
Aus dem Institut für Trainingswissenschaft und Sportinformatik
der Deutschen Sporthochschule Köln
Geschäftsführender Leiter: Prof. Dr. Daniel Memmert

KLASSIFIZIERUNG VON
INTERAKTIONSPROZESSEN IM SPORTSPIEL
MIT HILFE SELBSTORGANISIERENDER
KARTEN

Von der Deutschen Sporthochschule Köln
zur Erlangung des akademischen Grades

Doktor der Naturwissenschaft

angenommene Dissertation

vorgelegt von

Andreas Grunz

aus

Wiesbaden

Köln 2021

Erster Gutachter:	Univ.-Prof. Dr. Daniel Memmert
Zweiter Gutachter:	Univ.-Prof. Dr. Arnold Baca
Vorsitzender des Promotionsausschusses:	Univ.-Prof. Dr. Mario Thevis
Datum der Disputation:	26.01.2021

Eidesstattliche Versicherung gemäß § 7, Absatz 2, Nummer 4 und 5 der Promotionsordnung der Deutschen Sporthochschule Köln vom 20. Februar 2013:

Hierdurch versichere ich:

Ich habe diese Arbeit selbstständig und nur unter Benutzung der angegebenen Quellen und technischen Hilfen angefertigt; sie hat noch keiner anderen Stelle zur Prüfung vorgelegen. Wörtlich übernommene Textstellen, auch Einzelsätze oder Teile davon, sind als Zitate kenntlich gemacht worden.

Hierdurch erkläre ich, dass ich die „Leitlinien guter wissenschaftlicher Praxis“ der Deutschen Sporthochschule Köln eingehalten habe.

Datum, Unterschrift

Danksagung

An erster Stelle möchte ich mich bei meinen beiden Betreuern Prof. Dr. Daniel Memmert und Prof. Dr. Jürgen Perl bedanken. Durch ihre Expertise, Unterstützung, Motivation und das entgegengebrachte Vertrauen wurde diese Arbeit erst möglich. Die Zeit in der gemeinsamen sportinformatischen Forschungsgruppe wird mir immer als eine schöne Lebensphase in Erinnerung bleiben.

Des Weiteren gilt mein Dank Prof. Dr. Arnold Baca, der sich kurzfristig bereit erklärt hat das Zweitgutachten zu übernehmen.

Den Mitarbeitenden am Institut für Trainingswissenschaft und Sportinformatik der Deutschen Sporthochschule Köln und den Mitarbeitenden am Institut für Informatik der Johannes Gutenberg-Universität Mainz möchte ich für die hilfreichen Diskussionen und die freundschaftliche Atmosphäre danken. Mein besonderer Dank geht an Dr. Philip Furley, Dr. Marco Rathschlag und Dr. Sebastian Schwab.

Meinen Eltern danke ich, dass sie mir diesen Weg ermöglicht haben. Ich danke meiner Frau Melanie für ihre Liebe, Unterstützung und das Korrekturlesen.

Inhaltsverzeichnis

Inhaltsverzeichnis	ii
Abbildungsverzeichnis	ii
Tabellenverzeichnis	iii
1 Einleitung	1
1.1 Motivation	1
1.2 Zielsetzung	3
1.3 Aufbau der Arbeit	3
2 Grundlagen	6
2.1 Forschungsstand	6
2.1.1 Sportwissenschaft	6
2.1.2 Informatik	9
2.1.3 Sportinformatik	17
2.2 Selbstorganisierende Karten	18
2.2.1 <i>Self-Organizing Map</i>	18
2.2.2 <i>Dynamically Controlled Network</i>	20
2.2.3 <i>Merge Self-Organizing Map</i>	22
2.3 Zusammenfassung	23
3 Veröffentlichungen	25
3.1 Analysis and Simulation of Actions in Games by Means of Special Self-Organizing Maps	26
3.2 Tactical pattern recognition in soccer games by means of special self-organizing maps	27

3.3	Detecting tactical pattern in basketball: Comparison of merge self-organising maps and dynamic controlled neural networks	29
4	Zusammenfassung	30
5	Abschluss	36
5.1	Bezug zur aktuellen Forschung	36
5.2	Implikationen für die Wissenschaft	41
5.3	Implikationen für die Praxis	42
6	Abstracts	44
6.1	Zusammenfassung	45
6.2	Summary	46
	Literaturverzeichnis	47

Abbildungsverzeichnis

2.1	Intrinsische Dimension, lineare und nicht-lineare Daten . . .	13
2.2	Zwei Wege mit kleinem <i>Hausdorff Abstand</i> d_H (nach Alt und Godau (1995))	16
2.3	Schematische Darstellung einer zweidimensionalen <i>Self-Organizing Map</i>	19
2.4	Schematische Darstellung des <i>PerPot</i> Metamodells	21

Tabellenverzeichnis

4.1	Übersicht der Klassifizierungsergebnisse nach Kategorien	32
-----	--	----

Kapitel 1

Einleitung

1.1 Motivation

Die Spielregeln im Fußball sind leicht nachzuvollziehen und in der Regel in kurzer Zeit zu erlernen. Das sich aus dem Regelwerk ergebende Spiel dagegen zeigt sehr dynamische und komplexe Muster. Interaktionen finden zwischen einzelnen Spielenden, zwischen Gruppen von Spielenden und zwischen den Mannschaften statt. Selbst diese Betrachtungsweise stellt noch eine Vereinfachung dar. Betrachtet man einzelne Spielende zu einem gegebenen Zeitpunkt, so lässt sich nicht mit Bestimmtheit sagen, welche anderen Spielenden die Aktionen in welchem Maße beeinflussen. Zudem findet eine permanente Wechselwirkung zwischen den Spielenden statt. Das Verhalten des gegnerischen Teams wirkt sich auf die Aktionen der Spielenden aus und umgekehrt.

Für den Erfolg ist neben den technischen Fertigkeiten, der physischen Konstitution und psychologischen Faktoren auch taktisches Agieren entscheidend. Die videobasierte Spielbeobachtung durch Trainierende ist vermutlich die nach wie vor am häufigsten praktizierte Form der

Leistungsbewertung. Um die vielfältigen Aspekte von Leistung systematisch zu erfassen, wurden zahlreiche Verfahren entwickelt. Eine Übersicht findet man bei Hughes und Bartlett (2002).

Die digitale Transformation, die so gut wie jeden Lebensbereich unserer Gesellschaft erfasst, hat auch Einzug in den Sport gehalten. Die Leistung von Computern steigt stetig und ermöglicht damit computergestützte Analysen, die auf Grund ihres Rechenaufwandes früher nicht praktikabel gewesen wären. Die Transformation zeigt sich auch dadurch, dass mittlerweile in fast jeder Sportart automatisch Daten erhoben werden. Im Fußball lassen sich aus einem digitalen Video die Positionen der Spielenden und des Balls errechnen. Dieser Wandel ermöglicht die Entwicklung und Berechnung neuer komplexer Leistungsindikatoren basierend auf Positionsdaten. Die meisten traditionell genutzten Indikatoren werden videobasiert erhoben und beruhen auf einfachen statistischen Kennzahlen, z.B. die Zweikampfquote. Diese Kennzahlen aggregieren jedoch den gesamten Spielverlauf und haben wenig Aussagekraft. Durch den Einsatz informatischer Methoden kann dagegen auch der spielerische Kontext berücksichtigt werden.

Leistungsindikatoren ermöglichen die systematische Erfassung von Leistung. Jedoch spiegeln Leistungsindikatoren nicht unbedingt die Bewertung durch Trainierende wieder. Natürlich können Leistungsindikatoren auch aus der Erfahrung von Trainierenden entwickelt werden, aber die Erfassung dieses impliziten Wissens ist schwierig und aufwändig. Gleichwohl haben diese Erfahrungswerte ihre Berechtigung und können die Bewertung durch Leistungsindikatoren unterstützen. In der vorliegenden Arbeit wird ein sportinformatisches Verfahren entwickelt, das dieses implizite Wissen am Beispiel Mannschaftstaktiken aus Beispielen erlernt. Gegenüber einer Fachkraft benötigt das Programm nur Minuten, um gelernte Mannschaftstaktiken in neuen Spielen zu identifizieren. Ein Algorithmus errechnet zudem die Muster immer auf dieselbe Weise und ist damit ob-

jektiver als eine Fachkraft. Die entwickelte Klassifizierung kann auch als Basis für ein neuen Leistungsindikator dienen.

1.2 Zielsetzung

Das Ziel der vorliegenden Arbeit ist die Klassifizierung mannschaftstaktischer Interaktionsprozesse mit Hilfe *Selbstorganisierender Karten*. Als erstes Teilziel wird der komplexe mannschaftstaktische Interaktionsprozess im Sportspiel mit Hilfe *Selbstorganisierender Karten* modelliert. Dabei muss die Dynamik des Interaktionsprozesses erhalten bleiben. Exemplarisch soll am Beispiel Fußball eine Abbildung entwickelt werden, die den komplexen Interaktionsprozess auf charakteristische Prototypen abstrahiert. Als zweites Teilziel soll auf Basis des Modells auch eine automatische Klassifizierung mannschaftstaktischer Interaktionsmuster auf Basis von Positionsdaten im Fußball durchgeführt werden. Da sich die Problemstellung an der Schnittstelle zwischen Sport und Informatik befindet, wird zunächst der Forschungsstand aus sportwissenschaftlicher, informatischer sowie sportinformatischer Sicht dargestellt. Insbesondere wird untersucht, inwiefern sich bestehende Verfahren zur Modellierung der Interaktionsprozesse im Sportspiel oder zur Klassifizierung taktischer Interaktionsmuster eignen. Drittes Teilziel ist die Übertragung des entwickelten Verfahrens auf ein anderes Sportspiel. Zudem wird ein alternatives Verfahren zur Klassifizierung angewendet. Die Ergebnisse der beiden Verfahren werden abschließend verglichen.

1.3 Aufbau der Arbeit

Kapitel 2 stellt zum einen den Forschungsstand dar und zum anderen die Grundlagen, auf denen die vorliegende Arbeit beruht. In Abschnitt 2.1 wird der aktuelle Forschungsstand aus sportwissenschaftlicher, infor-

matischer sowie sportinformatischer Sicht dargestellt. *Selbstorganisierende Karten* (Abschnitt 2.2) wurden in verschiedenen sportinformatischen Untersuchungen erfolgreich eingesetzt (Jäger, Perl & Schöllhorn, 2007; Lames & Perl, 1999; Perl, 2002b; Pfeiffer & Perl, 2006; Schmidt, 2012). Daher werden diese als Basis für das eigene entwickelte Verfahren gewählt und im Folgenden näher betrachtet. In Abschnitt 2.2.1 wird die *Selbstorganisierende Karte* (SOM) nach Kohonen (1982, 1990, 2001, 2013) vorgestellt. Das aus der *Selbstorganisierenden Karte* und dem *PerPot* Modell entwickelte *Dynamically Controlled Network* (DyCoN) nach Perl (2001b) wird in Abschnitt 2.2.2 beschrieben. Als ein Vertreter der rekurrenten *Selbstorganisierenden Karten* wird in Kapitel 2.2.3 die *Merge Self-Organizing Map* (MSOM) nach Strickert und Hammer (2003) betrachtet. Zusammenfassend wird in Abschnitt 2.3 das sich aus dem Forschungsstand und der Zielsetzung ergebende Forschungsdefizit dargestellt, aus dem sich die weitere Arbeit ableitet.

Das darauf beruhende Forschungsvorhaben und die abgeleiteten Teilziele werden in den einzelnen Veröffentlichungen in Kapitel 3 behandelt. In der ersten Veröffentlichung 3.1 wird ein Modell entwickelt, um die mannschaftstaktischen Interaktionsprozesse im Fußball mittels *Selbstorganisierender Karten* abzubilden. In der zweiten Veröffentlichung 3.2 wird dieser Ansatz zu einer Hierarchie Künstlicher Neuronaler Netze weiterentwickelt, um kurze und lange Spieleröffnungen zu klassifizieren. Anhand der Positionsdaten des WM Finales 2006 wird der neue Ansatz evaluiert. In der dritten Veröffentlichung 3.3 wird das hierarchische Verfahren für die Erkennung taktischer Muster im Basketball genutzt. Gleichzeitig wird mit dem MSOM nach Strickert und Hammer (2003) auch ein Algorithmus zur Verarbeitung von Zeitreihen angewandt und dem entwickelten Ansatz gegenübergestellt. Anhand der Positionsdaten eines Basketballspiels und drei taktischen Mustern wird die Klassifizierung mit beiden Verfahren durchgeführt, und die Ergebnisse werden verglichen.

In Kapitel 4 werden die Ergebnisse im Gesamtzusammenhang diskutiert.

Schließlich wird ein Ausblick auf Bereiche zukünftiger Forschung gegeben.

Abschließend werden in Kapitel 5 die Veröffentlichungen aus Kapitel 3 in Bezug zu weiteren Untersuchungen gesetzt, die seitdem publiziert wurden (Abschnitt 5.1). Die Implikationen für Wissenschaft und Praxis werden in Abschnitt 5.2 und 5.3 dargestellt.

