

Alfred O. Effenberg (Hrsg.)

Bewegungs-Sonification und Musteranalyse im Sport

–

Sportwissenschaft trifft Informatik



Deutsche Sporthochschule Köln



universität**bonn**

Alfred O. Effenberg
(Herausgeber)

Bewegungs-Sonification und Musteranalyse im Sport

–

Sportwissenschaft trifft Informatik

Bibliografische Information Der Deutschen Bibliothek

Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.ddb.de> abrufbar.

1. Aufl. - Göttingen : Cuvillier, 2006

ISBN 3-86727-017-5

Anschrift des Herausgebers

PD Dr. Alfred Effenberg

Deutsche Sporthochschule Köln

c/o Institut für Sportwissenschaft und Sport

Universität Bonn

Nachtigallenweg 86

53127 Bonn

Email: effenberg@uni-bonn.de

Beitragsband zum Forschungssymposium am 15. Juli 2005 an der Deutschen Sporthochschule in Köln.

Redaktionelle Bearbeitung und Layout: Manuela Motschall

© CUVILLIER VERLAG, Göttingen 2006

Nonnenstieg 8, 37075 Göttingen

Telefon: 0551-54724-0

Telefax: 0551-54724-21

www.cuvillier.de

Alle Rechte vorbehalten. Ohne ausdrückliche Genehmigung des Verlages ist es nicht gestattet, das Buch oder Teile daraus auf fotomechanischem Weg (Fotokopie, Mikrokopie) zu vervielfältigen.

1. Auflage, 2006

Gedruckt auf säurefreiem Papier

ISBN 3-86727-017-5

GRÜßWORT DURCH DEN REKTOR DER DEUTSCHEN SPORTHOCHSCHULE KÖLN,
PROF. MULT. DR. WALTER TOKARSKI

Forschungs-Symposium "Bewegungs-Sonification und Musteranalyse im Sport", 15. Juli 2005

Liebe Teilnehmer des Forschungs-Symposiums "Bewegungs-Sonification und Musteranalyse im Sport",

im Namen der Deutschen Sporthochschule Köln heiße ich Sie alle sehr herzlich als Gäste der größten Sportuniversität Europas willkommen.

Auf den ersten Blick mag es vielleicht überraschen, dass in dieser Veranstaltung die Rolle des Gehörs für die Regulation und das Erlernen von Sportbewegungen thematisiert wird. Wie Sprache im Kontext mit Hören und vice versa funktioniert, das ist bereits umfassend untersucht worden. Wie allerdings akustische Information das Bewegungsverhalten beeinflusst, oder wie die komplexe Verbindung aller Sinne beim Zustandekommen von Bewegung funktioniert, das ist gegenwärtig in weiten Teilen noch ungeklärt.

An dieser Stelle mit einer solchen Veranstaltung einmal anzusetzen erscheint mir daher perspektivreich: Als Musikliebhaber haben mich die Fähigkeiten meines Gehörs bereits wiederholt in Erstaunen versetzt, etwa über das mühelose Erinnern einer zuvor erst wenige Male gehörten Melodie oder auch über die hohe zeitliche Differenzierungsfähigkeit, die uns das Raumrichtungshören überhaupt erst ermöglicht.

Die zusätzliche Vertonung von Bewegungsmustern soll zukünftig zu präziseren und effizienteren Methoden beim Bewegungslernen führen. Die Notwendigkeit interdisziplinärer Kooperationen für ein solches Unterfangen deutet sich in der zweiten Titelzeile zu diesem Symposium bereits an: Nicht sportwissenschaftliche Fragestellungen allein bestimmen das Veranstaltungsprogramm. Verschiedene Beiträge kommen aus der Informatik der Universität Bonn und der Universität Mainz, ergänzt durch Arbeiten aus den Neurowissenschaften: Die Neurologische Klinik des Universitätsklinikums Aachen ist hier genauso vertreten wie die Neuroinformatik der Universität Bielefeld und die Kognitive Neurowissenschaft der Universität Bochum. Sportwissenschaftliche Arbeiten von Kollegen von den Universitäten in Heidelberg, Mainz, Bonn und der Deutschen Sporthochschule Köln bilden den inhaltlichen Rahmen für das Symposiumsprogramm.

