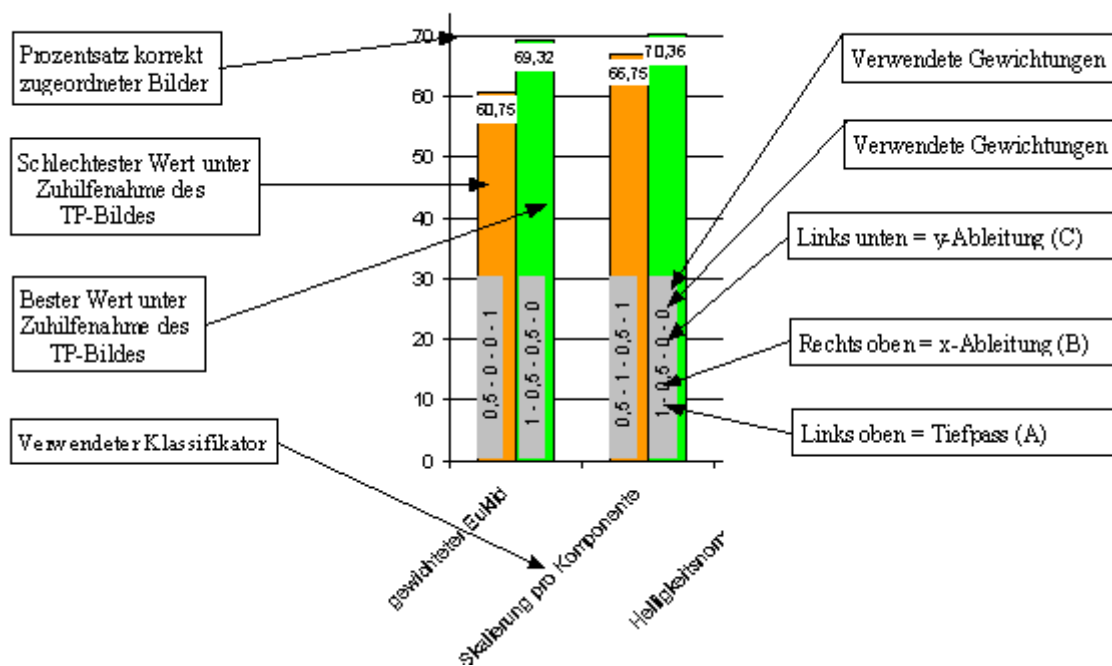
Da, wie auch in Abb. 5-5 zu sehen ist, die Fälle bei denen das Gewicht für links oben gleich Null ist, die Abstandsberechnung also nur auf den Kantenbildern ausgeführt wird, für alle Verfahren sehr schlechte Ergebnisse liefern, wurde dieser Teil aus den Graphiken ausgespart.

Die folgenden Abbildungen zeigen das durch verschiedene Gewichtungen erreichte beste und schlechteste Ergebnis für den jeweiligen Klassifikator, wobei unter schlechtestem Ergebnis jeweils der niedrigste Wert zu verstehen ist, der unter Zuhilfenahme des Tiefpassbildes auftrat.

Zum Vergleich ist der normale euklidische Abstand ohne Teilmatrizen und Gewichtungen sowie die einfache Korrelation des verkleinerten Originalbildes dargestellt. Auf der x-Achse sind die jeweiligen Klassifikatoren dargestellt, die y-Achse zeigt den Prozentsatz der korrekt zugeordneten Bilder. Hierbei wurde für jeden Klassifikator der schlechteste und der beste Wert notiert. Zusätzlich ist angegeben, bei welcher Teilmatrizen-Gewichtung dieser Wert erreicht wurde.

5.3.4 Ergebnisse für die unterschiedlichen Vektorgrößen

Abb. 5-6: Erklärung der Grafiken



1.1.1.4. Fünf Passes (32x32 Pixel)

5.3.4.1.1 Tiefpassbild und Kantenbilder

Der beste Wert für fünf Wavelet-Passes wurde in allen Fällen erreicht, wenn die Kantenbilder nicht in den Vergleich mit einbezogen wurden, der Vergleich also nur auf dem 16x16 Pixel großen Tiefpassbild beruhte. Wurde das Tiefpassbild nicht in den Vergleich mit einbezogen, also mit Null ge-



wichtet, so lag der Prozentsatz korrekt zugeordneter Bilder bei allen Verfahren zwischen 20 und 50 Prozent.

1.1.1.5. Beste und schlechteste Werte

In der Grafik (Abb. 5-7) ist zu erkennen, dass unter Zuhilfenahme des Tiefpassbildes die schlechtesten Werte für den gewichteten euklidischen Abstand, den euklidischen Abstand mit Normierung pro Komponente und den euklidischen Abstand mit Histogrammspreizung bei etwa 65 Prozent liegen. Bei Helligkeitsnormierung und Mahalanobis Abstand liegen die schlechtesten Werte bei etwa 50 bzw. 40 Prozent. Bei diesen Verfahren ist also nur durch den Einsatz des Tiefpassbildes noch keine Verbesserung festzustellen. Trotzdem erreichen auch diese Verfahren ihren besten Wert, wenn nur das Tiefpassbild verglichen wird.