Kapitel 2

Grundlagen

Im Hinblick auf die Zielstellung werden zunächst der sportwissenschaftliche [2.1.1](#), informatische [2.1.2](#) sowie sportinformatische [2.1.3](#) Forschungsstand dargelegt. Basierend auf diesen Forschungserkenntnissen werden die *Self-Organizing Map (SOM)* nach Kohonen ([2001](#)) und das daraus entwickelte *Dynamically Controlled Network (DyCoN)* nach Perl ([2001b](#)) als Grundlage für das eigene Verfahren in Abschnitt [2.2.1](#) bzw. [2.2.2](#) vorgestellt. Mit der *Merge Self-Organizing Map (MSOM)* nach Strickert und Hammer ([2003](#)) wird zudem eine rekurrente *Selbstorganisierende Karte* dargestellt (Abschnitt [2.2.3](#)).

2.1 Forschungsstand

2.1.1 Sportwissenschaft

Die videobasierte Spielbeobachtung durch Trainierende ist vermutlich die noch immer am häufigsten genutzte Form der Leistungsbewertung. In ihrer Untersuchung konnten Franks und Miller ([1991](#)) zeigen, dass Trainierende im Fußball im Nachgang maximal 40% der relevanten Infor-

mationen aus einem Video wiedergeben konnten. Gleichzeitig wird eine in der Regel subjektive Bewertung vorgenommen. Durch qualitative und quantitative Methoden kann eine fundierte systematische Spielbeobachtung realisiert werden. Mackenzie und Cushion (2013) sowie Sarmiento et al. (2014) stellen eine Übersicht verschiedener empirischer Studien im Fußball vor. Während früher in der Regel nur ein Video für die Analyse zur Verfügung stand, gibt es heute auch andere Datenquellen. Insbesondere werden in vielen Sportspielen mittlerweile Positionsdaten erhoben.

Der Definition von Hughes und Bartlett (2002) folgend, ist ein Leistungsindikator eine Auswahl oder Kombination verschiedener beobachteter Merkmale zur Bestimmung der gesamten Leistung bzw. verschiedener Aspekte von Leistung. Entsprechend dieser Definition gibt es sehr viele unterschiedliche Arten von Indikatoren, z.B. taktische Indikatoren, technische Indikatoren, physische Indikatoren etc., je nachdem welcher Aspekt der Leistung betrachtet wird. Während erste Indikatoren oft noch durch ein Video erhoben wurden, basieren neuere Leistungsindikatoren auf Positionsdaten.

Die Auswahl und der Einsatz von Leistungsindikatoren erfordern eine gewisse Expertise. Hughes und Bartlett (2002) weisen darauf hin, dass die Interpretation einzelner Leistungsindikatoren zu falschen Schlussfolgerungen führen kann. Betrachtet man z.B. die Anzahl der Tore als einen Indikator, so schneidet Mannschaft A mit vier Toren besser ab als Mannschaft B, die nur zwei Tore erzielt hat. Betrachtet man jedoch zusätzlich die Anzahl der Torschussversuche, so schneidet Mannschaft B mit zwei Spielabschlüssen gegenüber der Mannschaft A mit 10 Spielabschlüssen besser ab. Es wird empfohlen nach Möglichkeit normierte und dimensionslose Indikatoren zu verwenden. Die Untersuchung von Parmar, James, Hearne und Jones (2018) verdeutlicht ein weiteres Problem beim Einsatz mehrerer Leistungsindikatoren. In der Untersuchung werden 45 Indikatoren durch eine *Hauptkomponentenanalyse* (Dimensionsreduktion 2.1.2) auf 10

Komponenten reduziert. Verschiedene Indikatoren können miteinander korrelieren. Werden beispielsweise die Herz- und die Atemfrequenz jeweils als ein Indikator erfasst, dann gehen hohe Herzfrequenzen bei Belastung vermutlich mit einer hohen Atemfrequenz einher. Betrachtet man eine Menge von Indikatoren zur Bewertung von Leistung, kann das einzelne Leistungsaspekte hervorheben. Der Auswahl der richtigen Leistungsindikatoren je nach Fragestellung kommt daher besondere Bedeutung zu. Zudem muss eine Evaluierung der Leistungsindikatoren erfolgen, um ebenso für die Leistung irrelevante Indikatoren auszuschließen. Eine Evaluation taktischer Leistungsindikatoren auf Basis von Positionsdaten findet man in der Untersuchung von Memmert et al. (2016).

Neben einfachen statistischen Leistungsindikatoren, wie beispielsweise der maximalen Sprintgeschwindigkeit oder der Zweikampfquote, werden zunehmend auch multivariate statistische Verfahren, Methoden des maschinellen Lernens oder der künstlichen Intelligenz zur Berechnung von Leistungsindikatoren eingesetzt. Einfache statistische Kennzahlen aggregieren den gesamten Spielverlauf in einer Zahl. Die Situation, in der eine Aktion ausgeführt wurde, geht dabei verloren. Mackenzie und Cushion (2013), Sarmiento et al. (2014) sowie Carling, Wright, Nelson und Bradley (2014) heben in ihren Artikeln die Bedeutung des Kontextes zur Interpretation der Leistungsindikatoren hervor.

Der Begriff Taktik geht zurück ins Altgriechische (τακτική τέχνη) und kann übersetzt werden als die Kunst der Anordnung (des Heeres). Eng verwandt ist der Begriff der Strategie als die Kunst der Planung (der Feldzüge). In ihrer ursprünglichen Verwendung spiegeln diese Begriffe das Verhältnis zwischen Taktik und Strategie wieder. Die Strategie ist der vorher gefasste Plan zur Erreichung der Ziele. Die Taktik beschreibt die folgende Umsetzung unter den gegebenen Umständen. Im Kontext des Sports beschreiben Schnabel, Harre und Krug (2014) die Strategie als vorgedachten und festgelegten Verhaltensplan. Die Taktik wird definiert

als Zusammenfassung aller Verhaltensweisen, Handlungen und Operationen, die unter den gegebenen Umständen die Bedingungen für einen optimalen Spielverlauf beeinflussen.

Bei der Betrachtung von Taktiken im Mannschaftssport wird zwischen Individualtaktiken, Gruppentaktiken und Mannschaftstaktiken unterschieden. Diese Kategorisierung stellt jedoch eine Vereinfachung dar. In einem Spiel gehen diese verschiedenen taktischen Kategorien fließend ineinander über. Die Individualtaktik ist so auch immer Teil einer Gruppentaktik. Die Gruppentaktik wiederum beeinflusst die Individualtaktik. In der vorliegenden Arbeit wird die Klassifizierung von Mannschaftstaktiken erlernt. Insbesondere berücksichtigt das entwickelte Verfahren auch den Kontext.

2.1.2 Informatik

Aus informatischer Sicht steigt, ausgehend von statischen Daten über Zeitreihen hin zu räumlich-zeitlichen Daten, die Komplexität der Daten an. Die Komplexität wird durch die Analyse von Interaktionsmustern weiter gesteigert. Mit zunehmender Komplexität der Daten sind weitergehende Analysen möglich. Statische Daten verändern sich nicht mit der Zeit und können beispielsweise die Merkmale eines Gegenstandes oder einer Person beschreiben. Bei statischen Daten können z.B. die Verteilung der Daten oder der Zusammenhang zwischen den Merkmalen untersucht werden. Bei Zeitreihen gibt es zusätzlich eine Abhängigkeit zu vorhergehenden oder nachfolgenden Datensätzen. Hier kann zusätzlich untersucht werden, inwiefern sich etwa zukünftige Werte aus früheren Werten ergeben. Bei räumlich-zeitlichen Daten können auch räumliche Muster untersucht werden, etwa ob sich bestimmte Objekte oder Personen in ähnlichen Bereichen bewegen. Erfasst man für jedes interagierende Element einer Interaktion die räumlich-zeitlichen Daten, so wird die Interaktion durch die gesamte Menge dieser Daten beschrieben. Diese Menge an Daten gibt dabei nur das Resultat der Interaktion zwischen den Elementen wieder. Während der Ausführung kann es Wechselwirkungen

zwischen einzelnen oder mehreren Elementen geben.

Im Folgenden werden einige grundlegende informatische Begriffe, Konzepte und Verfahren erläutert. Zudem wird das eigene Verfahren in der Forschungslandschaft verortet.

Überwachte und unüberwachte Lernverfahren Betrachtet man die Verfahren zur Datenanalyse aus der Sicht des maschinellen Lernens, lässt sich zwischen überwachten und unüberwachten Lernverfahren unterscheiden. Die Klassifikation zählt zu den überwachten Lernverfahren. Hier liegen bereits kategorisierte Daten vor, d.h. den einzelnen Daten wurde aus einer Menge von Kategorien in der Regel eine Kategorie zugeordnet. Anhand der kategorisierten Daten lernt das Verfahren die Zuordnung zu den Kategorien. Im Anschluss kann das gelernte System dazu genutzt werden, neue Daten zu erkennen und diese einer Kategorie zuzuordnen. Eine Gesamtsicht überwachter Lernverfahren findet sich bei Iqbal und Yan (2015). Beim unüberwachten Lernen hingegen liegen keine kategorisierten Daten vor. Das System erlernt die Muster aus den ursprünglichen Daten. Zielvorgaben können in diesem Fall z.B. die Gruppierung nach Ähnlichkeit sein oder die Identifizierung von Prototypen, die eine Teilmenge der Daten repräsentieren. Wurden entsprechende Gruppen oder Prototypen von Daten bestimmt, so kann eine semantische Kalibrierung erfolgen. Dabei wird mit kategorisierten Daten getestet, welche Prototypen bzw. Gruppen den Daten jeweils am ähnlichsten sind. Die beispielsweise am häufigsten zugewiesene Kategorie wird dann dem Prototypen oder der Gruppe als Kategorie zugewiesen. Im ersten Schritt werden die explorativen Eigenschaften des unüberwachten Lernens genutzt, um geeignete Abstraktionen zu finden. Im zweiten Schritt kann mittels eines überwachten Schrittes eine Klassifizierung erfolgen.

Clusteranalyse Der Mensch besitzt die natürliche Fähigkeit, verschiedene Personen oder Objekte nach Ähnlichkeit zu gruppieren. Diese Fähigkeit wird informatisch mit der Clusteranalyse nachgeahmt. Eine Menge von

Objekten wird abhängig von der Ähnlichkeit in verschiedene Gruppen aufgeteilt. Die Ähnlichkeit wird dabei anhand verschiedener Eigenschaften des Objektes bestimmt. Je nachdem welche Eigenschaften gewählt wurden, ergeben sich trotz der gleichen Menge an Objekten unterschiedliche Gruppen.

Formal kann eine einfache Clusteranalyse der endlichen Menge X in k Gruppen C_1, C_2, \dots, C_k wie folgt definiert werden.

- $\bigcup_{i=1}^k C_i = X$
- $C_i \cap C_j = \emptyset, i \neq j, i, j = 1, \dots, k$

Die zweite Bedingung ist bei der *Fuzzy-Clusteranalyse* nicht erfüllt, da dort ein Datensatz auch zu mehreren Clustern gehören kann.

Auf Grund der Menge an verschiedenen Cluster-Verfahren ist es im Rahmen der vorliegenden Arbeit nicht möglich, eine vollständige Übersicht über alle Verfahren zu geben. Die meisten Verfahren zur Clusteranalyse wurden ursprünglich für die Analyse statischer Daten entwickelt. Eine Übersicht und Kategorisierung zu Verfahren der Clusteranalyse findet man bei Rui Xu und Wunsch (2005) sowie Tan, Steinbach, Karpatne und Kumar (2018).

Ähnlichkeits- und Distanzmaße Einen wesentlichen Baustein in vielen Verfahren zur Datenanalyse bilden Ähnlichkeits- oder Distanzmaße. Bei der Clusteranalyse wird das Maß dazu genutzt, die Ähnlichkeit zwischen Datensätzen zu messen. Basierend auf der Ähnlichkeit werden die Datensätze verschiedenen Gruppen zugeordnet. Bei der Klassifizierung wird über ein Maß die Ähnlichkeit zu schon bekannten Datensätzen und deren Kategorien bestimmt. Ein Ähnlichkeitsmaß erreicht sein Maximum (für gewöhnlich 1) bei identischen Datensätzen. Das Distanzmaß verhält sich genau entgegengesetzt, für identische Datensätze erhält man das Minimum (für gewöhnlich 0).

Ein Distanzmaß auf der Menge X ist eine Funktion von $d : X \times X \rightarrow \mathbb{R}$ mit

- $d(x, x) = 0, \forall x \in X$
- $d(x, y) = d(y, x), \forall x, y \in X.$

Ein Ähnlichkeitsmaß ist eine Funktion von $d : X \times X \rightarrow \mathbb{R}$ mit

- $d(x, x) = 1, \forall x \in X$
- $d(x, y) = d(y, x), \forall x, y \in X.$

Die Maße sind metrisch, wenn zusätzlich noch die Dreiecksungleichung erfüllt ist und aus dem Erreichen des minimalen bzw. maximalen Abstandes folgt, dass die Datensätze gleich sind.