Zwischen Sportwissenschaft und Informatik der Universität Bonn hat es bereits eine mehrjährige erfolgreiche Kooperation gegeben und es ist zukünftig eine weitere Zusammenarbeit im Rahmen einer größeren Forschungsinitiative der Informatik der Universität Bonn vorgesehen.

Die Universität Bonn und die Deutsche Sporthochschule Köln verbindet seit 2003 zudem die enge Kooperation beim Studiengang "Sportwissenschaft mit dem Schwerpunkt Alterssport", der seit drei Semestern als Zertifikationsstudiengang existiert und ab 2007/08 in den Master-Studiengang "Bewegung, Sport und Altern" an der DSHS überführt werden soll.

Ich freue mich in besonderem Maße über die intensive interdisziplinäre Kooperation zwischen Wissenschaftlern der Universität Bonn und der Deutschen Sporthochschule Köln und erhoffe mir, dass Resultate dieser Forschungsk Kooperation gerade auch in Form praxisrelevanter Erkenntnisse erzielt werden mögen: Ich denke dabei an audiovisuelle Trainingsformen im Bereich des Techniktrainings oder auch an neuartige Interventionsformen beim Sport mit Sondergruppen und in der Rehabilitation.

Bevor ich schließe, danke ich Ihnen, Herr Kollege Effenberg, und ihren Mitarbeitern für die Organisation dieser Veranstaltung. Ich wünsche dem Symposium nun einen angeregten erkenntnisreichen wissenschaftlichen Austausch und hoffe, dass Sie an der Deutschen Sporthochschule Köln nicht nur einen interessanten, sondern auch einen schönen Tag in Köln verbringen.

Bewegungs-Sonification und Musteranalyse im Sport – Sportwissenschaft trifft Informatik

Beim motorischen Lernen in Sport und Rehabilitation dominieren visuelle und verbale Informationen nicht nur in der Bewegungspraxis, sie bilden den Schwerpunkt ebenso in der Lerntheorie und der Motorikforschung. Doch können zentrale lernrelevante Informationen neben dem Auge auch über das Ohr gewonnen werden, wobei in unserem Zusammenhang ausschließlich nonverbale akustische Informationen behandelt werden. Und während neben der Videotechnik die Bewegungsvisualisierung auf der Grundlage von Bewegungsdaten bereits umfassend in der sportlichen Trainingspraxis etabliert ist, gilt dies für die akustische Umsetzung von Bewegungsdaten – die Bewegungs-Sonification – noch nicht in vergleichbarem Maß.

Der Forschungsansatz zur Bewegungs-Sonification eröffnet neuartige Perspektiven beim motorischen Lernen in Sport und Rehabilitation: mit der ergänzenden Nutzung auditiver Wahrnehmungsmechanismen ebenso wie über die gezielte Einbindung audiovisueller Integrationsinstanzen, wenn die Bewegungs-Sonification mit visuellen Bewegungsinformationen im Instruktions- wie auch im Feedback-Zusammenhang kombiniert wird. In diesen beiden Funktionsbereichen des Wahrnehmungssystems können Informationen generiert werden, die über den Funktionsumfang visueller Information hinausgehen. So bilden die Fragen nach Umfang, Form und zukünftigen Anwendungsbereichen der datenbasierten Bewegungsakustik in Sport und Rehabilitation den einen Schwerpunkt dieser Schrift.

Einen anderen Schwerpunkt bildet die Suche nach Bewegungsmerkmalen, die für das motorische Lernen eine hervorgehobene Bedeutung besitzen. Dieser Gegenstandsbereich wird aus unterschiedlichen Perspektiven wie auch aus verschiedenen Richtungen behandelt. Eine primär endogene Perspektive, und zwar mit einem Bezug auf Wahrnehmung und Informationsverarbeitung im lernenden und sich entwickelnden System, nimmt der Beitrag von Mechling ein, mit dem die grundlegende Bedeutung mehr- bzw. amodaler Information mit Redundanzanteilen für den Entwicklungs- und Lernzusammenhang betont wird. Mit dem Wahrnehmungs- und Informationsprozess zwischen einem agierenden und einem beobachtenden System erfolgt mit dem Beitrag von Westhoff und Troje ein Perspektivwechsel: Hier wird die Funktionsweise spezifischer visueller Mechanismen bei der Beobachtung biologischer Bewegungen thematisiert, indem die Bedeutung bestimmter Stimulusmerkmale wie Größe, Schrittfrequenz, Blickwinkel etc. auf die Identifikationsleistung überprüft wird.