Der schlechteste Wert mit Beachtung des Tiefpassbildes tritt in allen Verfahren auf, wenn das Tiefpassbild vergleichsweise schwach und die Kantenbilder vergleichsweise stark gewichtet werden. Der insgesamt beste Wert wird mit Hilfe der Korrelationsfunktion erreicht.

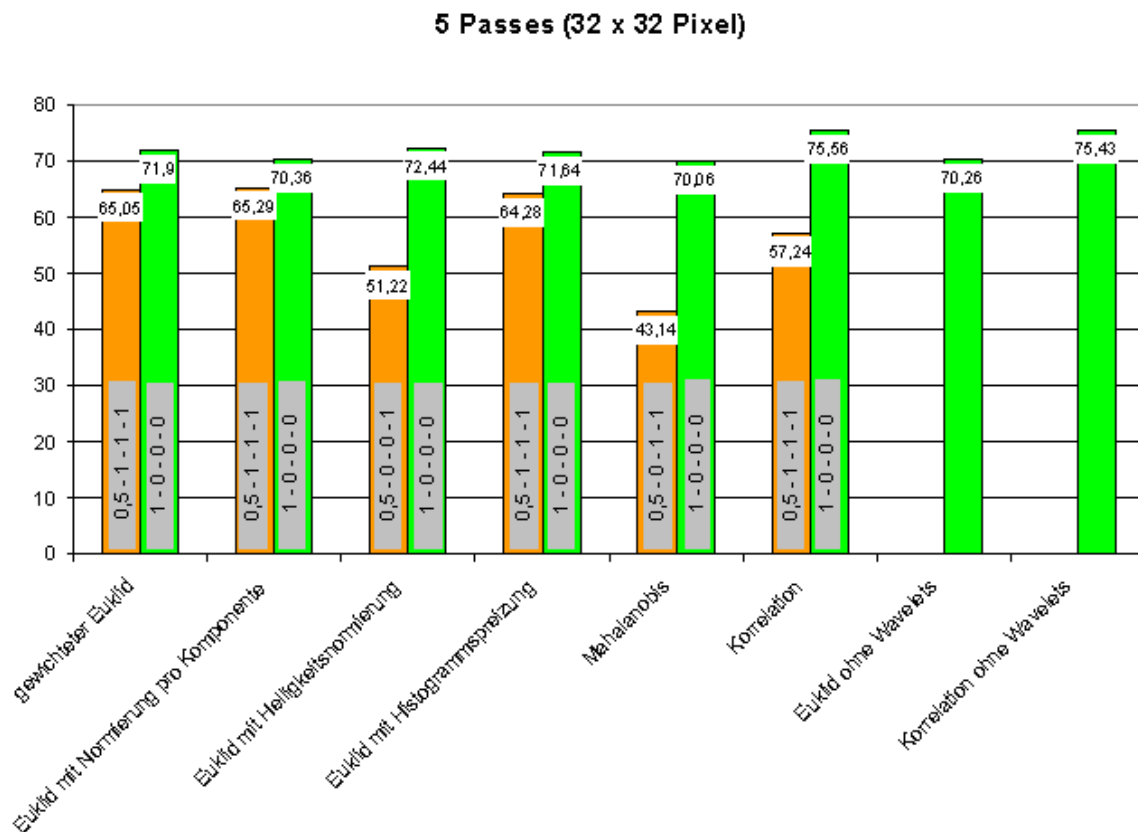
1.1.1.6. Euklidischer Abstand und Normierung

Sowohl die Normierung pro Komponente als auch die Histogrammspreizung führen nicht zu einer Verbesserung des gewichteten euklidischen Abstands, sondern sogar zu einer leichten Abnahme der korrekt zugeordneten Bilder. Nur mit Hilfe der Helligkeitsnormierung konnte eine leichte Verbesserung erreicht werden. Obwohl auch beim gewichteten euklidischen Abstand nur das Tiefpassbild genutzt wird, stellt sich eine Verbesserung gegenüber dem einfachen euklidischen Abstand ein, bei dem keine Wavelet-Transformation genutzt wurde.

1.1.1.7. Überblick

Alle Verfahren liefern recht ähnliche Ergebnisse und finden ihr Minimum und Maximum unter ähnlichen oder gleichen Gewichtungen, haben also im Allgemeinen ein sehr ähnliches Verhalten. Es lässt sich jedoch ein Unterschied feststellen zwischen den Verfahren, die unter Zuhilfenahme des Tiefpassbildes recht stetige Ergebnisse zwischen 65 und 70 Prozent liefern und denen, wo auch bei Nutzung des Tiefpassbildes durch die Kombination der Matrixgewichtungen eine hohe Variabilität im Ergebnis auftritt.