- $d(x, y) = 0$ bzw. $1 \Leftrightarrow x = y$
- $d(x, z) \leq d(x, y) + d(y, z), \forall x, y, z \in X$

Im statischen Fall wird für den Abstand zwischen zwei Punkten $x, y \in \mathbb{R}^n$ am häufigsten die durch die L_p Norm induzierte *Minkowski Metrik* verwendet.

$$d_p(x, y) = \left(\sum_{i=1}^n |x_i - y_i|^p \right)^{\frac{1}{p}}$$

Für $p = 1$ erhält man die *Manhattan Metrik*,

$$d_1(x, y) = \sum_{i=1}^n |x_i - y_i|$$

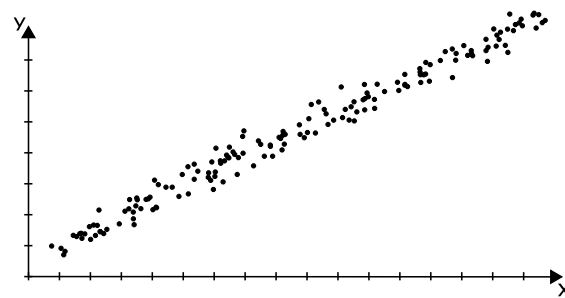
für $p = 2$ den *Euklidischen Abstand*,

$$d_2(x, y) = \sqrt{\sum_{i=1}^n |x_i - y_i|^2}$$

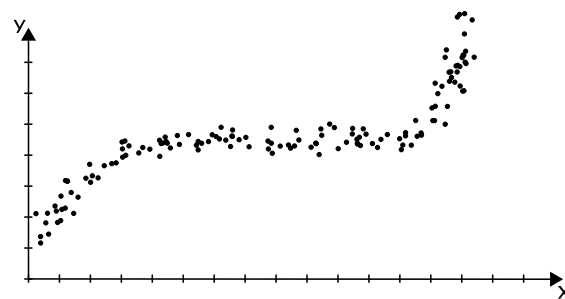
für $p = \infty$ die *Chebychev Metrik*

$$d_\infty(x, y) = \max_{1 \leq i \leq n} |x_i - y_i|.$$

Dimensionsreduktion Statische Daten liegen oft als Elemente eines in der Regel metrischen Raumes vor, z.B. als reelle Merkmalsvektoren. Oft sind einzelne Einträge der Vektoren nicht unabhängig voneinander. Beispielsweise kann es einen Zusammenhang zwischen der zurückgelegten Strecke und der durchschnittlichen Geschwindigkeit geben. Werden Vektoren verglichen, die mehrere abhängige Einträge enthalten, kann dies die Clusteranalyse und Klassifizierung erschweren. Bei der Abstandsberechnung gehen die abhängigen Merkmale mehrfach ein und überlagern so eventuell die unabhängigen Merkmale. Eine Gewichtung der einzelnen Merkmale in der Abstandsberechnung ist schwierig, da diese a priori unbekannt ist. Die intrinsische Dimension der Daten ist in diesem Fall kleiner als die Dimension der Rohdaten. Durch Verfahren zur Dimensionsreduktion werden die Rohdaten in einen niedriger dimensionalen Raum transformiert.



(a) Lineare Daten



(b) Nicht-lineare Daten

Abbildung 2.1: Intrinsische Dimension, lineare und nicht-lineare Daten

Es lassen sich lineare von nicht-linearen Dimensionsreduktionsverfahren unterscheiden. In [Abbildung 2.1 \(a\)](#) besteht ein linearer Zusammenhang zwischen den Variablen x und y . Die Daten können durch eine Gerade

approximiert werden, da die intrinsische Dimension nur 1-dimensional ist. Allgemein werden bei der linearen Dimensionsreduktion die Daten durch den Mittelwert, eine Gerade, eine Ebene usw. approximiert, je nachdem welche intrinsische Dimension vorliegt. Die wohl bekannteste lineare Dimensionsreduktion ist die von Pearson (1901) entwickelte *Hauptkomponentenanalyse* (PCA). Die *Hauptkomponentenanalyse* minimiert die Varianz der Abweichungen. Das von Torgerson (1958) entwickelte *Multi Dimensional Scaling* (MDS) versucht dagegen möglichst die Abstände zwischen den Daten zu erhalten. Es gibt zahlreiche weitere Verfahren zur mehrdimensionalen Dimensionsreduktion. Eine quantitative Übersicht findet sich bei Espadoto, Martins, Kerren, Hirata und Telea (2019). In den meisten Anwendungsfällen liegen nicht-lineare Zusammenhänge in den Daten vor. Wie in der Arbeit von Van Der Maaten, Postma und Van den Herik (2009) dargelegt, sind Verfahren zur nicht-linearen Dimensionsreduktion jedoch oft nicht in der Lage, bei natürlichen Datensätzen bessere Ergebnisse zu erzielen als die Hauptkomponentenanalyse.

Die im Folgenden verwendeten *Selbstorganisierenden Karten* bilden die hochdimensionalen Eingabedaten auf einen niedriger dimensional Raum ab. Dabei können nicht-lineare Daten abgebildet werden, gleichzeitig wird auch die lokale Topologie erhalten. Daten die bezüglich eines Distanz- oder Ähnlichkeitsmaßes im Eingaberaum nah beieinander liegen, sind auch im Ausgaberaum benachbart. Gegenüber den oben genannten Verfahren entsteht hier die Dimensionsreduktion iterativ aus dem unüberwachten Lernen der Eingabedaten. Die Dimensionsreduktion erfolgt somit nicht aus dem Minimieren einer Kostenfunktion über alle Datensätze. Es gibt zahlreiche Anwendungsfälle, in denen *Selbstorganisierende Karten* erfolgreich zur Dimensionsreduktion eingesetzt wurden.

Zeitreihen Verändern sich die Merkmale über die Zeit, ergeben sich sequenzielle Daten bzw. Zeitreihen. Das Ziel einer Zeitreihenanalyse kann z.B. die Gruppierung ähnlicher Zeitreihen sein. Auf Basis der gebildeten Gruppen können dann ungewöhnliche Zeitreihen identifiziert werden.

Fakhrazari und Vakilzadian (2017) geben einen Querschnitt über Verfahren zur Zeitreihenanalyse. Es gibt verschiedene Erweiterungen *Selbstorganisierender Karten* zur Verarbeitung von Zeitreihen. Eine Übersicht und Kategorisierung wurde von Guimarães, Lobo und Moura-Pires (2003) erstellt. Auf *Selbstorganisierende Karten* wird in Kapitel 2.2 noch gesondert eingegangen. Generell lässt sich feststellen, dass viele Verfahren der statischen Datenanalyse auf Zeitreihen erweitert werden, etwa indem der *Euklidische Abstand* durch Ähnlichkeitsmaße auf Zeitreihen ersetzt wird.

Räumlich-zeitliche Daten Verfahren zur räumlich-zeitlichen Datenanalyse werden eingesetzt, wenn zusätzlich zu den zeitlichen auch räumliche Eigenschaften untersucht werden. Kisilevich, Mansmann, Nanni und Rinzivillo (2009) teilen räumlich-zeitliche Daten anhand verschiedener Kriterien ein. In räumlicher Dimension werden Personen oder Objekte danach eingeteilt, ob sie sich an einem festen Ort befinden oder ihre Position verändern. Weiter wird unterschieden, ob die Daten einmalig erfasst oder fortlaufend aktualisiert werden. Im Fall fortlaufender Aktualisierung kann zudem unterschieden werden, ob jeweils nur der letzte Zeitpunkt oder die gesamte Historie bekannt ist. Schließlich können die vorliegenden Daten Punkte, Linien oder Flächen repräsentieren. Positionsdaten im Sportspiel sind entsprechend bei den Trajektorien einzuordnen.

Die Ziele, die mittels räumlich-zeitlicher Verfahren verfolgt werden, ähneln vielfach den Zielen einer Analyse von statischen Daten oder Zeitreihen. Beispielweise können räumlich-zeitliche Verläufe gruppiert werden, um typische Verläufe zu finden oder um Ausreißer zu identifizieren. Auch in diesem Fall werden statische Verfahren durch ein geeignetes Distanzmaß auf räumlich-zeitliche Daten übertragen.

Ein einfaches Distanzmaß zwischen zwei Mengen X, Y ist der durch Hausdorff (1914) definierte *Hausdorff Abstand*

$$d_H(X, Y) = \max \left\{ \sup_{x \in X} \inf_{y \in Y} d(x, y), \sup_{y \in Y} \inf_{x \in X} d(x, y) \right\},$$

wobei d eine Metrik ist.

Werden zwei Wege miteinander verglichen, so entspricht der *Hausdorff Abstand* nicht dem natürlichen Empfinden von Ähnlichkeit. In [Abbildung 2.2](#) sind zwei Wege abgebildet, die einen sehr kleinen *Hausdorff Abstand* haben, vom Verlauf jedoch sehr unterschiedlich sind.

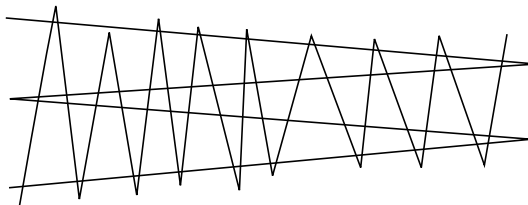


Abbildung 2.2: Zwei Wege mit kleinem *Hausdorff Abstand* d_H (nach Alt und Godau (1995))

Der durch Fréchet (1906) definierte *Fréchet Abstand* berücksichtigt auch den Verlauf der Wege. Der *Fréchet Abstand* ist ein Maß für die Ähnlichkeit von zwei parametrisierten Kurven. Zur Veranschaulichung wird oft das Beispiel eines Mannes gewählt, der einen Hund an der Leine führt. Der Mann und der Hund folgen jeweils ihrem eigenen Weg. Sie können unabhängig voneinander ihre Geschwindigkeit variieren oder auch stehen bleiben, jedoch nicht rückwärtsgehen. Der *Fréchet Abstand* entspricht dann der kürzesten benötigten Leine unter allen Verläufen. Diese zwei Maße stellen Extremfälle dar. Der *Hausdorff Abstand* ist schnell zu berechnen aber sehr ungenau. Der *Fréchet Abstand* entspricht dem natürlichen Empfinden ähnlicher Wege, ist aber schwierig zu berechnen. In der Praxis finden oft Vereinfachungen des *Fréchet Abstands* Anwendung. Eine Übersicht über räumlich-zeitliche Verfahren findet man bei Atluri, Karpatsne und Kumar (2018).

2.1.3 Sportinformatik

Zur Bewertung taktischen Verhaltens auf Basis von Positionsdaten sind verschiedene Leistungsindikatoren entwickelt worden. Eine Übersicht und Kategorisierung findet man bei Rein und Memmert (2016) sowie Memmert, Lemmink und Sampaio (2017). Einige der Verfahren wurden entwickelt, um einzelne Aspekte taktischer Interaktionsprozesse wie etwa die Raumkontrolle zu untersuchen. Dagegen sind *Künstliche Neuronale Netze* breiter anwendbar. Insbesondere die Klasse der *Selbstorganisierenden Karten* wird in einer Vielzahl von Untersuchungen erfolgreich zur Analyse von Sportspielprozessen angewendet.

Lames und Perl (1999) klassifizieren Ballwechselerlaufstypen mit *Neuronalen Netzen* am Beispiel Volleyball. Perl (2002b) erlernt mit Hilfe eines *Künstlichen Neuronalen Netzes* charakteristische Schlagpositionsfolgen im Squash. Die Abbildung des Spielprozesses durch eine Trajektorie auf dem Netz ermöglicht die Untersuchung intra-individueller und inter-individueller Eigenschaften. In Pfeiffer und Perl (2006) wird ein *DyCoN* zur Analyse taktischer Strukturen im Handball eingesetzt. In Jäger et al. (2007) werden zum einen durch ein *DyCoN* charakteristische Positionsmuster im Volleyball bestimmt und zum anderen werden die charakteristischen Positionsmuster beider Teams miteinander verglichen. Die Positionsmuster unterscheiden sich in ihrer Ausprägung und Häufigkeit. Diese Unterschiede deuten auf unterschiedliche taktische Konzepte hin.

Perl und Memmert (2011) untersuchen taktische Interaktionsmuster im Fußball auf Basis von Positionsdaten. Im ersten Schritt werden die Positionsdaten der Angreifenden und der Verteidigenden normalisiert. Über den Schwerpunkt lassen sich die Positionsdaten in einer Formation zusammenfassen, die unabhängig von der Position auf dem Spielfeld ist. Mit den Formationen der Angreifenden und Verteidigenden wird jeweils ein *DyCoN* trainiert. Die gesamte Analyse wird in dem entwickelten Tool SOC-

CER zusammengefasst. Über eine grafische Matrix lässt sich die Verteilung der Interaktionsmuster darstellen. Wählt man eines der Interaktionsmuster aus, so erhält man eine Liste der Sequenzen, die durch dieses Interaktionsmuster abgebildet werden. Das *SOCER* Tool ermöglicht darüber hinaus eine Bewertung unterschiedlicher taktischer Muster und die Darstellung des Interaktionsprozesses in einem Phasendiagramm.