Eine exogene Perspektive nehmen die drei folgenden Beiträge aus der Informatik und der aus der Sportspielforschung ein, mit denen neue Wege bei der strukturellen Analyse von Bewegungsmerkmalen skizziert werden: Damit treten zunächst einmal wahrnehmungsseitige Bezüge in den Hintergrund, da die Analyse auf der Basis von Sensordaten menschlicher Bewegungen realisiert wird. Grundtechniken der Bewegungsanalyse auf der Grundlage von Motion Capturing Daten, z.B. zu Aspekten der inter- und intraindividuellen Variabilität von Bewegungsmustern,

werden von Clausen, Müller und Röder beschrieben. Die Autoren stellen zudem mit dem ‚Dynamic Time Warping‘ eine neue Technik zur Berechnung von Ähnlichkeitsmaßen für verschiedene Bewegungen vor. Um Vergleiche und Klassifikationen von menschlichen Bewegungen auf der Basis von Motion Capturing Daten geht es auch in dem Beitrag von Röder und Müller. Die Autoren beschreiben einen Ansatz auf der Basis von geometrischen Relationen bestimmter Körperposen bzw. deren spezifischer Abfolge, der eine sehr schnelle Suche bestimmter Bewegungen in großen Datenbanken ermöglicht. Perl zeigt darauffolgend, dass mit geeigneten dynamischen neuronalen Netzwerkmodellen (DyCoN, Dynamically Controlled Network) Lernprozessverläufe qualitativ analysiert werden können - und das bereits auf Basis weniger Messzeitpunkte und geringer Datenumfänge. Memmert überträgt den Ansatz der dynamischen neuronalen Netzwerkmodelle in die Sportspielforschung und zeigt, dass sich DyCoNs auch für die qualitative Analyse von inter- und intraindividuellen Unterschieden bei dem taktischen Verhalten im Sportspiel eignen.

Die Beiträge im dritten Block behandeln verschiedene Aspekte der Sonification an der Schnittstelle von empirischen Wirksamkeitsbelegen, sensordatenbasierter Modellierung, funktionellen Aspekten von Sonificationssystemen und einer neurophysiologisch unterstützten Theorieentwicklung. Der Beitrag von Effenberg bezieht sich auf die Wirkungsweise der Bewegungs-Sonification, skizziert Anwendungsbereiche in Sport und Rehabilitation und führt empirische Belege zur Wirksamkeit bewegungsakustischer Information bei der Wahrnehmung, Beurteilung und dem Nachvollzug von Sportbewegungen an. Weber beschreibt in dem darauffolgenden Beitrag die inneren Kräfte eines spezifischen Bewegungsmusters als einen interessanten Bezugspunkt für die Bewegungs-Sonification, die sich mit der Inversen Dynamik aus kinematischen Parametern und massedefinierten Körpermodellen näherungsweise berechnen lassen. Kurth beschreibt mit der verlaufs- und der ereignisbezogenen Sonification unterschiedliche Sonificationsarten und diskutiert verschiedene Verfahren der Audiosignalverarbeitung und Sounderzeugung wie Parametrische Modelle, sample-basierte Synthesemodelle und das Physical Modelling. Effenberg, Fehse und Mechling stellen abschließend die Entwicklung einer fMRT-Untersuchungsmethode zur Wirkungsweise der Bewegungs-Sonification vor, die konsequent an einer verhaltenspsychologischen Methode orientiert wird. Neben der Option, Ergebnisse aus beiden Untersuchungslinien aufeinander zu projizieren und die Wirkungsweise der Bewegungsakustik vor dem Hintergrund der multisensorischen Integration im ZNS zu explorieren, wird mit diesem Beitrag auch für die neurophysiologische Ergänzung der gegenwärtig kognitiv dominierten Motorikforschung votiert.