Abb. 5-7: Ergebnisse bei fünf Wavelet-Passes



1.1.1.8. Sechs Passes (16x16 Pixel)

5.3.4.1.2 Tiefpassbild und Kantenbilder

Insgesamt ist zu bemerken, dass zwischen den Gewichtungen, bei denen das Tiefpassbild nicht in den Vergleich mit einbezogen wird und denen, die das Tiefpassbild mit für den Vergleich nutzen, kein so signifikanter Unterschied mehr besteht wie bei fünf Wavelet-Passes. Während bei fünf Passes die Einbeziehung des Tiefpassbildes eine Verbesserung von etwa 30 Prozent korrekt zugeordneter Bilder bewirkte, sind es nun nur noch ca. 20 Prozent.

5.3.4.1.3 Beste und schlechteste Werte

Immer noch zeigen alle Verfahren recht ähnliche Werte. Allerdings werden die schlechtesten und besten Ergebnisse nun bei den verschiedenen Verfahren unter verschiedenen Gewichtungen erreicht. Der beste Wert wird mit Hilfe der Korrelation erreicht, wobei dieser aber nur ein wenig größer ist als der Vergleichswert, der bei einer Korrelationsberechnung ohne vorherige Wavelet-Transformation erreicht wird. Der schlechteste Wert wird mit Hilfe des in [2] genutzten gewichteten euklidischen Abstands erreicht.



Sogar der einfache euklidische Abstand ohne Wavelet-Transformation liefert bessere Werte. Der Mahalanobis-Abstand, der bei fünf Wavelet-Passes noch einen mittelmäßigen Wert lieferte, zeigt nun schlechtere Ergebnisse, die etwa dem gewichteten Euklid ohne Skalierung entsprechen.

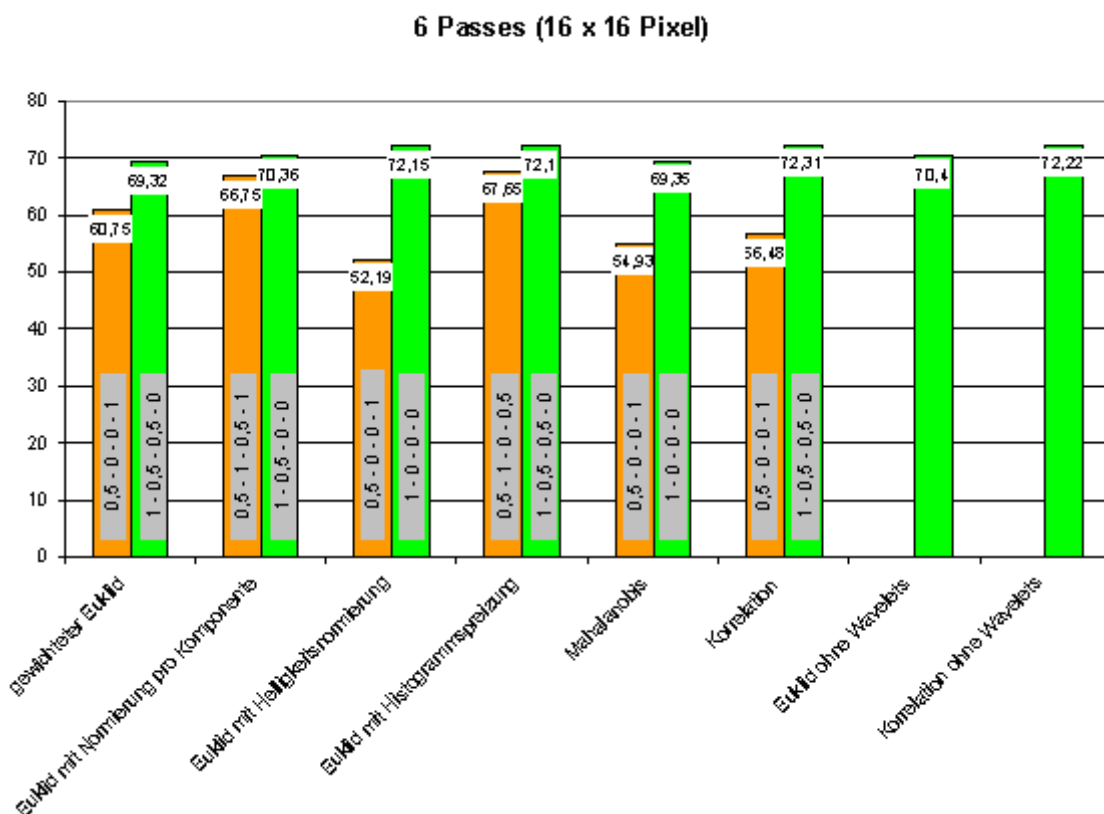
5.3.4.1.4 Euklidischer Abstand und Normierung

Diesmal sind alle Normierungen geeignet, um den Wert des euklidischen Abstands zu verbessern, wobei die Helligkeitsnormierung den besten Wert liefert.

