Klassifikation von Gangmustern mittels Support Vektor Maschinen

Vorgelegt am
Institut für Sportwissenschaft
Patrick Hegen
Johannes Gutenberg Universität Mainz

Von
Katja Kempter und Lisa Fank
Seminar „Bewegende Systeme analysieren und beschreiben“
M. Sc. Sportwissenschaft

Mainz, 18.April 2013

Inhaltsverzeichnis

Abbildungsverzeichnis II

1 Problemstellung 1

2 Theoretische Grundlagen 2

2.1 Die Ganganalyse 2

2.2 Mustererkennung und Klassifizierung 3

2.3 Forschungsstand 4

2.4 Support Vector Machines 6

2.4.1 Lineare Support Vektor Maschinen 6

2.4.2 Nicht-lineare Support Vektor Maschinen 7

2.4.3 Kernfunktionen 8

2.4.4 Multiklassifikation mit Support Vektor Maschinen 9

3 Methodik 10

3.1 Untersuchungsinstrumente 10

3.1.1 Qualisys Track Manager – QTM 10

3.1.2 VISUAL3D TM 12

3.2 Untersuchungsdurchführung 13

3.2.1 Datenverarbeitung 14

4 Klassifikation mittels Support Vektor Maschine 16

4.1 Training mit allen Datensätzen 16

4.2 Training mit einem Teil der Datensätze 18

4.3 Diskussion der Ergebnisse 19

5 Zusammenfassung und Ausblick 21

6 Literatur 22

Anhang 24

Erklärung 26

Abbildungsverzeichnis

Abbildung 1: Lineare Trennung mittels Hyperebene (vgl. Abe, 2005, S.17) 7
Abbildung 2: Trennung mittels Hyperebene durch Transformation in einen höherdimensionalen Raum (vgl. Schölkopf& Smola, 2002, S.15) 8
Abbildung 3: Markererkennung mit Qualisys 10
Abbildung 4: Bezeichnung der einzelnen Marker 11
Abbildung 5: Model-builder Visual 3D 13
Abbildung 6: Festlegen der Datenmenge für die Erstellung des Modells 18

1 Problemstellung

Das Erkennen und Einordnen von Mustern hat sowohl in unserem alltäglichen Handeln als auch in der Wissenschaft eine immense Bedeutung. Im Sport spielt die laienhafte Einordnung beispielsweise bei Ballsportarten eine große Rolle. Wir versuchen bestimmte Bewegungsmuster des Gegners oder Mitspielers zu kategorisieren, um möglichst adäquat und schnell auf ihre Aktion reagieren zu können. Vergleichbar werden auch bei wissenschaftlichen Fragestellungen versucht, Muster zu erkennen und einzuordnen. Das Ziel der Wissenschaft bei der Mustererkennung ist demnach, Muster zu identifizieren und klassifizieren. So können in der Sportwissenschaft beispielsweise abweichende Bewegungsmuster identifiziert werden, um eine individuell passende Intervention zu konzipieren. Das Ziel verschiedener Methoden und Verfahren ist deshalb eine möglichst fehlerfreie Klassifizierung von Objekten. In der vorliegenden Arbeit sollen Gangmuster, die durch dreidimensionale Vektoren beschrieben werden, den Probanden zugeordnet werden. Die Gelenkwinkel sind die Komponenten der Vektoren. Als Klassifikationsverfahren werden Support Vektor Maschinen (SVM) eingesetzt. Ziel der Anwendung des Verfahrens ist dabei die Minimierung der Fehlklassifikationen.

Zunächst einmal werden die elementaren Begrifflichkeiten in der Verwendung von Support Vektor Maschinen erläutert. Relevant für die vorliegende Arbeit sind dabei die Ganganalyse sowie die Mustererkennung. Selbstverständlich wird besonderer Wert auf die theoretischen Hintergründe von Support Vektor Maschinen gelegt, die Funktionsweise dargelegt und die unterschiedlichen Arten an SVM, die für die folgende Untersuchung relevant sind, beschrieben. In Kapitel 3 wird dann die Vorgehensweise der Untersuchung geschildert. Hierzu werden zu Beginn des Kapitels die Untersuchungsinstrumente vorgestellt und erläutert, um im Anschluss den Untersuchungsaufbau, die Untersuchungsdurchführung und die Vordatenverarbeitung darzulegen. Das anschließende Kapitel beschäftigt sich dann mit der Anwendung der Support Vektor Maschine auf das vorliegende Klassifikationsproblem. In diesem Kapitel werden zudem die Ergebnisse dargestellt, analysiert und interpretiert. Im Anschluss wird noch einmal die gesamte Untersuchung kritisch reflektiert.

2 Theoretische Grundlagen

2.1 Die Ganganalyse

Die Analyse und Beschreibung des menschlichen Gangs reicht bis in die Zeit Aristoteles (384 v. Chr. Bis 322 v. Chr.) zurück. Besonders hervorzuheben in der Entwicklung der Ganganalyse ist das Werk *de motuanimalum* von Borelli (1682). In dieser Untersuchung wird zum ersten Mal der Körperschwerpunkt bestimmt. (vgl. Mikut, 2001, S.4). Moderne wissenschaftliche Messungen und Beschreibungen von Prinzipien des Ganges sind in den Werken der Gebrüder Weber *Die Mechanik der Gehwerkzeuge* (1836) zu finden (vgl. Mikut, 2001,S.4). Die Methode zur Bestimmung von dreidimensionalen Ganganalysedaten aus fotografischem Material stellten erstmals Braune und Fischer 1889 vor (vgl. Mikut, 2001,S.4). Die bisher genannten Autoren hatten für Ihre Messergebnisse einen hohen Aufwand für die Messung und ihre Auswertung. Scherb entwickelte in der ersten Hälfte des letzten Jahrhunderts ein Laufband, welches es ermöglicht, den Probanden ohne Ortsänderung zu untersuchen (vgl. Mikut, 2001, S.4). Der größte Durchbruch gelang 1980 durch die Kombination moderner Computertechnologie und den Analysemethoden der fotografischen Technik (vgl. Mikut, 2001,S.4). Heute spricht man hier von der instrumentellen Ganganalyse (vgl. Mikut, 2001,S.4).

Die Ganganalyse heute ist eine Analysemethode zur Untersuchung des menschlichen Gangbildes mit dem Ziel, Störungen und ihre Ursachen festzustellen und mögliche Korrekturmaßnahmen festzulegen (vgl. Bachmann 2008,S.1). „Das Gangmuster eines jeden Menschen ist individuell und doch von Schritt zu Schritt zu einem gewissen Grad variabel. Deshalb ist die Grenze zwischen gesunden und pathologischen Gangbildern oft unklar“ (Bachmann 2008, S.1). Die visuelle Analyse kann die Komplexität des Ganges nicht vollständig erfassen. Für den Beobachter kommen in kürzester Zeit zu viele Informationen zusammen, sodass ein gleichzeitiges Erkennen und korrektes Verarbeiten der Informationen nicht möglich ist. Um das Gangmuster eines Menschen im vollen Umfang verstehen zu können sind technische Hilfsmittel notwendig. Dies führte zur Entwicklung der instrumentellen Ganganalyse (vgl. Bachmann 2008, S.1). Heute stehen uns dank der Entwicklung der instrumentellen Ganganalyse Messsysteme zur Verfügung, die die Bewegungen der einzelnen Körpersegmente präzise dreidimensional erfassen können (Kinematik), die Informationen der wirkenden Kraft vom Boden liefern (Kinetik) und die Aktivitäten

der Muskulatur aufzeigen (Elektromyografie). Ebenso können durch Druckmessplatten schnelle Bewegungen bei Stossbewegungen und Landungen erfasst werden (vgl. Bachmann 2008, S.1).

2.2 Mustererkennung und Klassifizierung

Der Informatiker Heinrich Niemann definiert die Mustererkennung wie folgt:

> „Die Mustererkennung beschäftigt sich mit den mathematisch-technischen Aspekten der automatischen Verarbeitung und Auswertung von Mustern. Es wird für ein physikalisches Signal (z.B. Sprache, Bild, Messwert) eine Symbolkette (bzw. in eine formale Datenstruktur) berechnet. Dazu gehört sowohl die Klassifikation einfacher Muster als auch die Klassifikation und Analyse komplexer Muster“ (Niemann 2003, S.13f).

Man unterscheidet in der Literatur verschiedene Ansätze der Mustererkennung: syntaktische, statistische und strukturelle Mustererkennung (vgl. Franz, 2008):

Syntaktisch

Dinge werden so durch Folgen von Symbole beschrieben, dass Objekte der gleichen Kategorie dieselben Beschreibungen aufweisen (vgl. Franz, 2008).

Statistisch

Ziel ist es hier, ein Objekt in die Kategorie mit der höchsten Wahrscheinlichkeit einzusortieren (vgl. Franz, 2008). Statt Merkmale nach vorgefertigten Regeln auszuwerten, werden sie hier einfach als Zahlenwert gemessen und in einem Merkmalsvektor zusammengefasst (vgl. Franz, 2008). Eine mathematische Funktion ordnet dann jedem denkbaren Merkmalsvektor eine Kategorie zu (vgl. Franz, 2008). Ein Beispiel sind hier die Support Vektor Maschinen (vgl. Franz 2008).