2.2 Selbstorganisierende Karten

Das Prinzip der *Selbstorganisierenden Karten* hat seinen Ursprung in der Biologie. Das *Zentrale Nervensystem* von Wirbeltieren erstellt in vielen Kontexten (auditiv, visuell, sensomotorisch, usw.) Topologie erhaltende Abbildungen. Ein Beispiel ist die Weiterleitung von Reizen der Retina in den Cortex (Talbot, S. A. & Marshall, W. H., 1941). Es wurde festgestellt, dass benachbarte Bereiche auf der Retina benachbarte Bereiche auf dem Cortex aktivieren. Die Bildung Topologie erhaltender Abbildungen war im Folgenden Gegenstand zahlreicher Untersuchungen. In ihrem Artikel stellen Willshaw und von der Malsburg (1976) eine Theorie der Selbstorganisation vor, um die Entstehung einer Topologie erhaltenden Abbildung zu erklären. Die von Kohonen (1982, 1990, 2001, 2013) entwickelte *Selbstorganisierende Karte* generalisiert das Konzept der Selbstorganisation. Ausgehend von der Modellierung physiologischer Prozesse entstand so ein Verfahren, das heute Anwendung in einer Vielzahl verschiedener Fachrichtungen findet.

2.2.1 Self-Organizing Map

Die *Selbstorganisierende Karte* besteht aus einer Menge von Neuronen

$$\mathcal{N} = \{1, \dots, m\}.$$

Jedem Neuron wird ein Gewichtsvektor $w_i \in \mathbb{R}^n$ zugeordnet. Die Neuronen werden in der Regel in einem zweidimensionalen Gitter angeordnet (siehe Abb. 2.3). Im Weiteren wird der Begriff Neuron der Einfachheit halber synonym für den Gewichtsvektor verwendet.

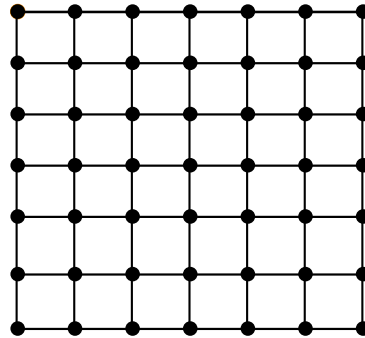


Abbildung 2.3: Schematische Darstellung einer zweidimensionalen *Self-Organizing Map*

Zunächst wird das Netz mit einer Menge X von Daten trainiert, wobei $\mathbf{x}(t) \in \mathbb{R}^n$ mit $t = 1, \dots, n$ ist. Die Datenvektoren $\mathbf{x}(t)$ werden in randomisierter Abfolge dem Netz präsentiert. Die Gewichtsvektoren werden vor dem ersten Training mit Zufallszahlen initialisiert. Zunächst wird das Neuron mit dem kleinsten Abstand bestimmt.

$$I = \arg \min_i \|\mathbf{x}(t) - \mathbf{w}_i(t)\|$$

Dieses Neuron wird im Folgenden als Gewinnerneuron bezeichnet. Die Gewichtsvektoren des Gewinnerneurons und seiner unmittelbaren Nachbarn werden im zweiten Schritt angepasst. Der Grad der Anpassung hängt dabei von der Entfernung zum Gewinnerneuron ab.

$$\Delta w_i(t) = \alpha(t) \cdot h(d_{\mathcal{N}}(i, I), t) \cdot (\mathbf{x}(t) - w_i(t))$$

Die Lernrate wird über eine monoton fallende Funktion $\alpha(t)$ gesteuert. Die Nachbarschaftsfunktion $h(\cdot)$ ist ebenfalls monoton fallend. Die Funktion

$$d_{\mathcal{N}}(i, I) : \mathcal{N} \times \mathcal{N} \rightarrow \mathbb{R}$$

gibt die Distanz zum Gewinnerneuron an. Im einfachsten Fall wird die Distanz über eine der Metriken (Kapitel 2.1.2) bestimmt. Für die Nachbarschaftsfunktion wird oft eine Gaußverteilung verwendet.

$$h(d_{\mathcal{N}}(i, I), t) = \exp \frac{-d_{\mathcal{N}}(i, I)}{2\sigma^2(t)}$$

Die Funktion $\sigma(t)$ fällt mit der Zeit monoton ab. Die Lernrate und die Nachbarschaftsfunktion zusammen bewirken zum Anfang des Trainings, dass zum einen idealerweise alle Neuronen angepasst werden und zum anderen, dass die Anpassungen entsprechend groß sind. Mit zunehmenden Training werden immer weniger Neuronen pro Trainingsschritt angepasst. Gleichzeitig werden die Anpassungen an die Eingabedaten immer kleiner. Dadurch kann das Netz anfangs eine globale Struktur ausbilden, die sich zunehmend stabilisiert. Durch das Training entsteht eine Topologie erhaltende Abbildung der Daten aus dem n-dimensionalen Raum in den zweidimensionalen Raum. Zudem spiegelt die Verteilung der Neuronen auch die ursprüngliche Verteilung der Daten wieder. Die Bereiche des Eingaberaumes, in denen sich die Daten gehäuft haben, werden entsprechend durch mehr Neuronen repräsentiert als die Bereiche, in denen sich fast keine Daten befunden haben.

2.2.2 Dynamically Controlled Network

Das *Dynamically Controlled Network (DyCoN)* wurde auf Basis der *Selbstorganisierenden Karte* von Perl (2001b) entwickelt. Beim SOM nehmen sowohl die Lernrate als auch die Nachbarschaftsfunktion mit der Zeit ab. Das

hat zur Folge, dass am Anfang des Trainings weite Bereiche des Netzes adaptiert werden und sich die Gewichtsvektoren an die Trainingsdaten anpassen. Gegen Ende des Trainings finden nur noch lokal begrenzte Anpassungen statt. Dies führt zu einer Stabilisierung der Gewichtsvektoren, verhindert gleichzeitig aber, dass sich das SOM an neue Daten anpassen kann.

Das *DyCoN* ist in der Lage kontinuierlich zu lernen. Dazu kommt ein *Performance Potential Metamodell (PerPot)* (Mester & Perl, 2000) zur dynamischen Steuerung der Nachbarschaftsfunktion zum Einsatz. Das *PerPot* wurde ursprünglich zur Analyse und Optimierung physiologischer Adaptionsprozesse entwickelt. Eine grundlegende Idee ist, dass sich eine Belastung immer antagonistisch auf die Leistung auswirkt. Durch die Belastung steigert sich die Leistung mit einer zeitlichen Verzögerung und gleichzeitig führt die Belastung auch mit einer zeitlichen Verzögerung zur Ermüdung. Ermüdung, Erholung und Leistung werden im *PerPot* Modell durch die entsprechenden Potentiale abgebildet (Abbildung 2.4).

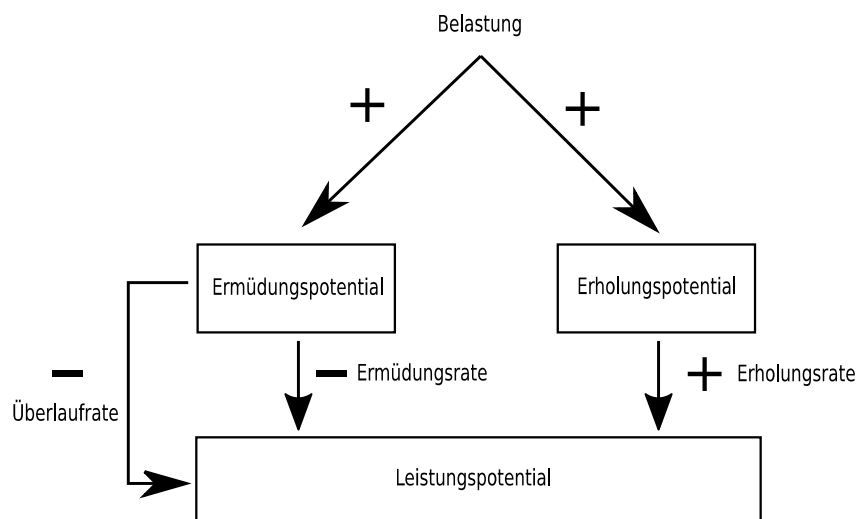


Abbildung 2.4: Schematische Darstellung des *PerPot* Metamodells

Über die Ermüdungs- und Überlafrate wirkt sich das Ermüdungspotential negativ auf das Leistungspotential aus. Im Gegensatz dazu wirkt sich die Erholungsrate positiv auf das Leistungspotential aus. Durch geeignete

Wahl der Verzögerungsparameter ist es möglich, unterschiedliche physiologische Phänomene, wie z.B. die Superkompensation nach einer hohen Belastung mit anschließender Ruhepause, abzubilden. Durch den Einsatz des *PerPot* Modells werden diese Eigenschaften auf die *Selbstorganisierende Karte* übertragen.

Über einen Aktivierungsradius werden die Neuronen um das Gewinnerneuron bestimmt, die sich an den Trainingsreiz anpassen. Das Leistungspotential wird zur Steuerung des Aktivierungsradius genutzt. Ein hohes Leistungspotential stellt dabei den Lernfortschritt des Neurons dar und resultiert in einem kleinen Aktivierungsradius. Dadurch ist das einzelne Neuron in der Lage seine Information zu stabilisieren. Wird das Neuron in den folgenden Lernschritten nicht mehr aktiviert, nimmt das Leistungspotential wieder ab (Atrophie). Die Lernrate wird ebenfalls über den Lernfortschritt gesteuert, beispielsweise wird die Distanz zwischen dem Neuron und dem Datenvektor bestimmt. Ein große Distanz resultiert dabei in einer kleinen Lernrate, die zu einer entsprechend kleinen Anpassung an den Trainingsdatensatz führt.

Durch die kontinuierliche Lerndynamik kann das *DyCoN* zunächst mit Monte-Carlo generierten Daten vortrainiert und anschließend mit den echten Daten trainiert werden. Dadurch sind auch Untersuchungen möglich, bei denen nur sehr wenige Daten vorliegen.

2.2.3 Merge Self-Organizing Map

Die von nach Strickert und Hammer (2003) entwickelte *Merge Self-Organizing Map (MSOM)* erweitert jedes Neuron um einen zusätzlichen Kontextvektor. Dadurch geht der vorhergehende Trainingsschritt in den nachfolgenden Schritt ein. Bei der Abstandsberechnung wird sowohl der aktuelle Datensatz als auch das letzte Gewinnerneuron in Form des Kontextvektors berücksichtigt. In jedem Trainingsschritt bestimmt man zunächst das Neuron mit dem kleinsten Abstand.

$$I = \arg \min_i ((1 - \alpha) \cdot \|\mathbf{x}(t) - \mathbf{w}_i(t)\| + \alpha \cdot \|\mathbf{c}(t) - \mathbf{c}_i(t)\|)$$

Der Kontext $\mathbf{c}(t)$ ist durch

$$\mathbf{c}(t) = (1 - \beta) \cdot \mathbf{w}(I_{t-1}) + \beta \cdot \mathbf{c}(I_{t-1})$$

gegeben. Der Kontextvektor ist dabei eine Linearkombination des Gewichts- und Kontextvektors des vorgehenden Gewinnerneurons. Für den ersten trainierten Datenvektor gibt es kein vorhergehendes Gewinnerneuron, hier wird $c(t) = 0$ gewählt. Der Parameter β liegt im Intervall $[0, 1]$, für gewöhnlich wird $\beta = 0.5$ gewählt. α wird derart gesteuert, dass der Einfluss des Kontextvektors auf die Berechnung des Abstands mit fortschreitendem Training zunimmt. Die Gewichts- und Kontextvektoren des Gewinnerneurons und seiner unmittelbaren Nachbarn werden im zweiten Schritt angepasst. Der Grad der Anpassung hängt dabei von der Entfernung zum Gewinnerneuron ab.

$$\begin{aligned} \Delta w_i(t) &= \gamma_1(t) \cdot h(d_{\mathcal{N}}(i, I), t) \cdot (x(t) - w_i(t)) \\ \Delta c_i(t) &= \gamma_2(t) \cdot h(d_{\mathcal{N}}(i, I), t) \cdot (c(t) - c_i(t)) \end{aligned}$$

Die Lernrate und Nachbarschaftsfunktion sind wie beim SOM monoton fallende Funktionen. Durch diese Erweiterung kann das MSOM sequenzielle Daten verarbeiten.

2.3 Zusammenfassung

Die Erfahrung von Trainierenden kann die Bewertung durch Leistungsindikatoren unterstützen. Gleichzeitig sind Trainierende nach einer Studie von Franks und Miller (1991) nur in der Lage 40% der relevanten Informationen eines Spiels wiederzugeben. Durch Verfahren des maschinellen

Lernens wird versucht, dieses Fachwissen aus Beispielen zu erlernen, um es dann auf neue Spiele anwenden zu können. Durch das Erlernen aus Beispielen vermeidet man zudem das schwierige und aufwändige Erfassen von Fachwissen durch ein Regelsystem.