5.3.4.1.5 Überblick

In den meisten Verfahren wird nun auch das Kantenbild für das beste Ergebnis herangezogen, wobei 1-0,5-0,5-0 eine häufig genutzte Gewichtung darstellt. Nur für die Helligkeitsnormierung und den Mahalanobis-Abstand konnte keine vorteilhafte Einbeziehung der Kantenbilder gefunden werden. Es verstärkt sich der Eindruck, dass bei einigen Verfahren die Schwankungen der Ergebnisse im Hinblick auf die unterschiedliche Gewichtung der einzelnen Kantenbilder recht gering sind, während bei anderen Verfahren hierbei hohe Schwankungen auftreten.

Abb. 5-8: Ergebnisse bei sechs Wavelet-Passes



1.1.1.9. Sieben Passes (8x8 Pixel)

Bei sieben Wavelet-Passes wurde auf die Berechnung der Korrelation verzichtet, da jede Teilmatrix nur noch aus 4x4 Pixeln bestand und eine Verschiebung der Matrizen gegeneinander somit nicht mehr sinnvoll ausgeführt werden konnte. Auch für den euklidischen Abstand ohne Wavelets wurde aufgrund der nur 4x4 Pixel keine Abstandsberechnung ausgeführt.

5.3.4.1.6 Tiefpassbild und Kantenbilder

Immer noch liefert die Ignorierung des Tiefpassbildes besonders schlechte Ergebnisse, wobei in allen Verfahren ein Minimum von ca. 40 Prozent korrekt zugeordneter Bilder erreicht wird. Allerdings ist bei Einbeziehung des Tiefpassbildes kein signifikanter Sprung mehr festzustellen, liegen doch sowohl mit als auch ohne Tiefpassbild die durchschnittlichen Werte bei etwa 65 Prozent.

5.3.4.1.7 Beste und schlechteste Werte

Der beste Wert wird diesmal mit Hilfe der Histogrammspreizung erreicht, der schlechteste Wert wiederum mit dem gewichteten Euklid ohne Normierung.

5.3.4.1.8 Euklidischer Abstand und Normierung

Mit Hilfe der Normierungen konnte in jedem Fall eine deutliche Verbesserung der euklidischen Abstandsberechnung festgestellt werden.

5.3.4.1.9 Überblick

Für sieben Wavelet-Passes ist keine allgemein gültige für alle Verfahren vorteilhafte Gewichtung zu erkennen, vielmehr finden alle Verfahren ihr Minimum und Maximum unter unterschiedlichen Gewichtungen. Nur bei Helligkeitsnormierung und Histogrammspreizung stimmen beide Gewichtungen überein.