Strukturell

Sie verbinden verschiedene syntaktische und/oder statistische Verfahren zu einem einzigen neuen Verfahren (vgl. Franz, 2008). Die grundlegende Merkmalserkennung wird dabei statistischen Verfahren überlassen (vgl. Franz, 2008).

Der Prozess der Mustererkennung wird in drei Schritte unterteilt:

1. Aufnahme (Messung) einschließlich Vorverarbeitung,
2. Merkmalsextraktion,
3. Klassifikation.

Ein Muster besitzt nach Niemann (2003, S.20) Merkmale, die für seine Zugehörigkeit zu einer Klasse charakteristisch sind. „*Die Merkmale bilden für Muster einer Klasse einen einigermaßen kompakten Bereich im Merkmalsraum. Die von Merkmalen verschiedener Klassen eingenommenen Bereiche sind einigermaßen getrennt*“ (Niemann 2003, S.14). Bei der Klassifikation von Mustern wird jedes Muster als ein Ganzes betrachtet und unabhängig von anderen Mustern einer genauen Klasse von vielen möglichen Klassen zugeordnet. (vgl. Niemann 2003, S.15f). Die anschließende Analyse von Mustern ordnet jedem Muster eine individuelle symbolische Beschreibung zu (vgl. Niemann 2003, S.16). Sie besteht aus einem Netzwerk, die aus dem Muster berechnete Instanzen von Konzepten eines Modells enthält (vgl. Niemann 2003, S.16).

2.3 Forschungsstand

Untersuchungen mittels moderner Software und einem Kamerasystem gehören zu einem jungen Bereich der Wissenschaft. Für einen besseren Einblick in das Themengebiet soll im Folgenden ein Überblick über bereits durchgeführte Studien im Bereich der Ganganalyse gegeben werden.

In einer Untersuchung von Schöllhorn, Nigg, Stefanyshyn & Liu (2002) konnten spezifische Charakteristika im Gangmuster einer Person erkannt werden. Dabei wurden kinetische und kinematische Daten der Gangmuster von 13 Frauen analysiert, die Schuhe mit unterschiedlichen Absätzen getragen haben (vgl Schöllhorn et al., 2002, S.181). Ebenso wurden die Vor- und Nachteile von zeitlich kontinuierlichen Daten und zeitdiskreten Daten herausgearbeitet (vgl. Schöllhorn et al., 2002, S.181). Mittels einer Kistler Kraftmessplatte (1000Hz) und vier Kameras (200Hz) wurden die kinematischen und kinetischen Daten erhoben (Schöllhorn et al., 2002, S.181f). Die Ergebnisse zeigten, dass bei beiden Ansätzen, mit den kontinuierlichen und den zeitdiskreten Daten, die individuellen Gangmuster erkannt werden konnten (Schöllhorn et al., 2002, S.180ff).

Im Jahre 2008 konnten Janssen et al. (S.84) die Personenerkennung und die Erkennung von Emotionszuständen anhand des Gangmusters mittels künstlicher neuronaler Netze zuordnen. Die Untersuchung bestand aus zwei Experimenten (vgl. Janssen et al., 2008, S.84). Zunächst wurden in einem Versuch die emotionalen Zustände der 22 Probanden gemessen (vgl. Janssen et al., 2008, S.84). Dabei wurden kinematische Daten erhoben, indem die Probanden mit unterschiedlichen Emotionszuständen (normal, glücklich, traurig, verärgert) über eine Kraftmessplatte gegangen sind (vgl. Janssen et al., 2008, S.84). Zu 95% konnten den Probanden die individuellen Gangmuster zugeordnet werden (Janssen et al., 2008, S.84). Im zweiten Versuch wurde derselbe Versuchsaufbau sowie die- durchführung wie im ersten Experiment gewählt (vgl. Janssen et al., 2008, S.86). Dieses Mal wurden sowohl kinematische als auch kinetische Daten untersucht (vgl. Janssen et al., 2008, S.86). Es stand die Frage im Mittelpunkt, ob es Unterschiede im Gangmuster von 16 Probanden gibt, die einmal keine Musik hören, dann beruhigende Musik und im nächsten Durchgang anregende Musik vorgespielt bekamen (vgl. Janssen et al., 2008, S.86). Auch hier zeigten die Ergebnisse hohe Zuordnungen. (Janssen et al., 2008,S.86)

In einer Studie im Jahre 2010 konnten Logan, et al. (S.339-345) nachweisen, dass das Sehen einen Einfluss auf den Gang hat. Dabei wurden 12 Probanden im Stehen und im Gehen (1 bzw.5 km/h) auf einem Laufband getestet, welches einen Meter vor einer Leinwand stand (vgl. Logan et al., S.339). Die Leinwand bestand aus einem schwarzen Hintergrund, auf welchem weiße Dreiecke in unterschiedlicher Geschwindigkeit sichtbar wurden (vgl. Logan et al., S.339). Das Sichtfeld des Probanden wurde durch eine Brille eingegrenzt (vgl. Logan et al., S.339). Durch am Körper platzierte Marker sollte die Rolle des Sehens beim Gehen nachgewiesen werden (Logan et al.,2010, S.339f). Die Ergebnisse zeigten, dass nur kleine Unterschiede zwischen dem Gehen und Stehen im Rumpf auftreten (vgl. Logan et al., S.343). Jedoch konnte eine horizontale Verschiebung von Hüfte und Schultern vom Stehen im Übergang zur Bewegung beobachtet werden (vgl. Logan et al., S.343). Die Autoren schlossen daraus, dass ein solcher Anstieg mit einer erhöhten Kopplung an das visuelle System einhergeht (Logan et al., 2010,S.343ff).

2.4 Support Vector Machines

Support Vektor Maschinen (SVM) gehören zum Feld der statistischen Mustererkennung und sind zu den Methoden des maschinellen Lernens einzuordnen. In diesem Gebiet sind sie von zunehmender Bedeutung (vgl. Gieseke 2012, S.25). Sie dienen der binären Klassifikation von Objekten mit dem Ziel, das Fehlklassifikationsrisiko zu minimieren (vgl. Decker, 2008, S.53). SVM werden dazu mit vorher klassifizierten Features, den Trainingsdaten, trainiert und werden deshalb zu den Verfahren des überwachten Lernens gezählt (Schölkopf& Smola, 2002, S.1f.). In vielen Methoden des maschinellen Lernens wird der Klassifikator so trainiert, dass die Klassifikationsleistung für die Trainingsdaten maximiert wird (vgl. Abe, 2005, S.2). In der Regel wird dadurch jedoch die Gesamtheit der Daten vernachlässigt und es resultiert eine Überanpassung an die Trainingsdaten, die mit einer verminderten Generalisierungsfähigkeit einhergeht (vgl. Abe, 2005, S.2). Im Gegensatz dazu ist die übergeordnete Idee der Support Vektor Maschine, für ein zwei-Klassen Problem die Generalisierungsfähigkeit durch die direkte Entscheidungsfunktion zu maximieren. (vgl. Abe, 2005, S.2). Die Entscheidungsfunktion bildet die sogenannte Hyperebene, die die Datensätze optimal in zwei Klassen teilen soll (vgl. Abe, 2005, S.2). Die Aufgabe der SVM ist also die optimale Trennung der Daten in zwei Klassen mittels Hyperebene. Lässt sich die Klassifizierung mittels Hyperebene im sogenannten input space –einem m-dimensionalen Raum, in dem die Trainingsdaten abgebildet werden- linear durchführen, spricht man von Linearen Support Vektor Maschinen (vgl. Abe, 2005, S.15). Lassen sich die die Trainingsdaten im input space linear nicht voneinander abgrenzen, werden diese in einen höherdimensionalen Raum abgebildet, um in diesem Raum eine Hyperebene zu bestimmen (vgl. Abe, 2005, S.15; Schölkopf& Smola, 2002, S.3). Diese Vorgehensweise wird mit nicht-linearen Support Vektor Maschinen ausgeführt.

2.4.1 Lineare Support Vektor Maschinen

Gegeben sind M m-dimensionale Trainingsinputs $x_i (i = 1, \dots, M)$ und die zugehörigen Labels $y_i = 1$ für Klasse 1 und $y_i = -1$ für Klasse 2 (vgl. Abe, 2005, S.15; Schölkopf& Smola, 2002, S.1f). Gesucht wird nun die optimale Lage der Hyperebene, die die Daten linear trennt (vgl. Abe, 2005, S.15f; Schölkopf& Smola, 2002, S.1f). Die Hyperebene soll dabei so platziert werden, dass sich zwischen den beiden Kassen ein möglichst breites Band aufspannen lässt (vgl. Abe, 2005, S.16f;

Schölkopf& Smola, 2002, S.11f). Der Abstand zwischen der Hyperebene und den der Hyperebene am nächsten gelegenen Trainingsdaten, wird *Margin* genannt. Die Hyperebene mit maximalem *Margin* wird als *Optimale Hyperebene* bezeichnet (vgl. Abb. 1). Da sich in diesem Fall keine Trainingsvektoren innerhalb des Bandes befinden, wird auch von sogenannten *Hard-Margin-SVM* gesprochen (vgl. Abe, 2005, S.15). Trainingsvektoren, die auf dem Rand des Bandes liegen, werden zur Darstellung der Hyperebene herangezogen und werden als Support Vektoren $\overrightarrow{x_{sp}}$ bezeichnet (vgl. Schölkopf& Smola, 2002, S.13f). Die Optimale Hyperebene entspricht der Entscheidungsfunktion $f(x_i)$, die die Trainingsvektoren $\overrightarrow{x_i}$ voneinander trennt und die im Trainingsprozess bestimmt wird (vgl. Schölkopf& Smola, 2002, S.14). Diese basiert auf der Berechnung des Skalarprodukts zweier Vektoren (vgl. Schölkopf& Smola, 2002, S.14).