Mit dem Forschungssymposium ist es gelungen, Aspekte der Musteranalyse und der Bewegungs-Sonification vor dem Hintergrund perzeptiv-motorischer Zusammenhänge und lerntheoretischer Erkenntnisse in einer echten interdisziplinären Expertenrunde wie in inspirierender Atmosphäre zu diskutieren und ein Stück voranzubringen. Dafür möchte ich mich bei allen Teilnehmern herzlich bedanken!

Inhalt

HEINZ MECHLING

Signalparameter, motorische Kontrolle und Handlungsregulation 8

CORD WESTHOFF, NIKOLAUS F. TROJE

**Personenidentifikation anhand von biologischer Bewegung –
strukturelle und kinematische Parameter 13**

MICHAEL CLAUSEN, MEINARD MÜLLER & TIDO RÖDER

Informatische Aspekte der Bewegungsanalyse 18

TIDO RÖDER, MEINARD MÜLLER

Geometrische Relationen für die Bewegungsanalyse 24

JÜRGEN PERL

Einsatz künstlicher neuronaler Netze zur Mustererkennung im Sport 29

DANIEL MEMMERT

**Kennzeichnung inter- und intraindividuelle taktischer Unterschiede
mit Hilfe Neuronaler Netze 37**

ALFRED O. EFFENBERG

**Bewegungs-Sonification und multisensorische Integration:
Empirische Befunde und zukünftige Perspektiven 42**

ANDREAS WEBER

**Rekonstruktion dynamischer Daten aus kinematischen Daten für die
Sonification: Schwerpunkttrajektorien, Summen einwirkender Kräfte
und Inverse-Dynamik-Berechnungen 48**

FRANK KURTH

Sonifikationssysteme 52

ALFRED O. EFFENBERG, URSULA FEHSE & HEINZ MECHLING

**Zur Verarbeitung audiovisueller Bewegungsinformation im Sport:
eine fMRT-Studie 59**

Autorenverzeichnis 66

HEINZ MECHLING

Signalparameter, motorische Kontrolle und Handlungsregulation

Einleitung

Wie andere verhaltens- und handlungsorientierte Wissenschaftsdisziplinen unterliegt die Sportwissenschaft und insbesondere die Teildisziplin Bewegungswissenschaft seit langem einem unaufhaltsamen Trend zur Spezialisierung. Dieser wird als konsequente Antwort auf die zunehmende Komplexität der Fragestellungen verstanden. Für die Sportwissenschaft insgesamt, aber auch für eine handlungsorientierte Bewegungswissenschaft hat diese Entwicklung auch desintegrierende Wirkungen. Diese führen wiederum zu einer gewissen Unzufriedenheit, weil die Frage, wie die Einzelteile wieder zusammengeführt werden, nicht beantwortet wird oder auch nicht beantwortet werden kann. Trotz aller funktioneller Spezialisierungen seiner einzelnen Systembestandteile ist unser Körper als Einheit nicht nur sichtbar, sondern muss auch in seinen untrennbaren Zusammenhängen analysiert werden. Dies gilt umsomehr für die Identität der Person, denn „We humans ... have no doubt whatsoever, that it is us, and us alone, that direct the motions of our bodies. ... the most fundamental kind of consciousness, the awareness of self, springs from the ground of spontaneous self-organized activity“ (Kelso, 2002, 370).