Aus informatischer Sicht sind Interaktionsdaten gleichzeitig statische, zeitliche und räumlich-zeitliche Daten. Daher kann grundsätzlich jedes der Verfahren für statische, zeitliche und räumlich-zeitliche Daten auch auf Positionsdaten im Sportspiel angewendet werden. Abhängig vom angewendeten Verfahren sind bestimmte Untersuchungen jedoch nicht möglich. Mit einer rein statischen Datenanalyse kann man keine Aussagen über den zeitlichen Verlauf der Daten machen. Räumlich-zeitliche Verfahren berücksichtigen nicht die Interaktion zwischen verschiedenen Elementen. Folglich kann eine Interaktionsanalyse mit diesen Verfahren allein nicht durchgeführt werden.

Selbstorganisierende Karten wurden in mehreren sportinformatischen Untersuchungen erfolgreich eingesetzt. Daher werden diese als Basis für das eigene Verfahren gewählt. In der vorliegenden Arbeit wird ein Verfahren entwickelt, das geeignet ist, vorgegebene mannschaftstaktische Interaktionsmuster auf Basis einer Hierarchie von *DyCoN* Netzen zu erlernen. Am Beispiel Fußball wird eine Klassifizierung von kurzen und langen Spieleröffnungen durchgeführt. Der hierarchische Ansatz wird auch auf den Basketballsport übertragen und zur Klassifizierung von drei taktischen Interaktionsprozessen verwendet. Schließlich wird auch eine Klassifizierung auf Basis eines *MSOMs* durchgeführt und mit dem hierarchischen Ansatz verglichen. Die Klassifizierung von mannschaftstaktischen Interaktionsprozessen auf Basis von Positionsdaten wurde vorher weder im Fußball- noch im Basketballspiel durchgeführt und stellt eine Neuerung dar.

Kapitel 3

Veröffentlichungen

Die folgenden Artikel wurden im Peer Review Verfahren begutachtet und veröffentlicht. Sie werden in chronologischer Reihenfolge aufgeführt und jeweils mit einer kurzen Zusammenfassung eingeleitet.

3.1 Analysis and Simulation of Actions in Games by Means of Special Self-Organizing Maps

Grunz, A., Memmert, D. & Perl, J. (2009). Analysis and simulation of actions in games by means of special self-organizing maps. *International Journal of Computer Science in Sport*, 8 (1).

Zusammenfassung

Selbstorganisierende Karten sind in der Lage, Prozesse im Sport abzubilden und dabei die wesentlichen Informationen und die Dynamik der Prozesse zu erhalten (Perl, 2001a, 2002a, 2008; Schmidt, 2012). Im ersten Teil der Publikation wird der Frage nachgegangen, wie sich mannschaftstaktische Spielprozesse im Fußball mit Hilfe von *Selbstorganisierenden Karten* modellieren lassen. Am Beispiel kurzer und langer Spieleröffnungen werden exemplarisch die Probleme erläutert und Lösungen vorgeschlagen. Im zweiten Teil wird mit den Positionsdaten des WM Finales 2006 auf Basis des zuvor entwickelten Ansatzes eine Interaktionsanalyse durchgeführt. Dabei lässt sich der komplexe Interaktionsprozess durch synchrone Phasendiagramme abbilden. Dies ermöglicht zum einen die Erkennung taktischer Muster und zum anderen die Bewertung von Aktionen. Zum Abschluss wird ein Ausblick auf die zukünftige Forschung geben. Auf Basis des entwickelten Modells lässt sich eine Simulation des Spiels durchführen. Das ermöglicht insbesondere auch die Untersuchung kreativen Verhaltens. Kreatives Verhalten ist durch die Seltenheit und Adäquatheit der Aktionen charakterisiert (Sternberg & Lubart, 1998). Es wird vorgeschlagen, die Adäquatheit über eine Simulation zu bestimmen.

3.2 Tactical pattern recognition in soccer games by means of special self-organizing maps

Grunz, A., Memmert, D. & Perl, J. (2012). Tactical pattern recognition in soccer games by means of special self-organizing maps. *Human Movement Science*, 31 (2), 334-343.

Zusammenfassung

Ausgehend von der ersten Untersuchung wird eine Architektur von *DyCoN* Netzen entworfen, um taktische Interaktionsmuster im Fußball zu klassifizieren. Dazu werden Konstellationen aus den Verteidigenden und den Angreifenden gebildet. Zusätzlich werden die Positionsdaten in x - und y -Koordinaten getrennt. Für jede Konstellation bzw. den Ball werden jeweils getrennt nach x - und y -Koordinaten *DyCoN* Netze trainiert und geclustert. Im folgenden Schritt werden aus den Netz-Trajektorien mit Hilfe der *Sliding Window* Technik neue Sequenzen generiert. Auf einer zweiten Ebene werden die Sequenzen wiederum zum Training weiterer *DyCoN* Netze verwendet. Im letzten Schritt werden die Aktivierungen der Netze der zweiten Ebene in einem neuen Datenvektor zusammengefasst. Mit diesen Daten wird ein drittes Netz trainiert. Durch den Einsatz der Netze auf verschiedenen Ebenen erhält man verschiedene Abstraktionen. Auf der ersten Ebene kodiert das neuronale Netz prototypisch verschiedene Konstellationen. Auf der zweiten Ebene werden prototypisch Verläufe der Konstellationen bzw. des Balls abgebildet. Die dritte Ebene abstrahiert schließlich eine Interaktion aus den Verteidigenden, den Angreifenden und dem Ball. Über eine semantische Kalibrierung kann das Interaktionsnetz mit verschiedenen Kategorien geprägt werden. Exemplarisch wird eine Klassifizierung von kurzen und langen Spieleröffnungen anhand der Positionsdaten des WM Finales 2006 durchgeführt. Dazu wurden im Vorfeld über ein entwickeltes Programm die Kategorien für das WM Finale 2006 erfasst. Diese Daten werden im Anschluss genutzt, um einerseits eine Kalibrierung durchzuführen und andererseits das kalibrierte

Netz zu testen. Als Ergebnis konnten 84% der kurzen und langen Spieleröffnungen richtig erkannt werden.

3.3 Detecting tactical pattern in basketball: Comparison of merge self-organising maps and dynamic controlled neural networks

Kempe, M., Grunz, A. & Memmert, D. (2014). Detecting tactical patterns in basketball: Comparison of merge self-organising maps and dynamic controlled neural networks. *European Journal of Sport Science*, 15 (4), 249-255.

Zusammenfassung

In der publizierten Studie wird der entwickelte hierarchische Ansatz zur Klassifizierung von drei taktischen Mustern im Basketball angewendet. Hierzu wird zunächst ein *DyCoN* mit den Positionsdaten der 5 Spielenden einer Mannschaft trainiert und anschließend eine Clusterung durchgeführt. Im Anschluss werden über die *Sliding Window* Technik die Trajektorien auf den Clustern in Sequenzen umgewandelt. Diese Sequenzen werden für das Training eines zweiten *DyCoN* Netzes verwendet. Durch eine semantische Prägung können die einzelnen Kategorien auf dem zweiten Netz verortet werden. Parallel wird auch ein alternativer Ansatz genutzt: Eine rekursive *Merge Self-Organizing Map* (*MSOM*, Strickert und Hammer (2005)) wird mit denselben Positionsdaten aus der ersten Halbzeit trainiert. Das *MSOM* erlaubt die direkte Verarbeitung von Zeitreihen. Dazu erhält jedes Neuron zusätzlich einen Kontextvektor. Dieser enthält die vorhergehende Aktivierung in Form einer gewichteten linearen Approximation der Gewichtsvektoren der letzten Gewinnerneuronen. Daraufhin wird auch das *MSOM* mit den kategorisierten Daten semantisch geprägt. Beide Ansätze werden schließlich mit den Positionsdaten der zweiten Halbzeit evaluiert. Dabei sind beide Verfahren in der Lage, die taktischen Muster im Basketball auch bei unbekanntem Daten zu klassifizieren. Das hierarchische Verfahren erreicht eine Genauigkeit von 97,5% gegenüber 80% beim *MSOM*.

Kapitel 4

Zusammenfassung

In diesem Kapitel werden zunächst die Ergebnisse im Gesamtzusammenhang diskutiert. Schließlich wird ein Ausblick auf Bereiche zukünftiger Forschung gegeben.

In der ersten Veröffentlichung wird ein neues Modell vorgestellt, um mannschaftstaktische Interaktionsprozesse im Fußball mit Hilfe *Selbstorganisierender Karten* abzubilden. Dazu werden Gruppen von Spielenden zu Konstellationen zusammengefasst, und mit diesen Daten wird jeweils ein *DyCoN* trainiert und geclustert. Mit der *Sliding Window* Technik kann der zeitliche Verlauf der Prozesse auf den Netzen erfasst werden. Die daraus entstehenden Sequenzen können für die weitere Analyse genutzt werden. Anhand der Positionsdaten des WM Finales 2006 werden Gruppen aus den Verteidigenden und den Angreifenden beider Mannschaften gebildet. Die sich daraus ergebenden Sequenzen können zusammen für das Training eines weiteren Netzes verwendet werden. Die Neuronen auf dem übergeordneten Netz bilden prototypisch die Interaktionen zwischen Verteidigung und Angreifenden ab. Die Trajektorie auf diesem Netz bildet die mannschaftstaktischen Interaktionsprozesse ab. Die Dynamik des Interaktionsprozesses bleibt dabei erhalten. Insgesamt wird durch das

entwickelte Modell die Modellierung mannschaftstaktischer Interaktionsprozesse erfolgreich umgesetzt und damit das erste Teilziel erreicht.

Die Ergebnisse aus dem ersten Teil des Forschungsvorhabens werden im zweiten Teil zu einer Hierarchie von *DyCoN* Netzen weiterentwickelt. Dadurch kann eine automatische Klassifizierung von kurzen und langen Spieleröffnungen realisiert werden. Durch die Abstraktion auf mehreren Ebenen wird die Komplexität der Interaktionsprozesse im Fußball handhabbar gemacht. Theoretisch gibt es bei einer kurzen oder langen Spieleröffnung unbegrenzt viele Ausführungsmöglichkeiten. Die Anzahl der am Spielzug Teilnehmenden kann sich unterscheiden, die Anzahl der ausgeführten Aktionen, die Dauer der Ausführung usw.. Die Idee ist, die interagierenden Elemente durch jeweils ein eigenes Netz abzubilden. Die Abbildung reduziert jeweils die Vielzahl an Möglichkeiten auf eine Menge Prototypen. Die vier Verteidigenden, die vier Angreifenden und der Ball werden jeweils durch eigene *DyCoN* Netze repräsentiert. Um den Verlust an relevanter Informationen durch die Dimensionsreduktion klein zu halten, werden zusätzlich noch x- und y-Koordinaten unterschieden.

Die Clusterung nach dem Training fasst die Neuronen in Gruppen zusammen, die ähnliche Informationen kodieren. Der komplexe Interaktionsprozess wird entsprechend über die drei Ebenen auf einzelne aktivierte Neuronen auf der dritten Ebene reduziert. Über eine semantische Prägung bildet das Netz auf der dritten Ebene spezifisch Neuronen aus, die einzelne Kategorien erkennen. Durch die *Sliding Window* Technik ist ein *post-processing* Schritt nötig, um auch längere Verläufe zu erkennen. Mehrere kurze oder lange Spieleröffnungen werden zusammengefügt, wenn diese zeitlich aufeinander folgen. Insgesamt werden 84% der kurzen und langen Spieleröffnungen im WM Finale 2006 richtig erkannt. Neben kurzen und langen Spieleröffnungen werden auch Einwürfe, Freistöße, Ecken und Konterspiele klassifiziert. In Tabelle 4.1 sind die Ergebnisse der Klassifizierung dargestellt.

Kategorie	Fachkraft	Hierarchie <i>DyCoNs</i>	Erkennungsrate
Spieleröffnungen	131	110	84%
Einwürfe	27	27	100%
Freistöße	16	14	88%
Ecken	12	12	100%
Konterspiele	49	29	59%

Tabelle 4.1: Übersicht der Klassifizierungsergebnisse nach Kategorien

Durch die erfolgreiche Klassifizierung von kurzen und langen Spieleröffnungen wird das zweite definierte Teilziel erreicht.

Kritisch muss gesehen werden, dass das Verfahren bei einigen Szenen Schwierigkeiten hatte, zwischen einer kurzen und langen Spieleröffnung zu unterscheiden. Das lässt sich dadurch erklären, dass sich kurze und lange Spieleröffnungen bis auf den charakteristischen langen Pass durchaus ähneln können. Dadurch kann eine Sequenz, die eine lange Spieleröffnung kodiert, auch Einträge einer kurzen Spieleröffnung enthalten. Mit der Wahl der Spieleröffnung wurde ein relativ strukturiertes taktisches Interaktionsmuster gewählt. Gerade im Abschluss wird ein Spiel nochmal deutlich dynamischer. Dieser Umstand erklärt auch die schlechtere Erkennungsrate bei Konterspielen. Durch die erhöhte Variabilität in diesen Interaktionsprozessen wird der Eingaberaum nicht ausreichend auf die *Künstlichen Neuronalen Netze* abgebildet.