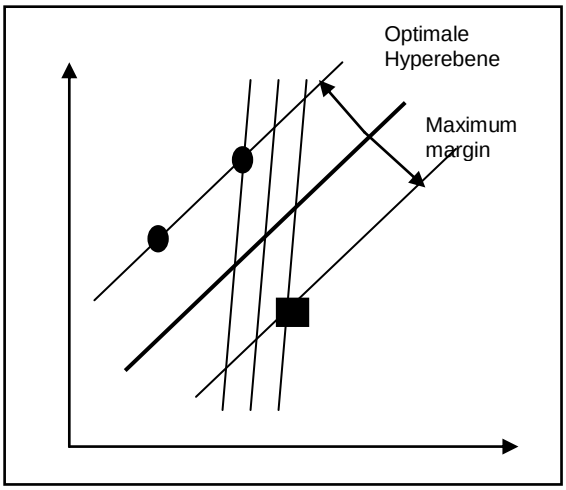


Abbildung 1: Lineare Trennung mittels Hyperebene (vgl. Abe, 2005, S.17)

2.4.2 Nicht-lineare Support Vektor Maschinen

Sind die Daten linear nicht trennbar, liefert die Hard-Margin-SVM keine Lösung (vgl. Abe, 2005, S.22). Die Support Vektor Maschine muss aus diesem Grund nicht-linear verallgemeinert werden (vgl. Abe, 2005, S.22). Die Trainingsdaten werden dabei mittels nicht-linearer Transformation in einen höherdimensionalen Feature- Raum abgebildet und die trennende Hyperebene wird dort mittels Skalarproduktberechnung konstruiert (s. Abb. 2), sodass weiterhin eine hohe Generalisierungsfähigkeit gegeben ist (vgl. Abe, 2005, S.25; Schölkopf& Smola, 2002, S.15). Dies ergibt wiederum eine nicht-lineare Grenzlinie im ursprünglichen Input- Raum (vgl. Abe, 2005, S.15f; Schölkopf& Smola, 2002, S.1f).

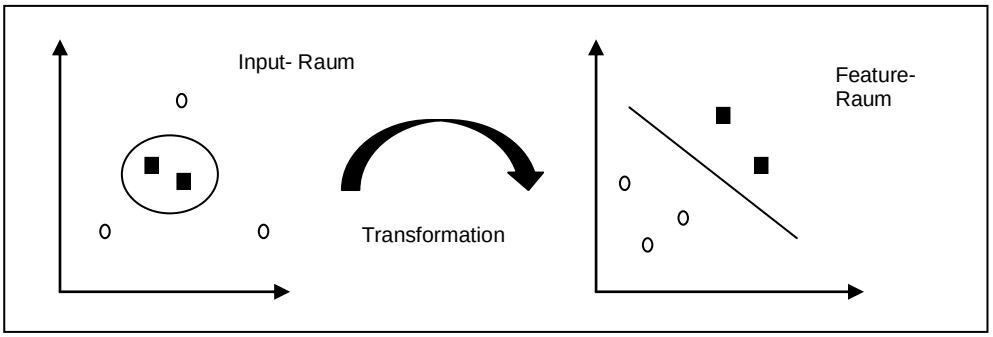


Abbildung 2: Trennung mittels Hyperebene durch Transformation in einen höherdimensionalen Raum (vgl. Schölkopf& Smola, 2002, S.15)

2.4.3 Kernfunktionen

Da die Berechnung der Entscheidungsfunktion in dem höherdimensionalen Raum sehr aufwendig ist, kann anhand spezieller Funktionen die Skalarprodukt-Berechnung der transformierten Vektoren im Input- Raum berechnet werden (vgl. Abe, 2005, S.35f, Schölkopf& Smola, 2002, S.25ff). Bei diesen Funktionen handelt es sich um sogenannte Kernfunktionen und das Verfahren wird auch als ‚Kern- Trick' bezeichnet (vgl. Abe, 2005, S.35f, Schölkopf& Smola, 2002, S.25ff). Kernfunktionen, die bei der Anwendung von SVM genutzt werden, sind u.a. die linearen, die polynomialen und die radialen Basisfunktions- Kernfunktionen, die im Folgenden näher beschrieben und für die Untersuchungsdurchführung genutzt werden (vgl. Abe, 2005, S.27, Schölkopf& Smola, 2002, S.45f).

Lineare Kernfunktion

Sind die Daten im ursprünglichen Input- Raum linear trennbar, müssen diese nicht in einen höherdimensionalen Raum abgebildet werden und die Klassifikation wird mit linearen Kernfunktionen durchgeführt:

$$H(x, x') = x^T x'$$

(vgl. Abe, 2005, S.27f, Schölkopf& Smola, 2002, S.45f).

Radiale Basisfunktions- Kernfunktion

Die radiale Basisfunktions- Kernfunktion mit gamma γ als positiver Parameter, der den Radius beeinflusst, ist gegeben durch:

$$H(x, x') = \exp\left(-\gamma |x - x'|^2\right)$$

(vgl. Abe, 2005, S.27f, Schölkopf& Smola, 2002, S.45f).

Polynomiale Kernfunktion

Die polynomiale Kernfunktion mit dregree *d* ($d \in \mathbb{N}$) und dem Koeffizienten (Kernparameter) *coef* ist gegeben durch:

$$\boxed{H(x, x') = (\gamma \cdot x^T x' + coef)^d}$$

(vgl. Abe, 2005, S.27f, Schölkopf& Smola, 2002, S.45f).

Sowohl die passende Kernfunktion, als auch die entsprechenden Kern-Parameter müssen für jeden Datensatz individuell ermittelt werden.

2.4.4 Multiklassifikation mit Support Vektor Maschinen

Bisher wurden nur binäre Klassifikationsprobleme und damit die Zuordnung in eine von zwei Klassen betrachtet. In den meisten Fällen jedoch soll die Klassifikation in mehrere Gruppen erfolgen. Entsprechend sollen in der vorliegenden Arbeit die Winkeldaten den Probanden zugeordnet werden. Es handelt sich also ebenfalls um ein Mehr-Klassen-Problem. Die Lösung solcher Mehr-Klassen-Probleme wird anhand der Multiclass Support Vektor Maschinen gelöst (vgl. Abe, 2005, S.83, Schölkopf& Smola, 2002, S.211-214). Prinzipiell lassen sich dabei vier Arten an Multiclass SVM unterscheiden:

1. One-gainst-all Support Vektor Maschinen
2. Pairwise Support Vektor Maschinen (One-against-one SVM)
3. Error-vorrecting output code Support Vektor Maschinen
4. All-at-once Support Vektor Machinen

(vgl. Abe, 2005, S.83, Schölkopf& Smola, 2002, S.211-214)

Da die libsvm toolbox die ‚one-against-one" Multiklassifikationsmethode anwendet, wird an dieser Stelle nur auf die Pairwise Support Vektor Maschinen eingegangen (vgl. http://www.csie.ntu.edu.tw/~cjlin/libsvm/faq.html#/Q3:_Data_preparation).

One-against-one

Die Pairwise Support Vektor Maschinen bzw. das one-against-one Verfahren basiert auf der Biklassifikation (vgl. Abe, 2005, S.96, Schölkopf& Smola, 2002, S.212). Bei dieser Methode wird paarweise die Entscheidungsfunktion bestimmt, die jeweils zwei Klassen unabhängig von den anderen Klassen voneinander trennt (vgl. Abe, 2005,

S.96, Schölkopf& Smola, 2002, S.212). Für jede mögliche Klassen-Paarung wird folglich ein Klassifikator trainiert (vgl. Schölkopf& Smola, 2002, S.212). Bei N Klassen sind dies $\frac{N(N-1)}{2}$ Klassifikatoren (vgl. Abe, 2005, S.96, Schölkopf& Smola, 2002, S.212). Bei zehn Probanden bzw. zehn Klassen wären dies 45 binäre Klassifikatoren. Für jede Trainingssession werden nur die Trainingsdaten der jeweiligen zwei Klassen verwendet (vgl. Abe, 2005, S.96).

3 Methodik

In diesem Kapitel wird die Vorgehensweise der Untersuchung geschildert. Zum Verständnis und zur Vervollständigung werden dazu zunächst die verwendeten Messinstrumente beschrieben, um danach die Durchführungsweise darzulegen.

3.1 Untersuchungsinstrumente

3.1.1 Qualisys Track Manager – QTM

Zur Erfassung des Gangmusters der Probanden wurde mit dem Programm Qualisys Track Manager (QTM) gearbeitet. QTM ist eine motion capture software die alle Arten von Bewegungen erkennt und darstellt (s. Abb.3).