„Selbstverständlich ist es dem Forscher völlig freigestellt, an welcher Stelle des Systems er seine Tiefbohrung ansetzt, sofern er nur weiß, wo sie in Bezug auf das ganze System gelegen ist“. Wie das Zitat von Konrad Lorenz (1978, 46) belegt sind solche Auffassungen keineswegs neu: In unterschiedlichen Wissenschaftsdisziplinen sind insbesondere in Verbindung mit der Erforschung und theoretischen Fundierung von kognitiven Vorgängen zunehmend empirische Ansätze zu finden, die einen Anspruch auf ihren Beitrag zu einem größeren theoretischen – oft noch nicht festgelegten – Zusammenhang beinhalten. Ansätze zu integrierenden Theorien sind mit den „Referential dynamics of cognition and action“ von Pressing (1999) der „Emulation theory of representation, motor control, imagery and perception“ von Grush (2003), einer „Integrated theory of mind“ von Anderson et al. (2004) oder der „Integrated science of the person“ von Mischel (2004) verbunden. Allen Newell hat über Jahrzehnte die Notwendigkeit der integrativen Theoriebildung eingefordert und selbst vorangetrieben. Für die Entwicklung komplexer Theorien, insbesondere der Kognitionstheorie, hat er vor separierten Betrachtungen gewarnt: „If a theory covers only part or component, it flirts with trouble from the start. Parts or components are ... important to understand and help to make that theory simple enough to use. But they don't remove a necessity of a theory, that provides a total picture and explains the role of the parts and why they exist.“ (Newell, 1990, 17-18). Die exakte Messung, die präzise Gestaltung und Erfassung von Signalen in empirischen Untersuchungen macht den ordnenden theoretischen Rahmen zu einer unverzichtbaren Notwendigkeit, auch um die Beiträge von Teilsystemen identifizieren, gewichten und für Prognosen nutzen zu können.

Signale - Signalparameter - Signalinventare

Für diese Betrachtung ist es bedeutsam, zwischen den Signalauffassungen der Nachrichtentechnik und der Handlungsregulation zu unterscheiden, um daraus Schlussfolgerungen für die Handlungsregulation und Handlungsansteuerung unter Hinzuziehung von Messparametern ziehen zu können.

In der Nachrichtentechnik wird unter Signal jeder Sachverhalt (Zeichen) verstanden, der bei festgelegter Codierung Informationen übermittelt. Dabei handelt es sich um wert- oder zeitkontinuierliche Funktionen. Bei deren Speicherung werden Zeitfunktionen (Spannung) in Ortsfunktionen (Physikalische Größen) umgewandelt.

Diese Signalauffassung mag für die unbelebte Natur und auch noch für psychophysikalische Betrachtungen ausreichend sein. Für den Wahrnehmungs-Handlungs-Zusammenhang und die Handlungsregulation haben Signale eine darüber hinausgehende Bedeutung. Sie sind funktional bedeutsame Reize mit handlungsnützlichen Informationen. Sie zeigen Handlungserfordernisse an, ihnen wird Bedeutung beigemessen und sie verlangen eine situationsangemessene Antwortreaktion (vgl. Hacker, 1998). Die Signalparameter als die für die Übertragung einer Information notwendigen Merkmale – deren Werte und der Werteverlauf – müssen in einen übergeordneten Zusammenhang integriert werden. Für die Handlungsregulation muss deren funktionale Bedeutsamkeit berücksichtigt werden. Signalinventare werden im Sinne von Signalverbindungen höherer Ordnung als besondere Form der Gedächtnisrepräsentation bezeichnet. Entscheidend für den Wahrnehmungs-Handlungs-Zusammenhang ist dann die Frage, wie diese bedeutsamen Reize für eine Objekt- und Zielorientierung zustande kommen und wie sie genutzt werden.

Dies bedeutet die Suche nach Präzisierungsangeboten für Zielorientierungen und Antworten auf die Frage, wie Signale einzelner Sinnes- und Informationskanäle genutzt und wie sie zu neuen Informationseinheiten integriert werden. Dies führt zu einer der Kernfragen der motorischen Kontrolle und Handlungsregulation: Wie entstehen aus Intentionen Befehle an die Muskulatur.

Handlungsregulation

Der Ausgangspunkt Zielorientierung führt eindeutig zu der Frage nach der Situation und der Veränderung der Situation durch Bewegungsaktivitäten und damit direkt zu Fragen der Wahrnehmung und Antizipation (vgl. Hoffmann, 1993). Die Wahrnehmung selbst muss daher als Grundlage und notwendige Voraussetzung für die Handlungssteuerung betrachtet werden (vgl. Gibson 1979; Rubinstein, 1977). Die Wahrnehmung bestimmt somit darüber, was und wie etwas getan werden muss. Die sensorischen Signale aus der Umwelt oder der Bewegung müssen so miteinander verbunden werden, dass Objekte und Situationen im Hinblick auf Handlungsoptionen richtig eingeschätzt werden können. Dabei können sensorische Vorgänge bzw. das daraus entstandene Wahrnehmungswissen geteilt vorliegen, oder kontrovers sein. Unser Erleben ist trotzdem immer eindeutig einheitlich. Daher ist anzunehmen, dass verschiedene interne Repräsentationen für die Handlungssteuerung und das Erleben vorliegen. Die *Wahrnehmungskonstanz* wird durch eine

Reihe von objektbezogenen Konstanzphänomenen wie *Form*, *Größe*, *Farbe*, *Helligkeit* und *Position* oder situativ durch *Funktion* und *Bedeutung* beschrieben (vgl. Walsh und Kulikowski, 1998).