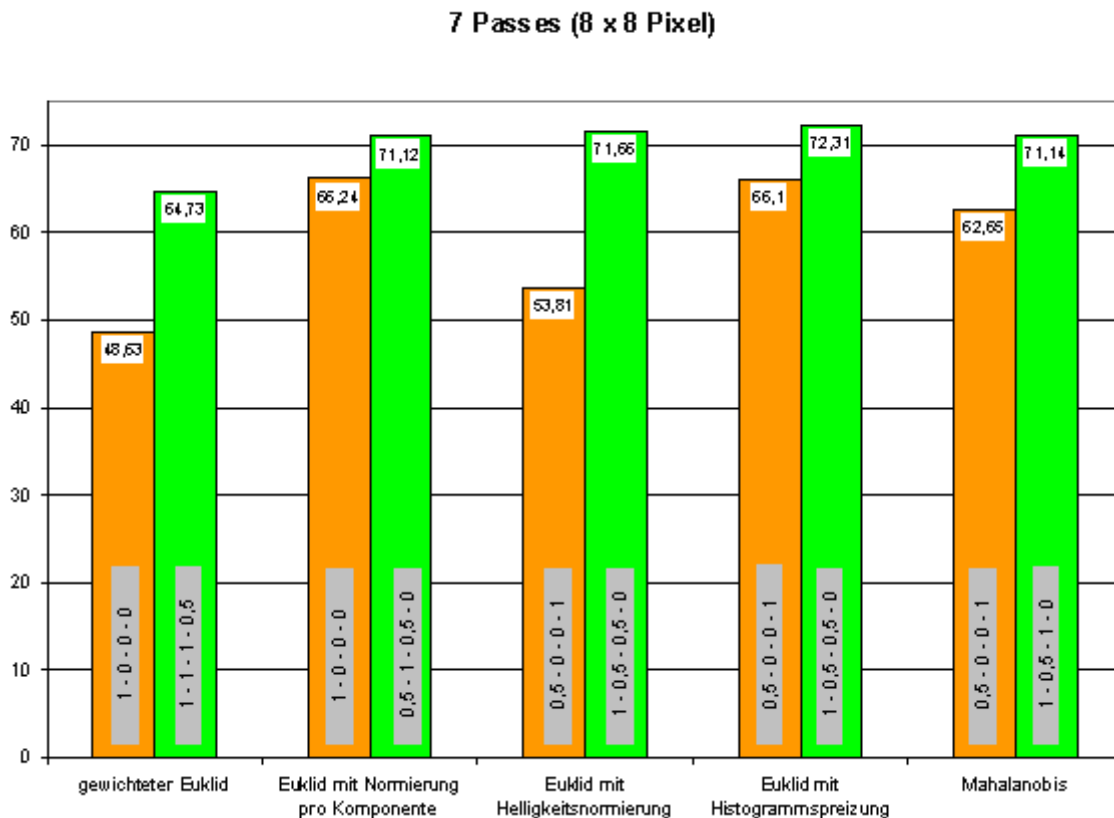
5.3.4.1.10 Gesamt

In Abb.5-10 wurden die besten Ergebnisse jedes Klassifikators für die verschiedenen Wavelet-Passes gegenübergestellt.

5.3.4.1.11 Tiefpassbild und Kantenbilder



Abb. 5-9: Ergebnisse bei sieben Wavelet-Passes



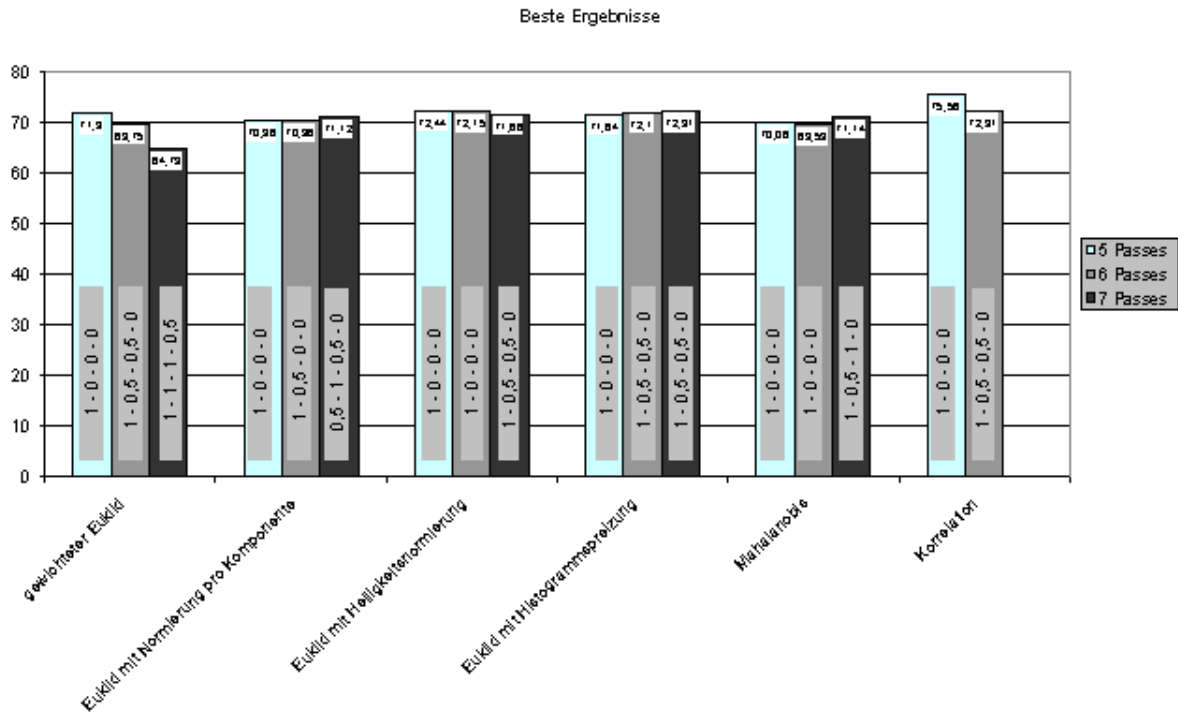
Auffällig ist, dass das Tiefpassbild um so wichtiger ist je mehr Pixel für den Vergleich zur Verfügung stehen. Bei fünf Wavelet-Passes beruht der Vergleich bei allen Verfahren nur auf dem Tiefpassbild, während bei weniger Pixeln auch die Kantenbilder gewinnbringend in den Vergleich einbezogen werden können. Auch der Einbruch, der bei Weglassen des Tiefpassbildes entsteht ist geringer, wenn insgesamt weniger Pixel vorhanden sind.