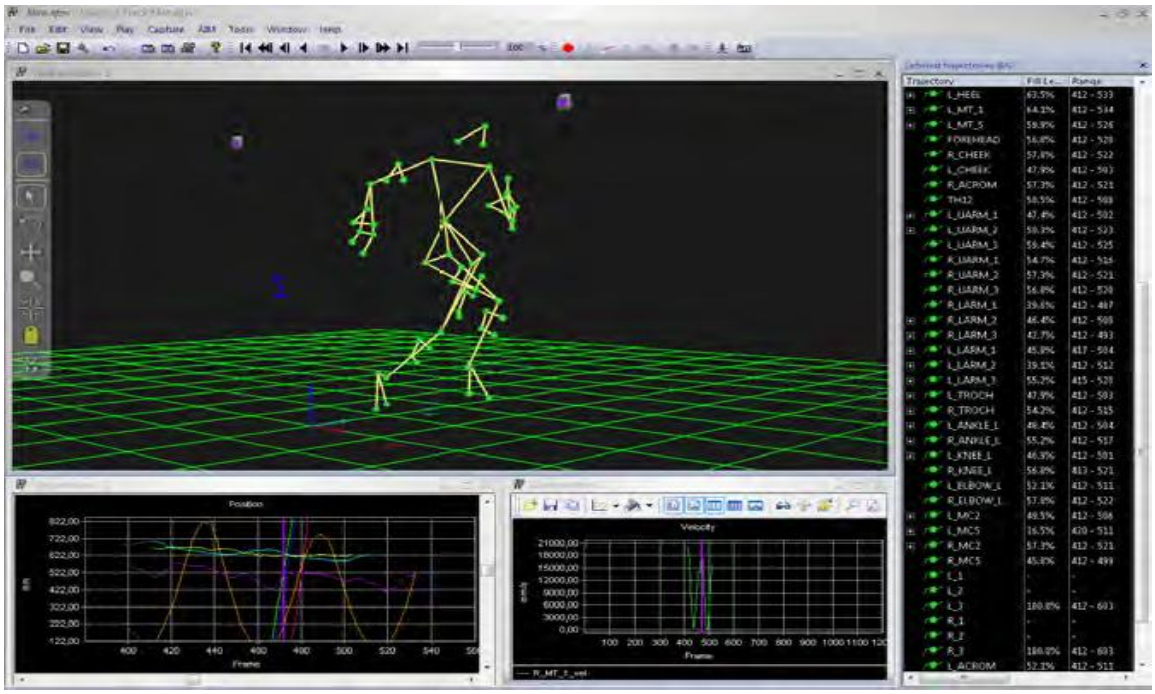


Abbildung 3: Markererkennung mit Qualisys

Kamerasystem

Das Infrarot-Kamerasystem des Herstellers QUALISYS kann nicht nur Infrarot-Licht aussenden und aufnehmen, sondern auch rechnerisch verarbeiten. Jede Kamera misst den Ort eines infrarotreflektierenden Objektes und berechnet die 2-dimensionalen Projektionen der Koordinaten des Punktes. Werden mindestens zwei Kameras kombiniert, werden 3D-Informationen des gemessenen Punktes automatisch erzeugt. Die Kameras besitzen einen Prozessor, der die gemessenen Informationen für jede Kamera einzeln in 2D-Koordinaten des kameraeigenen Koordinatensystems umgewandelt. Die Daten werden über das Netzwerk zum Rechner der Messung übertragen und in 3D-Daten übersetzt. Qualisys übermittelt die Daten in Echtzeit.

Marker

Zum Qualisys-System gehören ebenso Marker, welche das von der Kamera entsandte Licht mit hohem Reflexionsgrad reflektieren. Diese Marker sind kugelförmig, da jede Kamera im Raum sie erkennen muss. Nur an einer Stelle befindet sich eine kleine flache Stelle, die es möglich macht die Marker am Untersuchungsobjekt zu fixieren. Aufgrund Ihrer speziell angefertigten Farbe können die Marker das Infrarotlicht reflektieren. Die Marker haben einen Durchmesser von 7mm. Jeder Marker wird einem bestimmten an einer bestimmten Körperstelle platziert (s. Abb.4).

Abkürzung	***Bezeichnung***	***Abkürzung***	***Bezeichnung***
RH	*Right Heel*	*RKL*	*Right Knee Lateral*
LH	*Left Heel*	*RKM*	*Right Knee Medial*
RFL	*Right Foot Lateral*	*LKL*	*Left Knee Lateral*
RFM	*Right Foot Medial*	*LKM*	*Left Knee Medial*
LFL	*Left Foot Lateral*	*RT1*	*Right Thigh 1*
LFM	*Left Foot Medial*	*RT2*	*Right Thigh 2*
RAL	*Right Ankle Lateral*	*RT3*	*Right Thigh 3*
RAM	*Right Ankle Medial*	*RT4*	*Right Thigh 4*
LAL	*Left Ankle Lateral*	*RT5*	*Right Thigh 5*
LAM	*Left Ankle Medial*	*LT1*	*Left Thigh 1*
RS1	*Right Shank 1*	*LT2*	*Left Thigh 2*
RS2	*Right Shank 2*	*LT3*	*Left Thigh 3*
RS3	*Right Shank 3*	*LT4*	*Left Thigh 4*
RS4	*Right Shank 4*	*LT5*	*Left Thigh 5*
RS5	*Right Shank 5*	*RSAC*	*Right Sacrum*
LS1	*Left Shank 1*	*LSAC*	*Left Sacrum*
LS2	*Left Shank 2*	*RIL*	*Right Iliacum*
LS3	*Left Shank 3*	*LIL*	*Left Iliacum*
LS4		*Left Shank 4*	
LS5		*Left Shank 5*	

Abbildung 4: Bezeichnung der einzelnen Marker

Kalibrierung

Das System benötigt für die Ermittlung der dreidimensionalen Koordinaten erstens die Information, wie die Kameras zueinander aufgestellt sind und zweitens die Lage des Inertialsystems. Dies gelingt durch die Kalibrierung. Für die Kalibrierung sind spezielle Werkzeuge notwendig:

Kalibrier-Gitter

Es handelt sich hierbei um einen rechten Winkel, der mit vier Markern bestückt ist, die in definierten Abständen angebracht sind. Der Scheitelpunkt definiert den Koordinaten-Ursprung des Koordinatensystems.

Kalibrier-Stab

An diesem Stab befinden sich an dessen Ende zwei Marker. Der Kalibrier-Stab wird beim Kalibrieren im Bereich des Kalibrier-Gitters bewegt, um die notwendigen Informationen zur Kalkulation der z-Dimension an das System zu übermitteln.

3.1.2 VISUAL3D TM

Im Anschluss an die Datenerhebung mit Qualisys, werden die Daten mittels des Programms Visual 3D weiter verarbeitet.

Visual 3D ist ein erweitertes Analyseprogramm zur Verwaltung und Dokumentation von optischen 3D-Daten. Dieses mit dem Qualisys Bewegungsanalyse-System kombiniert, bietet eine gute Möglichkeit zur biomechanischen Analyse.

Eingabedaten

Das Standardeingabeformat ist das C3D-Format. Visual 3D liest und schreibt C3D-Dateien, die auf einer herkömmlichen Computerplattform erstellt wurden.

Model-Builder

Der Model-Builder ist ein Modellierungstool, das genutzt werden kann, um eine begrenzte Anzahl an festen Segmenten zu definieren und miteinander zu verbinden. Die Modellierungstechnik ermöglicht die virtuelle Analyse von Menschen, Tieren oder mechanischer Bewegungen (s. Abb. 5).

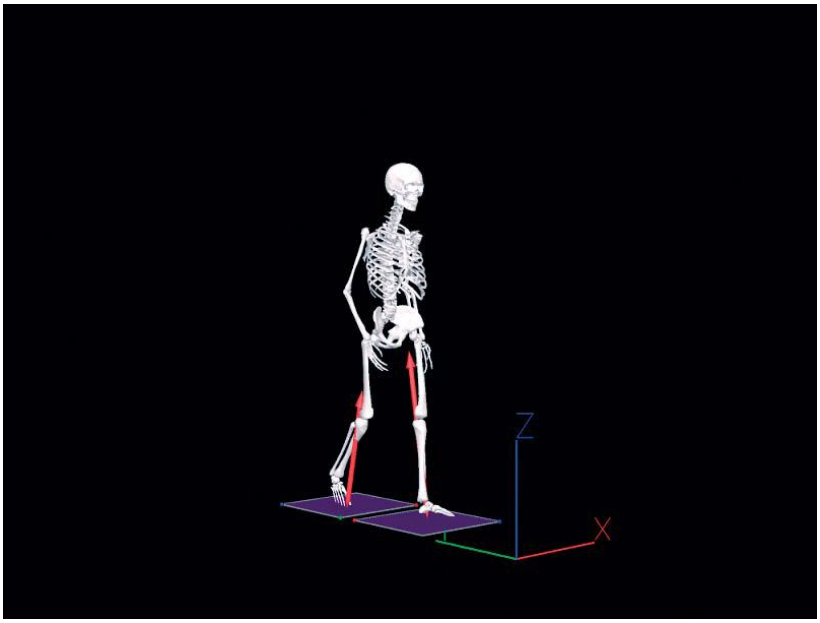


Abbildung 5: Model-builder Visual 3D

Skripte und Ablauf

Visual3D hat eine einfache Skriptsprache, welche die automatische Verarbeitung der Daten ermöglicht. Durch die Benutzung von Skripten kann eine benutzerdefinierte Datenverarbeitung für ausgewählte Anwender eingerichtet werden.