Da nicht alle den Sinnesorganen vorliegenden Informationen in gleichem Ausmaße verbreitet und verarbeitet werden, spielt für weitere Überlegungen die Verbindung zwischen Wahrnehmung und Aufmerksamkeit eine bedeutsame Rolle. Die Frage danach, was selektiert wird, führt zu der Unterscheidung zwischen willkürlicher (intentional) und unwillkürlicher (nicht intentional) Aufmerksamkeitsorientierung. Darüber hinaus führt die hier gewählte Thematik zu wichtigen Unterscheidungen zwischen unimodalen oder multimodalen/crossmodalen Sinnesmodalitäten und deren Einfluss.

Die Wirkungen und Zusammenhänge uni- und bimodaler Aufmerksamkeitseffekte wurden relativ isoliert und primär unter unimodalen Gesichtspunkten untersucht. Multimodale bzw. crossmodale Aufmerksamkeitseffekte wurden von Spence und Driver (1996) beschrieben. Nach ihren Aussagen ziehen auditive Hinweise unwillkürlich visuelle und taktile Aufmerksamkeit auf sich. Taktile Hinweise führen zu unwillkürlicher visueller und auditiver Aufmerksamkeitsorientierung. Visuelle Hinweisreize lösen unwillkürlich taktile Aufmerksamkeit aus. Von der Ausrichtung auditiver Aufmerksamkeit profitiert die Verarbeitung visueller Reize. Zur Wirkung multimodaler bzw. crossmodaler Informationsverarbeitung liegen allerdings widersprüchliche Ergebnisse vor. So postulieren Driver und Spence (1998), dass die Wirkung von visuell auf auditiv nicht vorhanden, bzw. viel geringer ist als umgekehrt. Sie bestätigen eine Asymmetrie der Auslösung unwillkürlicher Aufmerksamkeit zwischen visueller und auditiver Modalität. Ward et al. (2000) beschreiben deutliche Effekte von visueller Information auf auditiv räumliche Zusammenhänge und können keine generelle Asymmetrie feststellen.

„Amodale“ Information und Handlung

Die Crux des Wahrnehmungs-Handlungsproblems haben Carello und Turvey (2004) prägnant zusammengefasst „... people are aware of the unchanging properties of objects even though their contact with them is changing constantly. ... Reliable perception implies that the changing forces and motions have to be linked by some property that does not change as they do.“ (2004, S. 25). Damit wird nicht die Existenz physikalischer Parameter und ihres Einflusses auf die Rezeptoren in Frage gestellt. Die Überführung dieser Parameter in eine reliable und valide Wahrnehmung setzt allerdings ein System voraus, das nicht mit diesen Parametern variiert. Verbunden mit der Frage, wie der sensorische Input konvergiert, müssen zeitinvariante Parameter identifiziert und beschrieben werden.

Neuere empirisch gestützte Angebote sind in der intersensorischen Redundanzhypothese (IRH) von Bahrnick und Lickliter (2002; Bahrnick et al. 2004) zu finden. Mit dem *Redundanzbegriff* gehen sie über den nachrichtentechnisch definierten Begriff hinaus und betrachten das mehrfache Vorhandensein funktions-, inhalts- oder wesensgleicher Objekte als Redundanz. Diese Redundanz drückt sich aus in *amodalen Informationen*, die *Zeit* (Synchronität), *Raum* (Konstanz von Objekt und

Schallquelle) sowie *Intensität* (Tempo, Rhythmus) umfassen. Beispiel hierfür sind, dass Händeklatschen und das Ballprellen. Bei beiden entsteht eine visuelle und akustische Korrespondenz von Objekt bzw. Ereignis und Schallquelle. Die amodale Information kann somit als grundlegend für das Lernen von bedeutungsvollen (intermodal redundante Anreize), konsistent auftretenden Objekten und Ereignissen angenommen werden. Diese intersensorische Redundanz unterstützt die Aufmerksamkeitszuwendung zu situativ entscheidenden Reizen. Sie schwächt die Aufmerksamkeit für nicht redundante Stimuli und erleichtert Wahrnehmung, Lernen und Gedächtnis für redundante oder amodale Stimuluseigenschaften (vgl. Bahrick et al. 2004, S. 99; Effenberg, 2004; S. 14).