5.3.4.1.12 Beste und schlechteste Werte

Es ist auffällig, dass keine allgemeingültige Aussage gemacht werden kann, wie viele Wavelet-Passes das beste Ergebnis liefern, da dies für alle Klassifikatorkombinationen unterschiedlich ist.

Beim Mahalanobis-Abstand ist nicht einmal innerhalb desselben Verfahrens ein eindeutiger Trend auszumachen, ob weniger Pixel nun eine schlechtere oder bessere Zuordnung ermöglichen, zumindest wird dort das beste Ergebnis mit Hilfe der wenigsten Pixel erreicht

Abb.5-10: beste Ergebnisse



5.3.4.1.13 Euklidischer Abstand und Normierung

Während der gewichtete euklidische Abstand aus [2] eine markante Verschlechterung zeigt, sobald weniger Pixel für den Vergleich zur Verfügung stehen, zeigt dasselbe Verfahren mit einer vorherigen Normierung in zwei Fällen sogar eine Verbesserung, wenn die Wavelet-Transformation öfter durchgeführt wird. Im dritten Fall ist zumindest keine so gravierende Verschlechterung zu erkennen wie ohne Normierung. Obwohl Helligkeitsnormierung und Histogrammspreizung im Endeffekt sehr ähnliche Ergebnisse liefern, ist ihr Trend genau gegensätzlich.

5.3.4.1.14 Überblick

Stellt man die besten Ergebnisse aller Verfahren aller Wavelet-Passes nebeneinander, so ist zu erkennen, dass sich alle Verfahren in einem sehr ähnlichen Wertebereich bewegen. Der beste Wert der gesamten Tests wird mit Hilfe der Korrelationsfunktion bei fünf Wavelet-Passes erreicht.

5.3.5 Laufzeiten

Die Tests wurden auf einem Rechner mit 2,7 GHz Prozessorleistung und 1GB Arbeitsspeicher durchgeführt. Die Laufzeiten sind in Sekunden angegeben und gelten für einen vollständigen Leaving-One-Out Durchlauf, bei welchem also alle 5756 verwendeten Bilder der Da-



Abb.5-11: Durchschnittliche Laufzeiten der Klassifikatoren für einen vollständigen Leaving-One-Out Durchlauf bei 6 Wavelet-Passes

Klassifikator	Laufzeit in Sekunden
Gewichteter Euklid	170
Gewichteter Euklid mit Helligkeitsnormierung	180
Gewichteter Euklid mit Histogrammspreizung	180
Gewichteter Euklid mit Skalierung pro Komponente	180
Mahalanobis Abstand	3800
Korrelationsfunktion	9500

tenbank als Anfragebild genutzt und jeweils mit allen anderen Bildern verglichen werden. Die angegebenen Zeiten sind Durchschnittswerte aus wiederholten Zeitmessungen.

Abb. 5-11 zeigt eine Gegenüberstellung der Laufzeiten aller getesteten Klassifikatoren.

Es ist zu erkennen, dass die Normierungen vor der euklidischen Abstandsberechnung kaum Rechenzeit verursachen. Bei der Mahalanobis-Distanz, bei welcher eine Division pro Komponente ausgeführt werden muss, steigt der Rechenaufwand erheblich.

Die größte Rechenleistung wird durch die Verschiebungen bei der Korrelationsfunktion benötigt.

Abb. 5-12 zeigt am Beispiel des euklidischen Abstands mit Helligkeitsnormierung den veränderten Zeitaufwand bei einer unterschiedlichen Anzahl von Wavelet-Passes. Die unterschiedlichen Laufzeiten werden durch die Anzahl zu vergleichender Pixel hervorgerufen. Wieder wurden die angegebenen Laufzeiten aus mehreren Versuchen gemittelt.

5.4 Bewertung

5.4.1.1.1 Tiefpassbild und Kantenbilder

Vergleicht man das Verhalten der Verfahren über die verschiedenen untersuchten Wavelet-Passes, so fällt auf, dass die Kantenbilder umso erfolgreicher in den Vergleich einbezogen werden können, je mehr Wavelet-Iterationen durchgeführt wurden. Dies bedeutet, dass der Vergleich auf weniger Pixeln durchgeführt wird, welche jeweils eine größere Anzahl von Originalpixeln zusammenfassen. Kantenbilder sind bei der Abstandsberechnung sehr empfindlich gegenüber Verschiebungen. Schon eine kleine Verschiebung kann dazu führen, dass Kanten nicht mehr übereinanderliegen und somit der größtmögliche Abstand zwischen zwei Bildern berechnet wird. Je mehr Pixel des Ausgangsbildes nun durch ein Pixel des zu vergleichenden Bildes dargestellt werden, desto mehr werden die Verschiebungen kompensiert. Bei der Korrelation dagegen, bei welcher Verschiebungen durch das Verfahren ausgeglichen werden, ergibt sich der erwartete Effekt, dass mehr Pixel, welche eine genauere Darstellung des Bildes ermöglichen auch zu besseren Ergebnissen führen.