In der dritten Untersuchung wird das hierarchische Verfahren zur Klassifizierung von drei taktischen Interaktionsmustern im Basketball genutzt. Mit den Positionsdaten der ersten beiden Viertel und den drei jeweils 15-mal aufgezeichneten Spielzügen wird ein *DyCoN* trainiert und kalibriert. Gleichzeitig wird ein neuer Ansatz mit dem *MSOM* umgesetzt. Das *MSOM* gehört zu den rekurrenten *Selbstorganisierenden Karten* und ist damit in der Lage, Zeitreihen ohne die *Sliding Window* Technik zu

verarbeiten. Dadurch entfallen der *pre-* und *post-processing* Schritt, d.h. das Bilden von Sequenzen aus den Trajektorien und Zusammenfügen von zeitlich aufeinander folgenden Ergebnissen der gleichen Kategorie nach der Klassifizierung. Beide Ansätze sind dabei in der Lage, die drei Interaktionsmuster zu klassifizieren. Der hierarchische Ansatz mit den *DyCoN* Netzen erreicht dabei eine Genauigkeit von 97,5% gegenüber 80% beim *MSOM* Ansatz. Das *DyCoN* ist in der Lage, auch mit sehr wenigen Daten eine gute Abbildung des Eingaberaums auf die Neuronen zu erzielen. Für das Training und die Kalibrierung wurden nur Daten aus der ersten Halbzeit des Spiels verwendet. Daher liegt die Vermutung nahe, dass das *DyCoN* durch seine Lerndynamik eine differenziertere Abbildung der Daten erreichen konnte.

Mit der Erreichung des dritten Teilziels ist es insgesamt gelungen, das angestrebte Ziel, die Klassifizierung mannschaftstaktischer Interaktionsprozesse mittels Selbstorganisierender Karten, umzusetzen. Die erfolgreiche Umsetzung ermöglicht nun das Fachwissen eines Trainierenden basierend auf Beispielen zu erlernen und für die Klassifizierung neuer Spiele zu verwenden. Wie eingangs beschrieben, kann dies eine Analyse durch Leistungsindikatoren unterstützen.

Ausblick

Aus informatischer Sicht sind verschiedene Verbesserungen zur Steigerung der Genauigkeit denkbar.

Beim Festlegen des Start- und Endzeitpunktes einer klassifizierten Sequenz durch das hierarchische Verfahren kann es Abweichungen im Sekundenbereich geben. Fachleute legen Start und Ende genau durch den Zeitpunkt des Ballgewinns bzw. Ballverlustes fest. Das *Neuronale Netz* hingegen bestimmt die Zeitspanne nur indirekt durch die Klassifizierung. Um ein noch genaueres Ergebnis zu erzielen, könnte das Verfahren mit einem weiteren Ansatz kombiniert werden. Die Positionsdaten lassen sich

sehr gut mit Vektoren approximieren. Durch die Start- und Endpunkte der Vektoren erhält man die Punkte, an denen der Ball seine Richtung ändert. Nimmt man weiter an, dass dies nur durch Einwirken eines Spielenden passiert, kann man beispielsweise näherungsweise dem nächsten Spielenden den Ballbesitz zuschreiben. Auf diese Weise lassen sich die Zeitpunkte eines Ballgewinns bzw. -verlustes ermitteln. Nun kann man die klassifizierten Sequenzen an diesen Zeitpunkten ausrichten. Start- und Endzeitpunkte können somit zeitlich präziser festgelegt werden.

Ein weiterer vielversprechender Ansatz ist die Entwicklung einer Hierarchie aus rekurrenten *Selbstorganisierenden Karten*. Dadurch könnte man die Abstraktionsebenen aus der Hierarchie übernehmen und gleichzeitig auf das *pre-* und *post processing* verzichten. Nachteilig bei diesem Ansatz wäre jedoch der Verlust der Lerndynamik des *DyCoNs*.

Denkbar wäre auch die Verwendung einer *Selbstorganisierenden Karte* mit flexibler Netztopologie, beispielsweise das *Merge Neural Gas (MNG)* nach Strickert und Hammer (2003). Durch die flexible Netztopologie kann der Eingaberaum präziser abgebildet werden. Da das im Zuge dieser Arbeit entwickelte hierarchische Verfahren die Sequenzen für die nachfolgende Ebene immer aus den Cluster-Trajektorien gebildet werden, kann das Verfahren leicht auf Netze mit flexibler Topologie erweitert werden.

Das entwickelte Verfahren klassifiziert mannschaftstaktische Interaktionsprozesse, aber es wird keine Bewertung innerhalb einer Klasse durchgeführt. Insbesondere lassen sich keine guten von schlechten Spieleröffnungen unterscheiden. Auf der dritten Ebene kodiert ein Neuron eine komplette Interaktion aus Verteidigung, Angriff und Ball. Eine Spieleröffnung könnte beispielsweise als erfolgreich bewertet werden, wenn es gelänge den Ball in den gegnerischen Strafraum zu spielen und gleichzeitig im Ballbesitz zu bleiben. Damit würden lange Pässe, die in den gegnerischen Strafraum gespielt werden aber zum Ballverlust führen,

nicht als erfolgreich bewertet werden.

Wie in der ersten Veröffentlichung im Ausblick bereits dargelegt, lassen sich auch kreative Aktionen auf dem Netz untersuchen. Dazu werden Neuronen bestimmt, die zum einen selten aktiviert werden und zum anderen eine adäquate Aktion repräsentieren. Übertragen auf die Hierarchie kann man untersuchen, welche Neuronen auf dem Interaktionsnetz selten aktiviert werden und ob diese erfolgreiche Spieleröffnungen repräsentieren.

Kapitel 5

Abschluss

Seit der Veröffentlichung der drei vorliegenden Publikationen sind weitere Untersuchungen erschienen. Kapitel 5.1 stellt den Bezug zur aktuellen Forschung dar. In Kapitel 5.2 werden die Implikationen für die Wissenschaft dargelegt und in Kapitel 5.3 wird der Nutzen für die Praxis diskutiert.

5.1 Bezug zur aktuellen Forschung

Zentrales Thema der vorliegenden Arbeit ist die Klassifizierung von taktischen Interaktionsmustern basierend auf Positionsdaten. Die Abbildung wird an Hand kategorisierter Daten erlernt, welche im Vorfeld durch Fachkundige bestimmt wurden. Die Klassifizierung gehört zu den **überwachten Lernverfahren**. Eine entsprechende Recherche hat lediglich eine Veröffentlichung ergeben, die in diesen Punkten vergleichbar zu den publizierten Veröffentlichungen ist. Wang und Zemel (2016) klassifizieren offensive Strategien im Basketball an Hand von Positionsdaten. Zunächst werden die Positionsdaten der Spielenden und des Balls in eine bildliche Darstellung überführt. Dadurch übersetzen sich das Problem in eines der Bildklassifizierung. Aus der Menge kategorisierter Daten der NBA Saison

2013-2014 der Toronto Raptors werden 11 offensive Strategien mit insgesamt 1435 Spielsequenzen ausgewählt. Teilmengen der Sequenzen werden im Folgenden von Wang und Zebel für das Training und die Validierung verwendet. Ein *Rekurrentes Neuronales Netz* erlernt die Abbildung auf die verschiedenen Klassen. Mit dem automatisierten Verfahren lassen sich vergleichbare Ergebnisse erzielen wie bei der manuellen Klassifizierung. Der Algorithmus benötigt jedoch nur einen Bruchteil der Zeit zur Klassifizierung.

Die im Weiteren betrachteten Untersuchungen unterscheiden sich von diesem Ansatz in der Zielstellung und im verwendeten Verfahren.

Zur Analyse taktischer Interaktionsmuster werden vielfach **unüberwachte Lernverfahren** angewendet. Bialkowski et al. (2014) berechnen aus den Positionsdaten in einem iterativen Verfahren die Zuordnung der Trajektorien zu den einzelnen Rollen der Spielenden. Die Mannschaftsformation ergibt sich im Anschluss aus allen zugewiesenen Rollen. Bialkowski et al. (2015) entwerfen ein Verfahren zur Erkennung einer Mannschaft an Hand von drei Features. Auf Basis des zuvor entwickelten Verfahrens (Bialkowski et al. (2014)) definieren sie einen Deskriptor zur Bestimmung der Formation. Aus der Spielstatistik und dem Ballbesitz werden zwei weitere Deskriptoren ermittelt. Mit den skalierten und zusammengefassten Features führen die Autoren eine Lineare Diskriminanzanalyse durch. Über eine *k-Nearest-Neighbor* Methode wird schließlich eine Mannschaft zugewiesen. In der Veröffentlichung 3.2 werden die ausgewechselten Spielenden als Teil des *pre-processing* manuell in die entsprechenden Angriffs- bzw. Verteidigungskonstellationen eingesetzt. Ein automatisches Einsetzen über den Zeitpunkt der Auswechslung funktioniert nicht, da gleichzeitig zwei Spielende ausgewechselt wurden.

Knauf, Memmert und Brefeld (2016) entwickeln einen räumlich-zeitlichen *Kernel* auf mehreren Trajektorien. Auf Basis von fünf Bundesligaspie-

len führen die Autoren mit Hilfe des *Kernels* eine Clusteranalyse der Spieleröffnung und Torabschlüsse durch. Über einfache Regeln wie Ballbesitzwechsel und Spielzonen werden Sequenzen, die Spieleröffnungen und Torabschlüsse darstellen, aus den Positionsdaten extrahiert. Alle Sequenzen kürzer als 12 Sekunden werden verworfen. Im Ergebnis der Clusteranalyse lassen sich Unterschiede zwischen den Mannschaften in der Spieleröffnung und im Torabschluss ausmachen. Im Gegensatz zu diesem Vorgehen verwendet der hierarchische Ansatz der vorliegenden Arbeit kein Regelsystem, um einzelne Interaktionsmuster zu erkennen. Insbesondere komplexere Muster lassen sich durch ein Regelsystem nur schwer erfassen. Der hierarchische Ansatz kann zudem Verläufe jeglicher Länge untersuchen.

Decroos, Haaren und Davis (2018) bestimmen taktische Interaktionsmuster im Fußball basierend auf Ereignisdaten in einem fünfstufigen Prozess. Ein Ereignis setzt sich laut der Autoren zusammen aus einem Zeitstempel, den x - und y -Koordinaten, der Aktion, den am Spiel Beteiligten und optionalen zusätzlichen Informationen. Die Abfolge der Ereignisse über das gesamte Spiel wird in einzelne Phasen zerlegt. Eine neue Phase startet mit einer Spielunterbrechung über mindestens 10 Sekunden bzw. über einen Wechsel des Ballbesitzes. Die Phasen werden über eine *Hierarchische Agglomerative Clusteranalyse* gruppiert. Abhängig von den enthaltenen Torschüssen werden die Cluster in eine Rangfolge gebracht. Die *CM-SPADE* Methode hilft bei der Ermittlung häufiger Sequenzen innerhalb jeder Gruppe. Im letzten Schritt werden über eine Bewertungsfunktion die häufigsten Sequenzen gemäß ihrer Relevanz für den Trainierenden geordnet. Das entwickelte Verfahren wenden Decroos et al. auf die Ereignisdaten der *English Premier League* Saison 2015-2016 an. Die Autoren finden Unterschiede im taktischen Verhalten der Mannschaften. Der wesentliche Unterschied des hierarchischen Verfahrens im Vergleich zur Methode von Decroos et al. besteht in der Verwendung von Positionsdaten anstelle von Ereignisdaten.

In mehreren Veröffentlichungen werden Bewegungsmuster untersucht. Verwendete Verfahren lassen sich an Hand ihrer räumlichen und zeitlichen Invarianz bei der Ähnlichkeitsberechnung unterscheiden. Die von Feuerhake (2016) entwickelte Methode erlaubt die translations- als auch rotationsinvariante Erkennung von Bewegungsmustern aus Positionsdaten. Beernaerts, Baets, Lenoir und de Weghe (2020) stellen ein Verfahren dar, das zu gegebenen Referenz-Trajektorien ähnliche Trajektorien bestimmt. In räumlicher Dimension können sich die Trajektorien bezüglich der Position auf dem Spielfeld, der Rotation und der Skalierung unterscheiden (Translations-, Rotations- und Skaleninvarianz). In zeitlicher Dimension können die Trajektorien in der Dauer sowie in der Ausführungsgeschwindigkeit variieren. Für die Klassifikation taktischer Interaktionsmuster ist beispielsweise die Position der Trajektorie auf dem Spielfeld entscheidend. Die in dieser Arbeit entwickelte Hierarchie von *DyCoNs* zeichnet sich dadurch aus, dass sie keine räumlichen Invarianzen zulässt.

Hassan, Ramadan, Schrapf und Tilp (2016) bewerten den Erfolg eines Taktik-Trainings im Handball mit Hilfe *Selbstorganisierender Karten*. 723 offensive Spielsituationen der *EHF EURO-Men-18 Championship 2012* werden zunächst über eine *Selbstorganisierende Karte* in 27 typische Verläufe gruppiert. Aus den 12 häufigsten Verläufen leiten die Autoren ein Taktik-Training ab. In einer sechswöchigen Trainingsphase werden die Muster durch eine neue Mannschaft erlernt. Der Erfolg des Taktik-Trainings wird im Anschluss wieder über eine *Selbstorganisierende Karte* bestimmt. Im Ergebnis können 58% der Spielsituationen einem der Verläufe zugeordnet werden.