Visualisierung

Die Visualisierung von Bewegungsdaten, die von externer Software importiert wurden, können an Skeletten, Modellen oder Animationen dargestellt werden. Es können EMG-Signale und Bodenreaktionskräfte von Kraftmessplatten visualisiert werden.

3.2 Untersuchungsdurchführung

Die Untersuchung fand im Labor des Instituts für Sportwissenschaft in der Abteilung Trainings- und Bewegungswissenschaft statt. Da die Untersuchung an mehreren Tagen durchgeführt wurde, musste darauf geachtet werden, dass die Gegebenheiten im Labor an jedem Tag identisch sind. An der Untersuchung haben zehn Personen (fünf weiblich und fünf männlich) im Alter von 22 bis 29 Jahren teilgenommen. Alle Personen waren zum Zeitpunkt der Messungen Sportstudenten der Universität Mainz. Die Untersuchungsdurchführung erstreckt sich über drei meist aufeinanderfolgenden Tagen. Die Aufnahmen werden in Gruppen von zwei bis drei Personen durchgeführt, das heißt pro Termin müssen die Datensätze von

mindestens zwei Probanden aufgenommen werden. Nachdem der Raum kalibriert wurde und die Kameras eingestellt sind, wird der erste Proband mit Markern (siehe Kap. 3.1.1) beklebt.

Der Ablauf des Hauptteils der Untersuchung gliedert sich wie folgt:

Jeder Proband startet mit einer statischen Aufnahme und hat anschließend 15 Laufdurchgänge, die sich in jeweils drei mal fünf Durchgänge mit unterschiedlichen Konditionen aufteilen:

- 5 x Laufen normal
- 5 min Eingewöhnung an die Einlagen
- 5 x Laufen mit Einlagen
- 5 x Laufen ohne Einlagen

Da die Untersuchung an drei Tagen durchgeführt wird, hat jeder Proband am Ende 45 Läufe absolviert.

Die Probanden suchen sich einen Startpunkt x, an welchem sie jeden Durchgang beginnen. Dabei muss gewährleistet sein, dass der linke Fuß jeweils auf der zweiten Kraftmessplatte aufgesetzt wird. Der Blick ist dabei geradeaus gerichtet und der Gang sollte möglichst dem natürlichen Gehverhalten der Probanden entsprechen. Der Versuchsleiter, in diesem Fall der zweite Proband, gibt das Kommando für den Start der Durchgänge. Zwischen jedem Durchlauf liegen ca. 20 Sekunden Pause, die der Proband nutzt, um wieder auf seine Startposition zu gehen. Nach fünf Durchgängen ohne Einlagen, hat der Proband fünf Minuten Zeit, sich an die Einlagen zu gewöhnen. Anschließend werden weitere fünf Laufdurchgänge vollzogen, bei denen der Proband Einlagen trägt. Zum Schluss absolviert der Proband weitere fünf Läufe ohne Einlagen.

3.2.1 Datenverarbeitung

Die in der Untersuchung verwendete Videoanalyse dient der Bestimmung der Körperpunkte, die durch die Marker gekennzeichnet wurden. Nach der Erfassung der Marker durch das Programm Qualisys Track Manager werden diese in einem System bei allen Probanden gleich zugeordnet. Man spricht hier vom sogenannten *labeln*. Die fertig gelabelten Daten, werden im Anschluss in das Programm Visual3D geladen. Visual3D von der Firma C-Motion erstellt ein virtuelles Knochenmodell der

Probanden und kann die Winkelverläufe der Gelenke berechnet. Auch die mit der Kraftmessplatte erfassten Kraftverläufe werden mit Hilfe des Programms ermittelt. Die vorliegende Arbeit bezieht sich jedoch nur auf die Winkelverläufe der Probanden. Zur Weiterverarbeitung der Daten wird nun die entsprechende Software zur Klassifizierung mit SVM benötigt.

Softwaretools zu Support Vektor Maschinen sind reichlich als Freeware erhältlich. Für die vorliegende wissenschaftliche Arbeit wurde die freie Software Scilab v genutzt. Über die Funktion Modul Manager ATOMS müssen zur Normierung der Visual 3D Daten die csv_readwrite toolbox und zur Klassifizierung mittels Support Vektor Maschinen die libsvm installiert sein. Über den integrierten Editor SciNotes werden dann zunächst die beiden bereitgestellten Programm- Dateien V3D_readdate_wholefolder_1_1.sce und V3D_normieren_minmax01_1_1.sce zur Normierung der Visual 3D Daten ausgeführt. Zur Ausführung der ebenfalls bereitgestellten Programm- Datei V3D_main_1_1.sce ist die Anzahl der Durchgänge, die Anzahl der Parameter, die Anzahl der Normalisierungswerte sowie die Art der Werte anzugeben. Bei der Anzahl der Durchgänge wird die Anzahl der Durchgänge gesucht, die pro Person ausgewählt werden. Im vorliegenden Fall sind dies 15 Durchgänge. Sind einzelne Durchgänge bei einer Person fehlerhaft und können deshalb nicht verwendet werden, kann die Anzahl der Durchgänge jeweils beim Einlesen der Daten korrigiert werden. Die Anzahl der Gelenkwinkel, die in den Winkeldaten vorliegen, bestimmt die Anzahl der Parameter. Bei den vorhandenen Aufnahmen sind dies 30 Parameter. Beteiligte Gelenke sind Sprunggelenk, Knie, Hüfte, Ellbogen und Schulter, sowohl rechts als auch links in drei Ebenen. Die Anzahl der Normalisierungswerte beträgt in diesem Fall 101. Um die Winkelwerte der Visual 3D Daten auszuwerten, wird die Art der Werte auf Winkel festgelegt. Danach können die gewünschten Datensätze eingelesen werden. Für jeden Datensatz muss dabei der Name und das Körpergewicht des Probanden angegeben werden. Die Angabe des Körpergewichtes dient der Kraftnormalisierung. Von den vorhandenen Datensätzen waren jedoch einige fehlerhaft, sodass diese aussortiert werden mussten. Häufigste Fehlerquelle war dabei das Fehlen von Gelenkwinkeln und damit die Fehlerhafte Anzahl der Parameter. Die aussortierten Datensätze wurden gelöscht, sodass sich die Anzahl der Durchgänge bei einzelnen Personen verringerte. Bei zwei der zehn Probanden konnten keinerlei Datensätze verwendet werden. Von 450 durchgeführten Durchgängen, konnten nur 268 eingelesen werden.

Die Anzahl der verwertbaren Datensätze pro Person liegt zwischen 15 und 43 Datensätzen. Die vollständig eingelesenen Datensätze können dann gemeinsam ausgewertet und gespeichert werden. Für die Anwendung der libsvm toolbox wird später die Datei mit der Endung ‚.dat' benötigt.

4 Klassifikation mittels Support Vektor Maschine

Bei der Verwendung der libsvm toolbox soll nun das Programm mit den zuvor normierten Datensätzen, den Trainingsdaten, trainiert werden. Hierzu muss zunächst die abgespeicherte ‚.dat'- Datei eingelesen werden (s. Anhang1). Die Datensätze der Datei können dann in Klassen eingeteilt werden. In dem vorliegenden Fall bekommt jede Person eine Kategorie zugeordnet. Dies bedeutet, dass die gültigen Datensätze der drei Versuchstage für jeden Probanden zusammengezählt werden. Das Programm soll demnach die ersten 27 Datensätze zu Kategorie ‚1', die folgenden 41 Datensätze zu Kategorie ‚2', usw., zuordnen. Ausgeführt wird diese Zuordnung („target") durch den String *//target= [ones(a,1) ;2*ones(b,1); 3*ones(c,1); 4*ones(d,1); 5*ones(e,1); 6*ones(f,1); 7*ones(g,1); 8*ones(h,1)];*, sodass acht Klassen entstehen (vgl. Anhang 2). Die Variablen a bis h müssen definiert werden und beziehen sich auf die Anzahl der Datensätze pro Person.

4.1 Training mit allen Datensätzen

Mit *model= libsvm_svmtrain(target,daten);* kann nun ein Modell anhand aller vorhandenen Datensätze trainiert werden. Bei der Erstellung des Modells handelt es sich um das Aufstellen der Entscheindungsfunktion, die eine bestmögliche Trennung der Trainingsdaten erlaubt. Um diese Funktion aufzustellen, muss zunächst die Art der Funktion bzw. die Kernfunktion festgelegt werden. Der Parameter *t* bestimmt dabei den Typ der Kernfunktion. In der libsvm toolbox wird mit der Voreinstellung *-t 2* eine radiale Basisfunktion gewählt: *exp(-gamma*|u-v|^2)*. G*amma* legt dabei den exponentiellen Verlauf fest, liegt zwischen Null und Eins und besitzt die Voreinstellung *1*. Werden diese *svmoptions* nicht anders definiert, dienen diese Voreinstellungen als Grundlage für das Modell.

Unter Verwendung des Modells (*model*), der Zuordnung *(target*) und der Daten (*daten*) kann nun mittels *libsvm_predict* Funktion eine Vohersage getroffen werden.