Abb. 5-12: Durchschnittliche Laufzeiten bei Verwendung des Euklidischen Abstand mit Helligkeitsnormierung für einen vollständigen Leaving-One-Out Durchlauf

Wavelet Passes	Pixelanzahl	Laufzeit in Sekunden
5	32x32	570
6	16x16	180
7	8x8	65

Die meisten Verfahren zeigen keine gravierenden Änderungen, wenn die Gewichtung der Teilmatrizen verändert wird. Minimaler und maximaler Wert liegen hier dicht nebeneinander. Eine Ausnahme bildet der Mahalanobis-Abstand. Dies ist auf die problematische Abstandsberechnung bei den Kantenbildern zurückzuführen. Da in einem Kantenbild die meisten Pixel schwarz sind und die Kanten nur durch wenige helle Pixel charakterisiert werden, wird der berechnete Mittelwert und die Standardabweichung sehr klein. Bei der Abstandsberechnung hat eine leicht verschobene Kante nun einen doppelt negativen Effekt, da nicht nur der Abstand vom Hintergrund in die Distanz eingeht, sondern dieser noch durch Division durch die sehr kleine Standardabweichung vergrößert wird.

Der Mahalanobis-Abstand, der im Vergleich zum euklidischen Abstand rechentechnisch einen größeren Aufwand bedeutet, da die



Standardabweichung berechnet werden muss, liefert nicht die erwünschten besseren Ergebnisse, so dass stattdessen lieber der euklidische Abstand mit vorheriger Normierung verwendet werden sollte.

Aufgrund der gesamten Kantenbild-Problematik kann die Aussage getroffen werden, dass für eine sinnvolle Klassifikation das Tiefpassbild in jedem Fall für den Vergleich genutzt werden muss.

5.4.1.1.2 Beste und schlechteste Werte

Den mit Abstand größten rechnerischen Aufwand erfordert die Korrelationsfunktion, welche auch insgesamt das beste Ergebnis liefert. Bedenkt man allerdings die erhöhte Rechenzeit und vergleicht dies mit dem relativ geringen Genauigkeitsgewinn bei der Klassifikation, so stellt sich die Frage, ob nicht auch hier besser der euklidische Abstand mit einer Normierung verwendet werden sollte.

5.4.1.1.3 Euklidischer Abstand und Normierung

Der in [2] genutzte euklidische Abstand konnte durch die vorher durchgeführten Normierungen deutlich verbessert werden. Hierbei ist aber anzumerken, dass zumindest bei fünf Wavelet-Passes nicht jede Normierung automatisch zu einer Verbesserung führte. Dennoch scheint es sehr sinnvoll eine der Normierungen einzusetzen, da deutlich bessere Ergebnisse erzielt werden können, besonders wenn der Vergleich auf weniger Pixeln beruht.

5.4.1.1.4 Überblick

Insgesamt können keine verallgemeinernden Aussagen getroffen werden. Es zeigt sich weder eine allgemein gültige vorteilhafte Teilmatrizengewichtung noch eine besonders geeignete Anzahl von Wavelet-Passes. Ein besonders geeignetes Verfahren ist schwer zu bestimmen, da die Unterschiede im Ergebnis zwischen den einzelnen Verfahren nicht besonders groß sind.

Auch zu den bereits vorhandenen Verfahren, welche die Wavelet-Transformation nicht zur Merkmalsextraktion nutzen, ist kein gravierender Unterschied, sondern nur eine leichte Verbesserung festzustellen.

Soll das Verfahren der Wavelet-Transformation unbedingt eingesetzt werden, so sollten folgende Parameter genutzt werden:

- Sieben Wavelet-Passes

- Euklidischer Abstand
- Normierung durch Histogrammspreizung

Durch die sieben Wavelet-Passes erfolgt der Vergleich auf einem kleinen Merkmalsvektor, was einen Geschwindigkeitsvorteil bedeutet. Zwar liefern andere Verfahren bei weniger Wavelet-Passes ein minimal besseres Ergebnis, allerdings ist dies auch mit einer größeren Rechenzeit für den Vergleich verbunden. Das hierbei beste Ergebnis, das von der Korrelation bei fünf Wavelet-Passes erreicht wird, benötigt ein Vielfaches der Rechenzeit der anderen Verfahren.

6 Diskussion und Fazit

In dieser Arbeit wurde die Wavelet-Transformation zur Merkmalsextraktion von Bilddaten genutzt und ein Bildvergleich anhand des so berechneten Merkmalsvektors durchgeführt. Es wurde getestet, inwieweit das implementierte Verfahren zur Klassifikation medizinischer Bilddaten geeignet ist. Dazu wurde die Wavelet-Transformation, die Merkmalsextraktion zur Bildung der Merkmalsvektoren sowie deren Einbindung in die Datenbank und die Distanzfunktionen mit entsprechender Datenbankanbindung implementiert.