Schluss

Vorausgesetzt, dass die Invariantenbildung für Wahrnehmung und Handeln über amodale Informationen unterstützt und gestärkt wird, verbleibt nach wie vor eine Reihe von Fragen:

- * Wie beeinflusst die Intention die Aufmerksamkeitslenkung und die Entdeckung der amodalen Informationen?
- * Welche Invarianten nutzt der Handelnde bei seiner Bewegungsentscheidung bzw. glaubt er zu nutzen?
- * Wie werden diese Invarianten in Muskelinnervation umgesetzt bzw. wie werden sie durch neue Muskelstrukturen – neue Formen von Invarianten – beeinflusst?
- * Welche Rolle spielt das somatosensorische System bei der Verarbeitung der Rezeptoraktivitäten?
- * Wie ist das Verhältnis von Gesamtzustand des Systems zu lokalen Rezeptorvorgängen?

Literaturverzeichnis

- Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). Integrated theory of the mind. *Psychological Review*, 111(4), 1036–1060.
- Bahrick, L. E., & Lickliter, R. (2002). Intersensory redundancy guides early perceptual and cognitive development. In R. Kail (Ed.), *Advances in Child Development and Behavior* (pp. 153–187). New York: Academic Press.
- Bahrick, L. E., Lickliter, R., & Flom, R. (2004). Intersensory redundancy guides the development of selective attention, perception, and cognition in infancy. *Current Directions in Psychological Science*, 13(3), 99–102.
- Carello, C., & Turvey, M. T. (2004). Physics and psychology of the muscle sense. *Current Directions in Psychological Science*, 13(1), 25–28.
- Driver, J., & Spence, C. (1998). Crossmodal links in spatial attention. *Philosophical Transactions of the Royal Society Section B*, 353, 1319–1331.
- Effenberg, A. O. (2004). *Synergien der Sinne für die Bewegungsregulation*. Frankfurt a. M.: Peter Lang.
- Hacker, W. (1998). *Arbeitspsychologie*. Bern: Huber
- Hoffmann, J. (1993). *Vorhersage und Erkenntnis*. Göttingen: Hogrefe.

- Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Austin: Houghton Mifflin.
- Grush, R. (2003). The Emulation Theory of Representation: Motor Control, Imagery, and Perception. *Behavioral and Brain Sciences* (in Press).
- Kelso, J. A. S. (2002). The Complementary Nature of Coordination Dynamics: Self-Organization and Agency. *Non Linear Phenomena in Complex Systems*, 5(4), 364-371.
- Lorenz, K. (1978). *Vergleichende Verhaltensforschung*. Wien/New York: Springer
- Mischel, W. (2004). Toward an Integrated Science of the Person. *Annual Review of Psychology*, 55, 1-23.
- Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press
- Pressing, J. (1999). The Referential Dynamics of Cognition and Action. *Psychological Review*, 106(4), 714-747.
- Rubinstein, S. L. (1977). *Sein und Bewusstsein*. Berlin: Akademie-Verlag.
- Spence, C., & Driver, J. (1996). Biovisual links in endogenous covert spatial attention. *Journal of Experimental Psychology: Human Perception and Performance*, 22, 1005-1030.
- Walsh, V., & Kulikowski, J. (Eds). (1998). *Perceptual Constancy: Why Things look as they do*. New York: Cambridge University Press.
- Ward, L. M., McDonald, J. J., & Lin, D. (2000). On asymmetries in cross-modal spatial attention orienting. *Perception and Psychophysics*, 62, 1258-1264.