Diese bezieht sich auf die vorhergesagten Kategorien (*pred*) –beschrieben durch einen Vektor mit den vorhergesagten Kategorien-, die Entscheidungswerte (*dec*) – dargestellt als Matrix mit den Entscheidungswerten- und die Genauigkeit der Vorherage (*acc*). *Acc* wird angegeben als dreikomponentiger Vektor bestehend aus der Genauigkeit in Prozent, dem mittleren quadratischen Fehler und dem quadrierten Korrelationskoeffizienten. Für diese Untersuchung reicht allerdings die Betrachtung der Genauigkeit der Vorhersage in Prozent aus, die aus dem jeweiligen Modell resultiert.

Der String *[pred,acc,dec]=libsvm_svmpredict(target,daten,model)* führt die gesamte Vorhersage aus, *disp(acc)* zeigt die Genauigkeit an (s. Anhang 3). Der Screenshot zeigt, dass unter Verwendung aller Datensätze als Trainingsdaten bei Einsatz einer radialen Basisfunktion mit *gamma=1* die Genauigkeit der Vorhersage bei 100% liegt. Auch bei Änderung des exponentiellen Verlaufs durch die Modifikation von *gamma* ändert sich die Vorhersagekraft durch eine radiale Basisfunktion nicht.

Eine weitere Möglichkeit zur Modellerstellung ist die Verwendung einer polynomialen Kernfunktion. Dies muss in den *svmoptions* mit *-t 1* definiert werden. Die ponynomiale Funktion *(gamma*u'*v + coef0)^degree* besitzt, wenn nicht anders definiert, die Voreinstellung *1* für *gamma*, *0* für *coef0* und *3* für *degree*. Werden wieder alle Datensätze als Trainingsdaten und eine polynomiale Kernfunktion genutzt, ergibt sich ebenfalls eine Genauigkeit bei der Vorhersage von 100%. Auch in diesem Fall stellen sich keine Veränderungen der Genauigkeit ein, wenn der Grad des Polynoms durch *degree* oder der exponentielle Verlauf durch *gamma* geändert wird.

Als dritte Möglichkeit soll nun eine lineare Kernfunktion zur Generierung des Modells genutzt werden. Das Skalarprodukt *u‘*v* wird durch *-t 0* angewendet. Bei Verwendung aller Datensätze als Trainingsdaten zeigt sich mittels linearen Kernel erneut eine 100%ige Genauigkeit der Vorhersage.

Folglich kann festgehalten werden, dass die vorliegenden Datensätze sowohl durch lineare als auch durch radiale Basisfunktions- und polynomiale Kernfunktionen klassifiziert werden können, wenn als Trainingsdaten alle Daten verwendet werden. Die Datensätze sind also linear und nicht-linear mittels Support Vektor Maschinen trennbar.

4.2 Training mit einem Teil der Datensätze

Zur Überprüfung der Prognosequalität werden die vorhandenen Datensätze nun in Trainingsdaten und Testdaten unterteilt. Zunächst werden zwei Drittel der Datensätze als Trainingsdaten genutzt (s. Abb. 6). Bei jeder neuen Ausführung

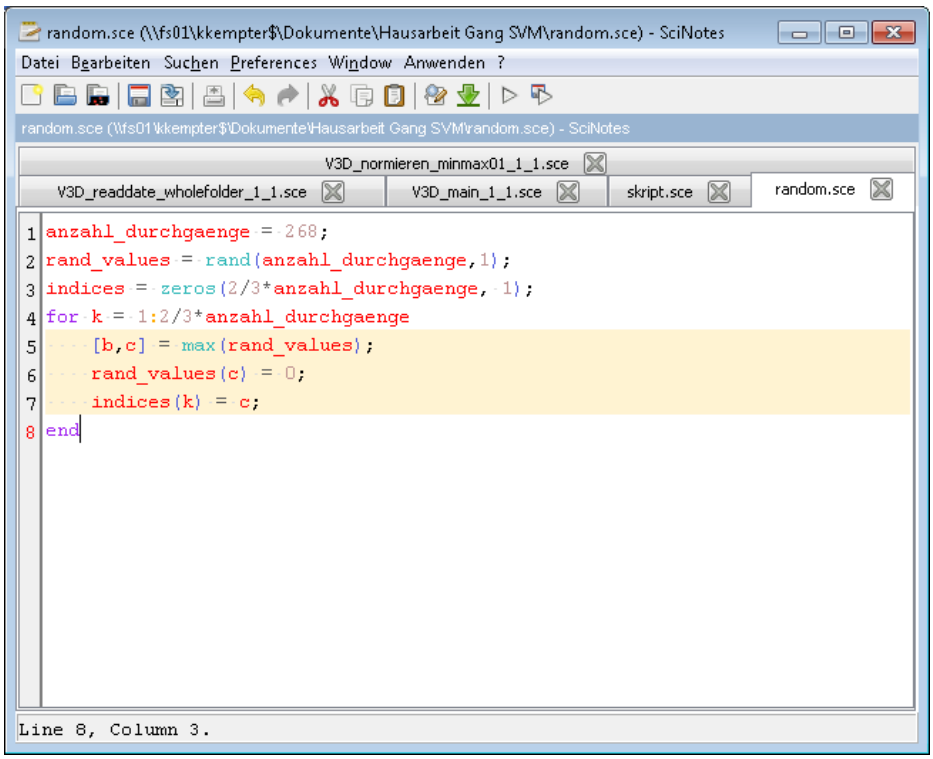


Abbildung 6: Festlegen der Datenmenge für die Erstellung des Modells

variieren die Datensätze. Das Modell wird nun mit den jeweiligen Datensätzen trainiert. Die Vorhersage wird schließlich sowohl mit den Trainings- als auch mit den Testdaten anhand des generierten Modells erstellt. Dabei zeigt sich auch bei der Verwendung einer geringeren Anzahl an Trainingsdaten und der Überprüfung mittels Trainings- und Testdaten bei dem Modell mit linearer Kernfunktion eine nahezu fehlerlose Klassifizierung. In den meisten Ausführungen liegt die Genauigkeit der Vorhersage bei 100%, in seltenen Fällen besteht eine 99,62%ige Präzision. Zu dem gleichen Ergebnis kommen die Modelle mit polynomialer Kernfunktion zweiten oder dritten Grades. Größere Fehlerraten ergeben sich hingegen bei Modellen mit radialen Basisfunktions- Kernfunktionen. In diesem Fall scheint der exponentielle Verlauf eine tragende Rolle zu spielen. Dies äußert sich folgendermaßen: nähert sich *gamma* dem Wert *1* an, wird die Klassifizierung unpräziser. Dahingegen kann die Genauigkeit der Vorhersage bei einem Wert von *gamma=0,1* bei 100% liegen.

Wird die Anzahl der Trainingsdaten weiter verringert und damit gleichzeitig die Anzahl der Testdaten erhöht, nimmt dadurch auch die Präzision der Klassifizierung

ab. Diese Abnahme zeigt sich gleichermaßen bei allen drei Kernfunktionen. Je weniger Trainingsdaten dem Modell zur Verfügung stehen, desto höher liegt die Fehlerquote.

4.3 Diskussion der Ergebnisse

Die Anwendung der Support Vektor Maschine zur Mustererkennung bei der Ganganalyse lieferte bei den vorhandenen Datensätzen eine hohe Präzision der Vorhersage. Die Zuordnung zu den Probanden (Labels) erfolgte dabei durch die Trainingsinputs der Winkeldaten. Hierzu wurden alle Datensätze jeder Person zusammengenommen. Es wurde keine Unterscheidung aufgrund der Tage oder dem Tragen der Einlagen gemacht. Der Einsatz der Einlagen war für diese Arbeit unerheblich. Bei Einsatz aller Datensätze als Trainingsdaten zeigte sich, dass die Daten mit den drei Kernfunktionen linear, radiale Basisfunktion und polynomial prinzipiell trennbar sind. Dies konnte durch die 100%ige Vorhersagekraft bei Verwendung der drei Kernfunktionen gezeigt werden. Interessant für empirische Zwecke ist jedoch, die Generalisierbarkeit der Methode. Die zentrale Frage ist demnach, inwieweit das trainierte Modell (die Entscheidungsfunktion) zuverlässig neue Datensätze mit unbekanntem Label klassifizieren kann. Zu diesem Zweck wurden zur Erstellung des Modells nur noch zwei Drittel der gesamten Datensätze als Trainingsdaten eingesetzt. Dies ermöglicht es, die Generalisierbarkeit mittels der nicht verwendeten Testdaten zu überprüfen. Die Untersuchung zeigte, dass bei Verwendung von 179 Trainingsdaten die gesamten 268 Datensätze –Trainingsdaten plus Testdaten- sowohl mittels linearer Kernfunktion als auch mittels polynomialer Kernfunktion zweiten oder dritten Grades sehr präzise klassifiziert werden konnten. Die Verwendung einer radialen Basisfunktions-Kernfunktion konnte die gesamten Datensätze hingegen weniger genau vorhersagen. Die Präzisionsquote lag so beispielsweise bei rund 90%. Da das Training mit den gesamten Daten gezeigt hat, dass eine 100%ige Vorhersage der Trainingsdaten möglich ist, bedeutet dies, dass die Vorhersagekraft bei den Testdaten nur noch bei rund 70% liegt. Das Modell, das auf einer radialen Basisfunktion basiert, kann Daten mit unbekanntem Label demnach nur zu 70% richtig zuordnen. Die radiale Kernfunktion ist folglich nicht die optimale Kernfunktion zur Lösung unseres Klassifikationsproblems.