Schrapf, Alsaied und Tilp (2017) untersuchen taktische Interaktionsmuster im Handball. Die Forschungsgruppe bildet einen Datensatz aus den letzten fünf Pässen und dem abschließenden Wurf auf das Tor

und stellen damit den Angriff dar. Die Positionen der Verteidigenden während einer Angriffsphase werden jeweils in einem eigenen Datensatz erfasst. Anschließend wird jeweils ein *DyCoN* mit den Angriffsdaten und den Verteidigungsdaten trainiert. Es ergeben sich 25 offensive und 13 defensive Muster. Die am häufigsten auftretenden Kombinationen aus einem offensiven und einem defensiven Muster werden im Folgenden bezüglich ihrer Effizienz untersucht. Als Effizienzkriterien dienen dabei die Trefferrate, der Abstand zum nächsten defensiven Spielenden sowie der Abstand zum Tor. Insgesamt ermöglicht das Verfahren die Analyse und Bewertung taktischer Interaktionsmuster im Handball. Das Verfahren verwendet keine kategorisierten Daten und ist daher dem unüberwachten Lernen zuzuschreiben.

In einer Übersicht stellen Herold et al. (2019) verschiedene Verfahren des maschinellen Lernens vor, die das Verständnis taktischer Interaktionsprozesse erweitern. Die Untersuchungen werden nach verschiedenen Kriterien eingeteilt: überwachtes oder unüberwachtes Lernen, angewendete Methode, Ereignis- gegenüber Positionsdaten und ob die Veröffentlichung eine Reproduktion der Ergebnisse zulässt. Die Verfahren auf Basis von Positionsdaten werden weiter unterteilt in Verfahren zur Erkennung und Klassifizierung von Pässen, Verfahren die das Passverhalten des Teams untersuchen und Verfahren die Interaktionsprozesse bezüglich ihrer räumlich-zeitlichen Dimension und dem Torerfolg analysieren. Die in Kapitel 3 vorgestellten Untersuchungen werden in der letzten Kategorie verortet. Andere Publikationen dieser Kategorie haben allesamt eine zu der vorliegenden Arbeit unterschiedliche Zielsetzung, beispielsweise werden Pässe hinsichtlich ihres Risikos bewertet.

Goes et al. (2020) unterscheiden in ihrer Übersicht zwischen Verfahren aus dem Bereich der Sportwissenschaft und aus dem Bereich der Informatik. In den Veröffentlichungen beider Fachbereiche spiegeln sich die unterschiedlichen Forschungsparadigmen wider. In sportwissenschaftlichen

Untersuchungen werden überwiegend Hypothesen abgeleitet und experimentell untersucht. Dagegen werden in den informatischen Publikationen hauptsächlich neue Methoden entwickelt und evaluiert. [Goes et al.](#) schlagen ein iteratives Vorgehen vor, um die unterschiedlichen Ansätze erfolgreich zusammenzubringen. Sportwissenschaftliche Untersuchungen können zur Überprüfung von Hypothesen neue informatische Methoden verwenden. Gleichzeitig können sportwissenschaftliche Fragestellungen zur Entwicklung neuer informatischer Methoden genutzt werden. Die in der vorliegenden Arbeit entwickelte Hierarchie von *DyCoNs* beispielsweise wurde zur automatischen Erkennung von taktischen Interaktionsmustern im Fußball entwickelt. Gleichwohl stellt dies keine typisch sportwissenschaftliche Untersuchung im obigen Sinne dar, da hier keine Hypothese überprüft wird.

5.2 Implikationen für die Wissenschaft

Aus sportinformatischer Sicht konnte gezeigt werden, dass es grundsätzlich möglich ist taktische Interaktionsmuster in Sportspielen mit Hilfe *Selbstorganisierender Karten* zu klassifizieren. Die Abbildung wird aus kategorisierten Daten erlernt, die eine Fachkraft im Vorfeld erstellt hat. Die in den Neuronen enthaltene Information auf den verschiedenen Ebenen kann für die Beantwortung sportwissenschaftlicher Fragestellungen genutzt werden. Auf der dritten Ebene kann beispielsweise untersucht werden inwiefern sich Spieleröffnungen verschiedener Mannschaften voneinander unterscheiden. Auf der zweiten Ebene bilden die Neuronen die typischen Bewegungsmuster der Konstellation ab. Die Verteilung der Bewegungsmuster über die Neuronen kann Aufschluss darüber geben, welche Eigenschaften eine taktische Interaktion wie die Spieleröffnung charakterisieren.

Aus informatischer Sicht lassen sich vielversprechende Ansätze ausmachen. Die Hierarchie von *DyCoNs* nutzt zunächst ein unüberwachtes Lernverfahren um geeignete Abstraktionen der Daten zu bilden. In einer zweiten überwachten Lernphase wird dann die Klassifikation erlernt. Wie in Goes et al. (2020) dargestellt, werden im Vergleich zu anderen Untersuchungen keine Features im Vorfeld festgelegt. Die *Selbstorganisierende Karte* findet selbstständig geeignete Repräsentationen. Insbesondere für eine Klassifikation ist die Auswahl geeigneter Features entscheidend. Eine Klassifizierung von Spieleröffnungen an Hand des Schwerpunkts aller Spielenden beispielsweise würde vermutlich keine guten Ergebnisse liefern. Betrachtet man jedes Detail einer Interaktion, so ist jede Situation einmalig. Das andere Extrem erhält man, wenn man jedes Detail vernachlässigt. Abhängig von der Fragestellung müssen geeignete Abstraktionen gefunden werden, um passende Aussagen treffen zu können. Die verschiedenen Ebenen der Hierarchie bilden selbstständig Abstraktionen aufsteigender Komplexität ab. Dieses Prinzip ähnelt Verfahren aus dem Bereich des *Deep Learnings* (Schulz und Behnke (2012)). In einem unüberwachten Schritt werden Features aufsteigender Komplexität berechnet, und anschließend wird eine Klassifizierung durchgeführt. Die Komplexität entsprechender Lernverfahren hat einen breiten Einsatz von *Deep Learning* Verfahren in der Vergangenheit verhindert. Erst die gesteigerte Rechenleistung der letzten Jahre ermöglicht eine praktikable Berechnung. Zusammen mit der sehr guten Klassifikationsleistung hat das zu einem starken Interesse an *Deep Learning* Verfahren geführt.

5.3 Implikationen für die Praxis

Durch das Erlernen taktischer Interaktionsmuster aus kategorisierten Beispielen wird das Fachwissen eines Trainierenden in einen Algorithmus überführt. Dem Trainierenden kann damit ein Werkzeug zur Unterstützung seiner Arbeit an die Hand gegeben werden. Statt mehrere

Stunden Videomaterial zu sichten, können interessante Interaktionsmuster in Minuten automatisch gefunden werden. Im Idealfall kann jeder Trainierende die ihn interessierenden Interaktionsmuster selbst bestimmen. Dazu weist der Trainierende im Vorfeld Spielszenen verschiedenen Kategorien zu. Im Anschluss erlernt das hierarchische Verfahren die Muster. Für den Praxiseinsatz müsste die Hierarchie von *DyCoNs* komplett automatisiert werden. Insbesondere die für das Training und die Clusteranalyse benötigten Parameter müssten durch ein automatisiertes Verfahren bestimmt werden. Schwierigkeiten könnten dann noch auftreten, wenn etwa die Menge der kategorisierten Interaktionsmuster zu klein ist, um die Varianzen in dem Muster abzubilden. Solche Aufgaben dürften sich nur schwer automatisch lösen lassen. In diesen Fällen könnte es zum Aufbau entsprechender Kenntnisse hilfreich sein, Verfahren des maschinellen Lernens in die Ausbildung von Spielanalysierenden aufzunehmen.

Ein weiterer häufig genannter Kritikpunkt im Zusammenhang mit *Künstlichen Neuronalen Netzen* ist die Nachvollziehbarkeit der Ergebnisse. Wie die Hierarchie von *DyCoNs* zu einem Klassifikationsergebnis kommt, ist von außen nicht unmittelbar nachvollziehbar. Erst eine genaue Kenntnis der einzelnen Berechnungsschritte zusammen mit den in den Neuronen enthaltenen Informationen ermöglicht eine Nachvollziehbarkeit einzelner Ergebnisse.

Mit dem hierarchischen Verfahren wird eine neue Methode bereitgestellt, um taktische Interaktionsmuster im Sportspiel zu untersuchen. Handlungsanweisungen für ein Training oder den Wettkampf ergeben sich daraus nicht automatisch. Goes et al. (2020) schlagen vor, neue informativische Methoden zur experimentellen Überprüfung von Hypothesen einzusetzen. Erst durch die Kombination der beiden Forschungsparadigmen können neue Erkenntnisse gewonnen werden. Rein und Memmert (2016) weisen jedoch darauf hin, dass das Fehlen einer übergreifenden Theorie des Sportspiels einen breiten Einsatz neuerer Methoden verhindert.

Kapitel 6

Abstracts

Im Folgenden wird die vorliegende Arbeit in Deutsch und Englisch zusammengefasst.

6.1 Zusammenfassung

In der vorliegenden Arbeit wird ein Verfahren zur Klassifizierung von mannschaftstaktischen Interaktionsprozessen mit Hilfe *Selbstorganisierender Karten* entwickelt. Die Klassifizierung erlernt die unterschiedlichen Muster aus kategorisierten Beispielen. Das entwickelte Verfahren kann die Bewertung durch Leistungsindikatoren unterstützen oder als Basis eines neuen Indikators dienen. Zunächst wird der aktuelle Forschungsstand aus sportwissenschaftlicher, informatischer sowie sportinformatischer Sicht dargestellt. Im Anschluss werden die *Selbstorganisierende Karte* (SOM) nach Kohonen (1982), das *Dynamically Controlled Network* (DyCoN) nach Perl (2001b) sowie die *Merge Self-Organizing Map* (MSOM) nach Strickert und Hammer (2003) näher beschrieben. Als erstes Teilziel wird der komplexe mannschaftstaktische Interaktionsprozess im Sportspiel mit Hilfe *Selbstorganisierender Karten* modelliert. Dabei muss die Dynamik des Interaktionsprozesses erhalten bleiben. Ausgehend von der ersten Untersuchung wird eine Hierarchie von DyCoN Netzen entworfen, um taktische Interaktionsmuster im Fußball zu klassifizieren. Anhand der Positionsdaten des WM Finales 2006 wird der neue Ansatz evaluiert. Schließlich wird das entwickelte Verfahren auf ein Basketballspiel übertragen. Zudem wird mit dem MSOM ein alternatives Verfahren zur Klassifizierung angewendet. Die Ergebnisse der beiden Verfahren werden abschließend verglichen.

Schlagworte: Analyse, Bewegungsmuster, Fußball, Fußballspiel, Selbstorganisierende Karte, Spielanalyse, Taktik, Mannschaft

6.2 Summary

In the present work a method for classification of team tactical interaction processes based on *Self-organizing maps* is developed. The classification learns the different patterns from examples that an expert has categorized in advance. The developed procedure can support the evaluation with performance indicators or serve as the basis for a new indicator. First the current state of research is presented from the point of view of sports science, informatics and sport informatics. Afterwards the *Self-organizing map* (SOM) developed by Kohonen (1982), the *Dynamically Controlled Network* (DyCoN) developed by Perl (2001b) and the *Merge Self-organizing map* (MSOM) developed by Strickert und Hammer (2003) are introduced. As first subgoal the complex team tactical interaction process in sport games is modeled with *Self-organizing maps*. The dynamics of the interaction process must be preserved by the model. Based on the results of the previous work a hierarchy of DyCoNs is developed in order to classify tactical interaction patterns in soccer. An evaluation of the new approach is carried out on the position data of the 2006 World Cup Finals. Finally the developed method is applied to a basketball game. In addition the MSOM is adapted as an alternative method for classification. The results of the two methods are finally compared.