Durch die Anwendung der Wavelet-Transformation bei der Merkmalsextraktion konnten leichte Verbesserungen in der anschließenden Klassifikation erreicht werden. Dies war jedoch nur möglich, da zu dem eigentlich vorgegebenen gewichteten euklidischen Abstand Verbesserungen bzw. Alternativen getestet wurden. Insgesamt ist die Verbesserung gegenüber ähnlichen Verfahren, die nicht die Wavelet-Transformation zur Merkmalsextraktion nutzen aber so gering, dass in Frage gestellt werden muss, ob die benötigte, zeitaufwändige Berechnung aller Wavelet-Koeffizienten überhaupt sinnvoll ist.

Ein weiteres Problem besteht in der Verwendung der durch die Wavelet-Transformation erzeugten Kantenbilder. Diese sind sehr empfindlich gegenüber Verschiebungen und erschweren deshalb eine korrekte Distanzberechnung. Die Tests bestätigen, dass Kantenbilder nur gewinnbringend in den Vergleich eingebracht werden können, wenn sie sehr stark komprimiert sind und nur noch aus wenigen Pixeln bestehen. Es stellt sich deshalb die Frage, ob es überhaupt sinnvoll ist, diese größtenteils nicht genutzten Bildwerte im Merkmalsvektor zu speichern.

Das vorliegende Verfahren liefert also nur minimal bessere Ergebnisse, wobei der Rechenaufwand erheblich erhöht wird. Als Fazit sollte vom hier getesteten Verfahren also eher abgeraten werden.

Normalerweise liefert das Verfahren der Wavelet-Transformation sehr gute Ergebnisse im Bereich der Bildkompression, wobei nach



einem speziellen Algorithmus kleine Wavelet-Koeffizienten durch null ersetzt werden, während die größeren gespeichert werden. Durch dieses Verfahren ist eine relativ hohe Kompression bei relativ wenig Informationsverlust möglich. Das hier vorliegende Verfahren macht von diesen Möglichkeiten keinen Gebrauch. Statt die wichtigsten Kanteninformationen zu speichern, werden alle verworfen. Die Wavelet-Transformation wird nur zur Verminderung der Bildpunkte und zur Erzeugung der drei Kantenbilder genutzt. Da diese drei Kantenbilder, wie durch die Tests erwiesen, keinen besonderen Zweck erfüllen, könnten sie eigentlich ebenfalls verworfen werden, ohne das Klassifikationsergebnis besonders zu verschlechtern.

7 Literaturverzeichnis

- [1] <http://www.irma-project.org/>
- [2] James Ze Wang, Gio Wiederhold, Oscar Firschein, Sha Xin Wie, Content based image indexing and searching using daubechies' wavelets, International Journal on Digital Libraries, 1997, Seiten 311-328
- [3] Thomas M. Lehmann, Henning Schubert, Daniel Keysers, Michael Kohlen, Berthold B.Wein, The IRMA code for unique classification of medical images, Proceedings SPIE 2003
- [4] Ingrid Daubechies, Ten Lectures on Wavelets CBMS-NSF Lecture Notes Nr. 61, SIAM, 1992
- [5] Thomas Lehmann, Walter Berschelp, Erich Pelikan, Rudolf Repges, Bildverarbeitung in der Medizin, Kap. 9 Springer 1997
- [6] Brani Vidakovic, Peter Mueller, Wavelets for Kids – A Tutorial Introduction, ISDS, Duke University, 1991
- [7] William H. Press, Saul A. Teukolsky, William T. Vetterling, Brian P. Flannery, Numerical Recipes in C++, Cambridge, 2002, Seiten 596-607
- [8] Mark Oliver Güld, Daniel Keysers, Thomas Deselaers, Marc Leisten, Henning Schubert, Hermann Ney, Thomas M. Lehmann, Comparison of Global Features for Categorisation of Medical Images, Proceedings SPIE 2004
- [9] Yves Nievergelt, Wavelets Made Easy, Birkhauser Boston, 1999
- [10] Wavelets-Transformationscodierung <http://www.fh-jena.de/contrib/fb/et/personal/ansorg/ftp/wavelet/wavelet.htm>.
- [11] W. Hillen, Digitale Bildverarbeitung in der Medizintechnik, Fachhochschule – Aachen (Jülich), Umdruck zur Vorlesung, 1998.

Referent: Prof. Dr. rer. nat. W. Hillen

Koreferent: Dipl.-Inform. Christian Thies

Erklärung

Hiermit erkläre ich, dass ich die vorliegende Diplomarbeit selbständig angefertigt habe. Es wurden nur die in der Arbeit ausdrücklich benannten Quellen und Hilfsmittel benutzt. Wörtlich oder sinngemäß übernommenes Gedankengut habe ich als solches kenntlich gemacht.

Ort, Datum

Unterschrift