Wie bereits beschrieben, nimmt die Vorhersagekraft der Modelle mit abnehmender Anzahl an Trainingsdaten ab. Bei Verwendung von 134 Trainingsdaten besteht bei

der Modellerstellung mittels linearer Kernfunktion und polynomialer Kernfunktion beispielsweise immer noch eine minimal 96,6%ige Prognosefähigkeit aller 268 Daten. Dies bedeutet wiederum eine richtige Klassifizierung zu 92%. Je weniger Trainingsdaten eingesetzt werden, desto größer wird die Fehlerquote. Dies kann zum einen bedeuten, dass die parallele Zunahme der Testdaten auch gleichzeitig eine erhöhte Fehlerquote induziert. Dies lässt auf eine Abschwächung der Generalisierungsfähigkeit der verwendeten Entscheidungsfunktion schließen. Auf der anderen Seite lässt sich aber auch nicht ausschließen, dass die abnehmende Präzision teilweise auf der Verringerung der Trainingsdaten basiert. Bei Verwendung von 134 Trainingsdaten dienen durchschnittlich nur 17 Trainingsdaten pro Person der Zuordnung. Möglicherweise beeinträchtigt die geringere Anzahl an Trainingsdaten die Bestimmung der Entscheidungsfunktion erheblich.

Eine optimale Hyperebene, die die Klassen fehlerlos trennt und eine maximale Generalisierbarkeit aufweist, konnte jedoch nicht gefunden werden. Die Datensätze waren gleichermaßen durch lineare sowie durch polynomiale Kernfunktionen klassifizierbar.

5 Zusammenfassung und Ausblick

In der vorliegenden Arbeit wurden die Möglichkeiten von Support Vektor Maschinen zu Klassifizierung von Gangmustern untersucht. Viele der erhobenen Datensätze mittels Qualisys und Visual 3D erwiesen sich in der Weiterverarbeitung der Daten als fehlerhaft. Datensätze zweier Probanden konnten aus diesem Grund überhaupt nicht verwendet werden, sodass letztendlich die Winkelwerte von acht Probanden klassifiziert werden sollten. Die Klassifizierung mittels libsvm toolbox in Scilab v wurde über die one against one Support Vektor Maschine durchgeführt. Unter Verwendung verschiedener Kernfunktionen wurde zunächst die Trennbarkeit der Klassen überprüft, die eine erfolgreiche Klassifizierung unter Einsatz von Kern- Tricks darlegte. Zur Überprüfung der Generalisierbarkeit wurden die Daten schließlich in Trainings- und Testdaten unterteilt. Dies brachte zum Ausdruck, dass die erstellten Entscheidungsfunktionen nur bedingt generalisierbar sind, wobei die Problematik der geringen Anzahl an Datensätzen zu beachten ist. Dennoch demonstriert die Verwendung der Support Vektor Maschine zur Lösung des vorliegenden Klassifikationsproblems die hohe Generalisierungsfähigkeit dieser Lernmaschine. Der Einsatz der unterschiedlichen Kernfunktionen zeigte in dieser Untersuchung nur geringe Unterschiede in der Genauigkeit. Trotz relativ geringer Anzahl an Trainingsdaten bei acht potentiellen Kategorien konnten sowohl mittels linearem als auch polynomialem Kernel Modelle erzeugt werden, die eine nahezu 100%ige Vorhersage ermöglichten.

Interessant wäre an dieser Stelle die Überprüfung der Generalisierbarkeit anhand weiterer Testdaten. Darüber hinaus wäre eine binäre Klassifikation der Datensätze denkbar, die prüfen soll, ob sich die Daten der Durchgänge mit Einlagen von den Daten der Durchgänge ohne Einlagen mittels Kernfunktionen voneinander trennen lassen. Dies könnte Aufschluss darüber geben, ob sich das Gangmuster in einer bestimmten Weise durch die Einlagen beeinflussen lässt.

6 Literatur

Abe, S. (2005). Support Vector Machines for Pattern Classification. London: Springer-Verlag

Bachmann, C. (2008). Messsysteme, Messmethoden und Beispiele zur instrumentierten Ganganalyse. *Schweizerische Zeitschrift für Sportmedizin und Sporttraumatologie* 56: S.29-34.

Chang, C., Lin, C. (2013). Libsvm FAQ. In *Libsvm –A library for Support Vector Machines*. Zugriff am 12.04.2013 unter http://www.csie.ntu.edu.tw/~cjlin/libsvm/faq.html#/Q3:_Data_preparation

Decker, M. (2007). Entwicklung eines ganzheitlichen Prognosemodells zur Kompensation von Varianzen in Prozessfolgen mittels Support Vektor Maschinen. Dissertation, Universität Stuttgart

Duda, R.O., Hart, P.E., Stork, D.G. (2000). Pattern Classification (2.Auflage). New York: Wiley Interscience

Franz, M.O. (o. A). Mustererkennung und Klassifikation. Zugriff am 19.03.2013 unter www.ios.htwg-konstanz.de/joomla_mof/index.php

Gieseke, F. (2011). From Supervised to Unsupervised Vector Machines and Applications in Astronomy. Dissertation, Universität Oldenburg

Janssen, D., Schöllhorn, W.I., Lubienetzki, J., Fölling, K., Kokenge, H. & Davids, K. (2008). Recognition of Emotions in Gait Patterns by Means of Artificial Neural Nets. *Journal of Nonverbal Behavior, 32, (2).* Zugriff am 16.04.2013 unter http://www.springerlink.com/content/u7x5v402561872u4/fulltext.pdf

Logan, D., Kiemel, T., Dominici, N., Capellini, G., Ivanenko, Y., Lacquanit, F. & Jeka, J. (2010). The many roles of vision during walking. *Experimental Brain Research, 2010, (3).* Zugriff am 14.03.2013 unter http://www.springerlink.com/content/t3318tur67101651/fulltext.pdf

Mikut, R. (2001). Diagnoseunterstützung für die instrumentelle Ganganalyse. Zugriff am 21.03.2013 unter http://bibliothek.fzk.de/zb/berichte/FZKA6613.pdf

Niemann, H. (2003). Klassifikation von Mustern (2. Überarbeitete Auflage).Zugriff am 20.03.2013 unter http://www5.informatik.uni-erlangen.de/fileadmin/Persons/NiemannHeinrich/klassifikation-von-mustern/m00-www.pdf

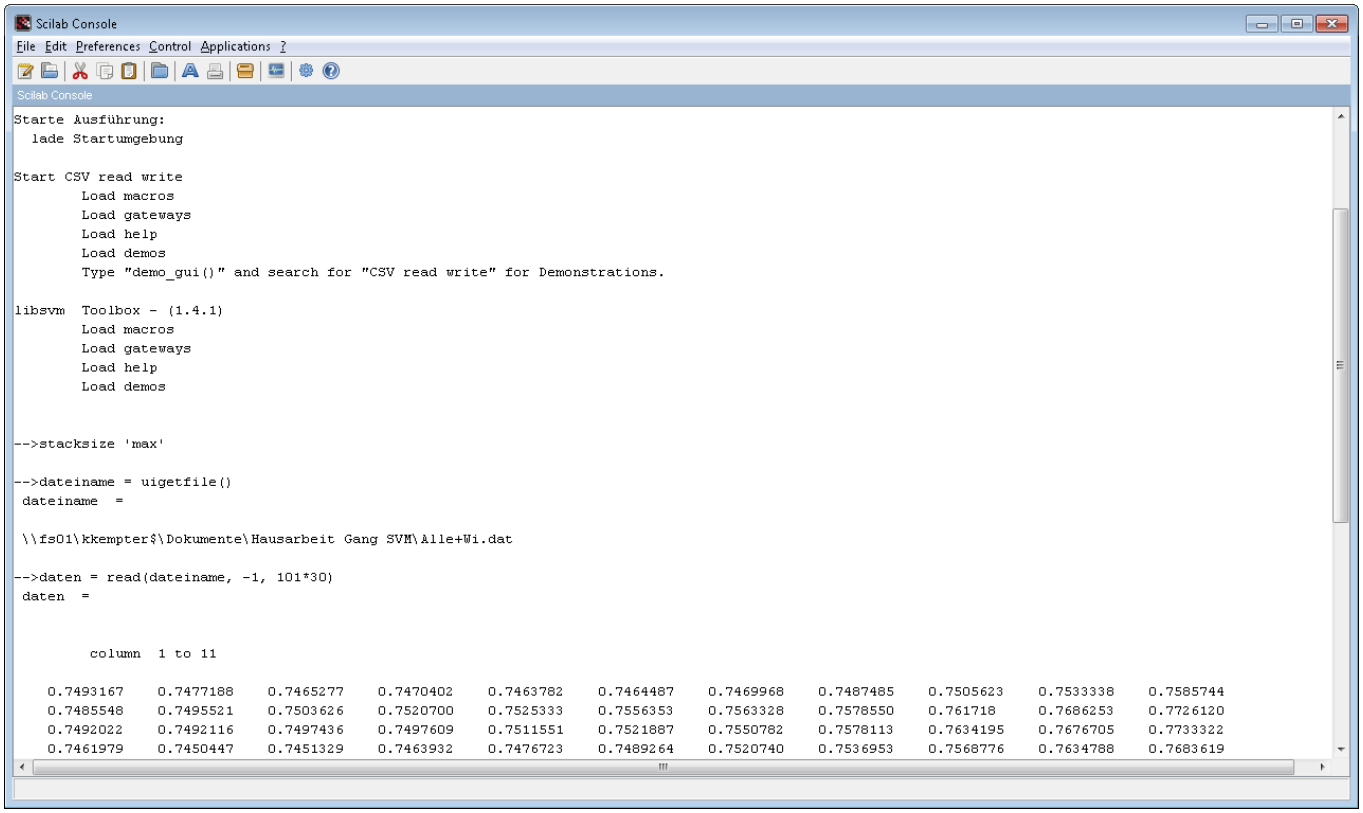
Schölkopf, B., Smola, A.J. (2002). Learning with Kernels. London: MIT Press