Keywords: analysis, movement patterns, soccer, self-organizing map, game analysis, tactics, team

Literaturverzeichnis

- Alt, H. & Godau, M. (1995). Computing the fréchet distance between two polygonal curves. *International Journal of Computational Geometry & Applications*, 05 (01n02), 75-91. Zugriff auf <https://doi.org/10.1142/S0218195995000064>
- Atluri, G., Karpatne, A. & Kumar, V. (2018). Spatio-temporal data mining: A survey of problems and methods. *ACM Comput. Surv.*, 51 (4). Zugriff auf <https://doi.org/10.1145/3161602>
- Beernaerts, J., Baets, B. D., Lenoir, M. & de Weghe, N. V. (2020). Spatial movement pattern recognition in soccer based on relative player movements. *PLoS ONE*, 15, e0227746. Zugriff auf <https://doi.org/10.1371/journal.pone.0227746>
- Bialkowski, A., Lucey, P., Carr, P., Yue, Y., Sridharan, S. & Matthews, I. (2014). Large-scale analysis of soccer matches using spatiotemporal tracking data. In *2014 IEEE International Conference on Data Mining* (S. 725-730).
- Bialkowski, A., Lucey, P., Carr, P., Yue, Y., Sridharan, S. & Matthews, I. (2015). Identifying team style in soccer using formations learned from spatiotemporal tracking data. In *IEEE International Conference on Data Mining Workshops, ICDMW* (Bd. 2015, S. 9-14). Zugriff auf <https://doi.org/10.1109/ICDMW.2014.167>
- Carling, C., Wright, C., Nelson, L. J. & Bradley, P. S. (2014). Comment on 'Performance analysis in football: A critical review and implications for future research'. *Journal of Sports Sciences*, 32, 2-7. Zugriff auf <https://doi.org/10.1080/02640414.2013.807352>
- Decroos, T., Haaren, J. V. & Davis, J. (2018). Automatic Discovery of Tactics in Spatio-Temporal Soccer Match Data. In *Proceedings of the 24th International Conference on Knowledge Discovery and Data Mining* (S. 223-232). ACM. Zu-

- Zugriff auf <https://doi.org/10.1145/3219819.3219832>
- Espadoto, M., Martins, R. M., Kerren, A., Hirata, N. S. T. & Telea, A. C. (2019). Towards a quantitative survey of dimension reduction techniques. *IEEE Transactions on Visualization and Computer Graphics*. Zugriff auf <https://doi.org/10.1109/TVCG.2019.2944182>
- Fakhrzari, A. & Vakilzadian, H. (2017). A survey on time series data mining. In *2017 IEEE International Conference on Electro Information Technology (EIT)* (S. 476-481). Zugriff auf <https://ieeexplore.ieee.org/document/8053409>
- Feuerhake, U. (2016). Recognition of repetitive movement patterns - the case of football analysis. *ISPRS International Journal of Geo-Information*, 5, 208. Zugriff auf <https://doi.org/10.3390/ijgi5110208>
- Franks, I. M. & Miller, G. (1991). Training coaches to observe and remember. *Journal of Sports Sciences*, 9 (3), 285-297. Zugriff auf <https://doi.org/10.1080/02640419108729890>
- Fréchet, M. M. (1906). Sur quelques points du calcul fonctionnel. *Rendiconti del Circolo Matematico di Palermo*, 22 (1), 1-72. Zugriff auf <https://doi.org/10.1007/BF03018603>
- Goes, F., Meerhoff, L., Bueno, M., Rodrigues, D., Moura, F., Brink, M., ... Lemmink, K. (2020). Unlocking the potential of big data to support tactical performance analysis in professional soccer: A systematic review. *European Journal of Sport Science*, 1-16. Zugriff auf <https://doi.org/10.1080/17461391.2020.1747552>
- Grunz, A., Memmert, D. & Perl, J. (2009). Analysis and simulation of actions in games by means of special self-organizing maps. *Int. J. Comput. Sci. Sport*, 8 (1).
- Grunz, A., Memmert, D. & Perl, J. (2012). Tactical pattern recognition in soccer games by means of special self-organizing maps. *Human Movement Science*, 31 (2), 334-343. Zugriff auf <https://doi.org/10.1016/j.humov.2011.02.008>
- Guimarães, G., Lobo, V. & Moura-Pires, F. (2003). A taxonomy of self-organizing maps for temporal sequence processing. *Intelligent Data Analysis*, 7, 269-290. Zugriff auf <https://doi.org/10.3233/IDA-2003-7402>

- Hassan, A., Ramadan, W., Schrapf, N. & Tilp, M. (2016). Evaluation of tactical training in team handball by means of artificial neural networks. *Journal of Sports Sciences*, May. Zugriff auf <https://doi.org/10.1080/02640414.2016.1183804>
- Hausdorff, F. (1914). *Grundzüge der Mengenlehre*. Leipzig: Von Veit.
- Herold, M., Goes, F., Nopp, S., Bauer, P., Thompson, C. & Meyer, T. (2019). Machine learning in men's professional football: Current applications and future directions for improving attacking play. *International Journal of Sports Science & Coaching*, 14 (6), 798–817. Zugriff auf <https://doi.org/10.1177/1747954119879350>
- Hughes, M. & Bartlett, R. (2002). The use of performance indicators in performance analysis. *Journal of Sports Sciences*, 20, 739-54. Zugriff auf <https://doi.org/10.1080/026404102320675602>
- Iqbal, M. & Yan, Z. (2015). Supervised machine learning approaches: A survey. *International Journal of Soft Computing*, 5, 946-952. Zugriff auf <https://doi.org/10.21917/ijsc.2015.0133>
- Jäger, J., Perl, J. & Schöllhorn, W. (2007). Analysis of players' configurations by means of artificial neural networks. *International Journal of Performance Analysis in Sport*, 7, 90-105. Zugriff auf <https://doi.org/10.1080/24748668.2007.11868413>
- Kempe, M., Grunz, A. & Memmert, D. (2014). Detecting tactical patterns in basketball: Comparison of merge self-organising maps and dynamic controlled neural networks. *European Journal of Sport Science*, 15 (4), 249–255. Zugriff auf <https://doi.org/10.1080/17461391.2014.933882>
- Kisilevich, S., Mansmann, F., Nanni, M. & Rinzivillo, S. (2009). Spatio-temporal clustering. In O. Maimon & L. Rokach (Hrsg.), *Data mining and knowledge discovery handbook* (S. 855-874). Boston, MA: Springer US. Zugriff auf https://doi.org/10.1007/978-0-387-09823-4_44
- Knauf, K., Memmert, D. & Brefeld, U. (2016). Spatio-temporal convolution kernels. *Machine Learning*, 102 (2), 247–273. Zugriff auf <https://doi.org/10.1007/s10994-015-5520-1>
- Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43 (1), 59-69. Zugriff auf <https://doi.org/>

[10.1007/BF00337288](https://doi.org/10.1007/BF00337288)

- Kohonen, T. (1990). The self-organizing map. *Proceedings of the IEEE*, 78 (9), 1464-1480. Zugriff auf <https://doi.org/10.1109/5.58325>
- Kohonen, T. (2001). *Self-organizing maps, third edition*. Springer. Zugriff auf <https://doi.org/10.1007/978-3-642-56927-2>
- Kohonen, T. (2013). Essentials of the self-organizing map. *Neural Networks*, 37, 52 - 65. Zugriff auf <http://www.sciencedirect.com/science/article/pii/S0893608012002596> (Twenty-fifth Anniversary Commemorative Issue)
- Lames, M. & Perl, J. (1999). Identifikation von Ballwechseltypen mit Neuronalen Netzen. In Roth, K., Pauer, Th., Reichle, K. (Hrsg.) *Dimensionen und Visionen des Sports*. Hamburg, 133.
- Mackenzie, R. & Cushion, C. (2013). Performance analysis in football: A critical review and implications for future research. *Journal of Sports Sciences*, 31 (6), 639–676. Zugriff auf <https://doi.org/10.1080/02640414.2012.746720>
- Memmert, D., Lemmink, K. A. P. M. & Sampaio, J. (2017). Current approaches to tactical performance analyses in soccer using position data. *Sports Medicine*, 47 (1), 1–10. Zugriff auf <https://doi.org/10.1007/s40279-016-0562-5>
- Memmert, D., Raabe, D., Knyazev, A., Franzen, A., Zekas, L., Rein, R., ... Weber, H. (2016). Big Data im Profi-Fußball: Analyse von Positionsdaten der Fußball-Bundesliga mit neuen innovativen Key Performance Indikatoren. *Leistungssport*, 46 (5), 21–26. Zugriff auf https://fis.dshs-koeln.de/portal/files/2743866/Memmert_2016_Big_Data_im_Profifussball.pdf
- Mester, J. & Perl, J. (2000). Grenzen der Anpassungs- und Leistungsfähigkeit des Menschen aus systemischer Sicht: Zeitreihenanalysen und ein informatisches Metamodell zur Untersuchung physiologischer Adaptationsprozesse. *Leistungssport*, 30, 43-51.
- Parmar, N., James, N., Hearne, G. & Jones, B. (2018). Using principal component analysis to develop performance indicators in professional rugby league. *International Journal of Performance Analysis in Sport*, 18, 1-12. Zugriff auf

<https://doi.org/10.1080/24748668.2018.1528525>

- Pearson, K. (1901). LIII. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2 (11), 559-572. Zugriff auf <https://doi.org/10.1080/14786440109462720>
- Perl, J. (2001a). Artificial neural networks in sports: New concepts and approaches. *International Journal of Performance Analysis in Sport*, 1, 106-121. Zugriff auf <https://doi.org/10.1080/24748668.2001.11868253>
- Perl, J. (2001b). DyCoN: Ein dynamisch gesteuertes Neuronales Netz zur Modellierung und Analyse von Prozessen im Sport. In J. Perl (Hrsg.), *Sport & Informatik VIII* (S. 85-98). Strauß, Köln.
- Perl, J. (2002a). Adaptation, antagonism, and system dynamics. In G. Ghent, D. Kluka & D. Jones (eds.), *Perspectives - The Multidisciplinary Series of Physical Education and Sport Science* (Hrsg.), (Bd. 4, S. 105-125). Meyer & Meyer Sport, Oxford.
- Perl, J. (2002b). Game analysis and control by means of continuously learning networks. *International Journal of Performance Analysis in Sport*, 2, 21-35. Zugriff auf <https://doi.org/10.1080/24748668.2002.11868258>
- Perl, J. (2008). Modelling. In Dabnichki, P. and Baca, A., *Computers in Sport* (Hrsg.), (S. 121-160). WIT Press.
- Perl, J. & Memmert, D. (2011). Net-based game analysis by means of the software tool SOCCER. *International journal of computer science in sport*, 10, 77-84. (Wien)
- Pfeiffer, M. & Perl, J. (2006). Analysis of tactical structures in team handball by means of artificial neural networks. *International Journal of Computer Science in Sport*, 5, 4-14.
- Rein, R. & Memmert, D. (2016). Big data and tactical analysis in elite soccer: future challenges and opportunities for sports science. *Springerplus*, 5 (1), 1410. Zugriff auf <https://doi.org/10.1186/s40064-016-3108-2>
- Rui Xu & Wunsch, D. (2005). Survey of clustering algorithms. *IEEE Transactions on Neural Networks*, 16 (3), 645-678. Zugriff auf <https://doi.org/10.1109/TNN.2005.845141>
- Sarmiento, H., Marcelino, R., Anguera, M. T., Campaniço, J., Matos, N. & Leitão,

- J. C. (2014). Match analysis in football: a systematic review. *Journal of Sports Sciences*, 32, 1831 - 1843. Zugriff auf <https://doi.org/10.1080/02640414.2014.898852>
- Schmidt, A. (2012). Movement pattern recognition in basketball free-throw shooting. *Human movement science*, 31 (2), 360–382. Zugriff auf <https://doi.org/10.1016/j.humov.2011.01.003>
- Schnabel, G., Harre, H.-D. & Krug, J. (2014). *Trainingslehre - Trainingswissenschaft*. Meyer + Meyer Fachverlag.
- Schrapf, N., Alsaied, S. & Tilp, M. (2017). Tactical interaction of offensive and defensive teams in team handball analysed by artificial neural networks. *Mathematical and Computer Modelling of Dynamical Systems*, 23, 363-371. Zugriff auf <https://doi.org/10.1080/13873954.2017.1336733>
- Schulz, H. & Behnke, S. (2012). Deep learning. *KI - Künstliche Intelligenz*, 26 (4), 357-363. Zugriff auf <https://doi.org/10.1007/s13218-012-0198-z>
- Sternberg, R. J. & Lubart, T. I. (1998). The concept of creativity: Prospects and paradigms. In R. J. Sternberg (Hrsg.), *Handbook of creativity*. Cambridge University Press. Zugriff auf <https://doi.org/10.1017/CBO9780511807916.003>
- Strickert, M. & Hammer, B. (2003). *Neural gas for sequences*.
- Strickert, M. & Hammer, B. (2005). Merge som for temporal data. *Neurocomputing*, 64, 39–71. Zugriff auf <https://doi.org/10.1016/j.neucom.2004.11.014>
- Talbot, S. A. & Marshall, W. H. (1941). Physiological studies on neural mechanisms of visual localization and discrimination*. *American Journal of Ophthalmology*, 24 (11), 1255 - 1264. Zugriff auf [https://doi.org/10.1016/S0002-9394\(41\)91363-6](https://doi.org/10.1016/S0002-9394(41)91363-6)
- Tan, P.-N., Steinbach, M., Karpatne, A. & Kumar, V. (2018). *Introduction to data mining (2nd edition)* (2nd Aufl.). Pearson.
- Torgerson, W. S. (1958). *Theory and methods of scaling*. New York: John Wiley and Sons.
- Van Der Maaten, L., Postma, E. & Van den Herik, J. (2009). Dimensionality reduction: a comparative review. *Journal of Machine Learning Research*, 10, 66-71. Zugriff auf <http://citeseerx.ist.psu.edu/viewdoc/summary>

?doi=10.1.1.112.5472

Wang, K. & Zemel, R. (2016). Classifying NBA offensive plays using neural networks. In *MIT sloan sports analytics conference*. Zugriff auf <http://www.sloansportsconference.com/wp-content/uploads/2016/02/1536-Classifying-NBA-Offensive-Plays-Using-Neural-Networks.pdf>

Willshaw, D. J. & von der Malsburg, C. (1976). How patterned neural connections can be set up by self-organization. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 194, 431 - 445. Zugriff auf <https://doi.org/10.1098/rspb.1976.0087>