Schöllhorn, W.I., Nigg, B.M., Stefanyshyn, D.J. & Liu,W. (2002). Identification of individual walking patterns using time discrete and time contionuous data sets. *Gait and Posture, 15 (2)*. Zugriff am 10.03.2013 unter http://www.ncbi.nlm.nih.gov/pubmed/11869912

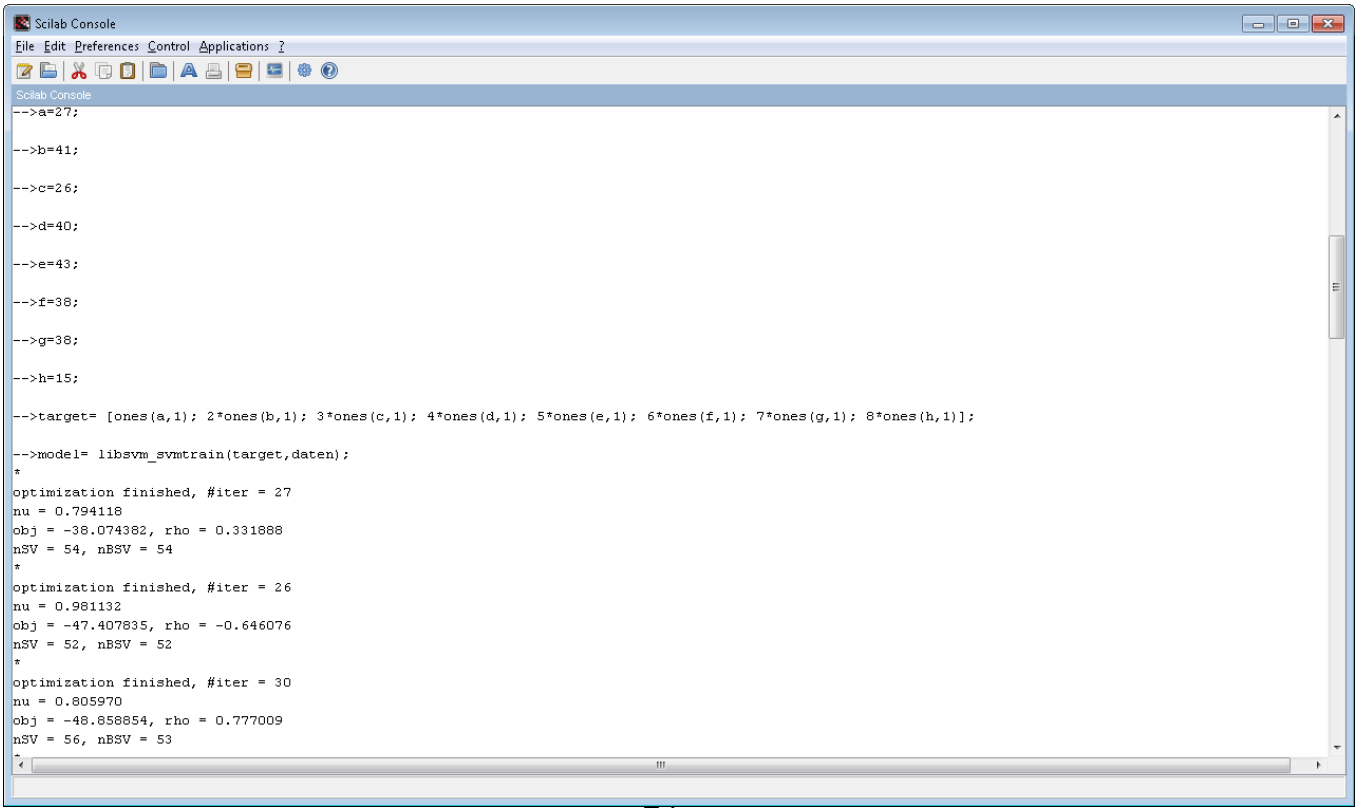
Qualisys, AB. (2013). Gait analysis, an objective method for the analysis of walking patterns. Zugriff am 10.03.2013 unter http://www.qualisys.com/wp-content/uploads/2011/05/AN_Gait.pdf

Anhang

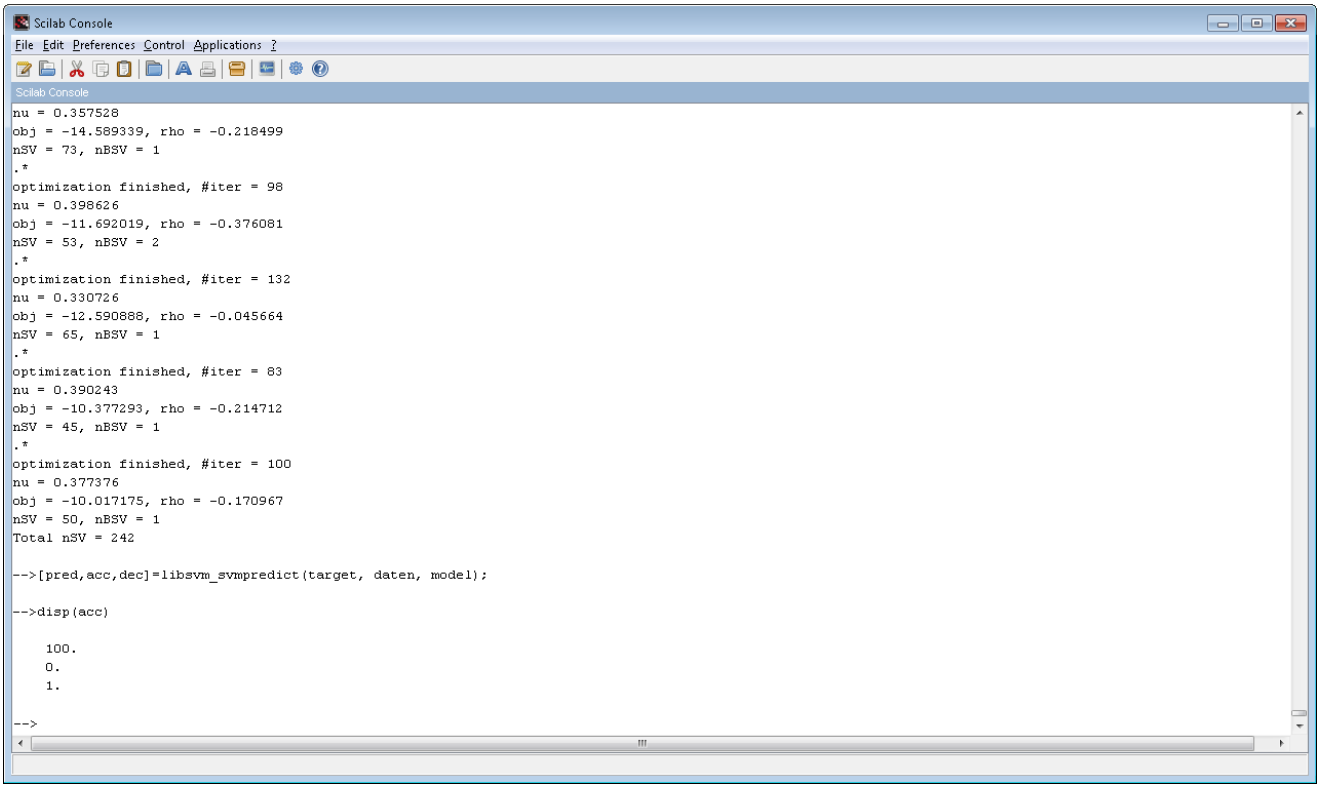
Anhang 1:Laden der Dateien in Scilab



Anhang 2: Zuordnung der Datensätze zu Klassen und Erstellung des Modells



Anhang 3: Datenvorhersage mittels *target*, *daten* und *modell*



Erklärung

Ich erklären wir, dass wir die Arbeit selbständig und nur mit den angegebenen Hilfsmitteln angefertigt haben und dass alle Stellen, die dem Wortlaut oder dem Sinne nach anderen Werken entnommen sind, durch Angabe der Quellen als Entlehnungen kenntlich gemacht worden sind.

(Ort, Datum, Unterschrift) (Ort, Datum, Unterschrift)