CORD WESTHOFF & NIKOLAUS F. TROJE

Personenidentifikation anhand von biologischer Bewegung – strukturelle und kinematische Parameter

Einleitung

Die Untersuchung biologischer Bewegung wurde in den 70er Jahren vom schwedischen Forscher Gunnar Johansson in die experimentelle Psychologie eingeführt. Johansson bezeichnete biologische Bewegung als die „charakteristische Art der Fortbewegung von Lebewesen“. Als Gegenstand der Untersuchung benutzte er sog. „Point-Light Displays“ (PLDs). Diese basieren auf Aufnahmen menschlicher Bewegung, wobei die zu Grunde liegende Person nur durch einige wenige Lichtpunkte an den Hauptgelenken des Körpers dargestellt wird. In einer statischen Aufnahme erscheinen diese Lichtpunkte als inhaltslose Punktwolke. Ein kurzes Abspielen der Aufnahmen (200 ms) führt jedoch zu einer stabilen und eindeutigen Wahrnehmung menschlicher Bewegung (Johansson, 1973, 1976).

Seit den ursprünglichen Arbeiten von Johansson haben sich eine Reihe von Studien der Untersuchung der Wahrnehmung biologischer Bewegung gewidmet. Es konnte z. B. gezeigt werden, dass Beobachter die dargestellte Aktion aus PLDs erkennen können (Dittrich, 1993). Darüber hinaus kann die zu Grunde liegende Emotion erkannt werden, sowohl hinsichtlich ihrer Valenz wie auch des Grades der Aktivierung (Pollick et al., 2002). Weiterhin kann das Geschlecht einer Person (Troje, 2002a) sowie ihre Identität wahrgenommen werden. Während in verschiedenen Studien gezeigt werden konnte, dass Beobachter in der Lage sind, die Identität einer Person anhand von PLDs zu erkennen (z. B. Beardsworth & Buckner, 1981; Cutting & Kozlowski, 1977; Stevenage et al. 1999), wurde den dabei relevanten Faktoren bisher wenig Aufmerksamkeit geschenkt. Daher ist nicht klar, ob Personen z. B. anhand ihrer Größe, Körperstruktur, Schrittfrequenz, oder anhand der Art ihrer Bewegungen identifiziert werden können.

Methode

Durch den Einsatz eines „Motion Capture“ Systems wurden die Gangmuster verschiedener Personen aufgenommen. Hierbei wird zunächst ein Set von reflektierenden Markern am Körper der Person befestigt, deren Position im dreidimensionalen Raum von neun Kameras mit einer räumlichen Auflösung von 1 mm und einer zeitlichen Auflösung von 120 Hz bestimmt wird. Nach biomechanischer Modellierung und Fourier-Transformation der Rohdaten (vgl. Troje, 2002a; 2002b), ist jedes Gangmuster als ein Punkt in einem multidimensionalen Raum repräsentiert. Der entsprechende Vektor erhält Einträge für die Größe einer Person, ihre Körperstruktur (d. h., die durchschnittliche Position der Marker über die Zeit), ihre fundamentale Schrittfrequenz sowie die übrigen kinematischen Parameter in den Fourier-Komponenten. Diese Repräsentation erlaubt die unabhängige Variation der Parameter, z. B. die Normalisierung einzelner Faktoren über eine Gruppe von Gangmustern.

Studie 1

In der ersten Studie wurde untersucht, welche Rolle die Parameter Größe, Struktur und Schrittfrequenz bei der Wiedererkennung von individuellen PLDs spielen (Troje et al., 2005). Die VPn (N = 18) lernten, sieben verschiedene männliche Gangmuster aus einer von drei Perspektiven (Frontal, FV; Halb-Profil, HV; Profil, PV) zu benennen. Das Training fand an fünf aufeinander folgenden Tagen statt. In jedem Trial mussten die VPn das gezeigte Gangmuster einem von sieben Namen unter Feedback zuordnen. Im Laufe der Trainingssitzungen wurde nacheinander die Größe, Struktur und Schrittfrequenz der PLDs normalisiert, so dass diese Faktoren keinen individuellen Wiedererkennungswert mehr besaßen. Am Ende des Experiments fand ein zusätzlicher Testdurchgang statt, in dem die Generalisierung der Wiedererkennung auf die anderen Blickwinkel überprüft wurde.

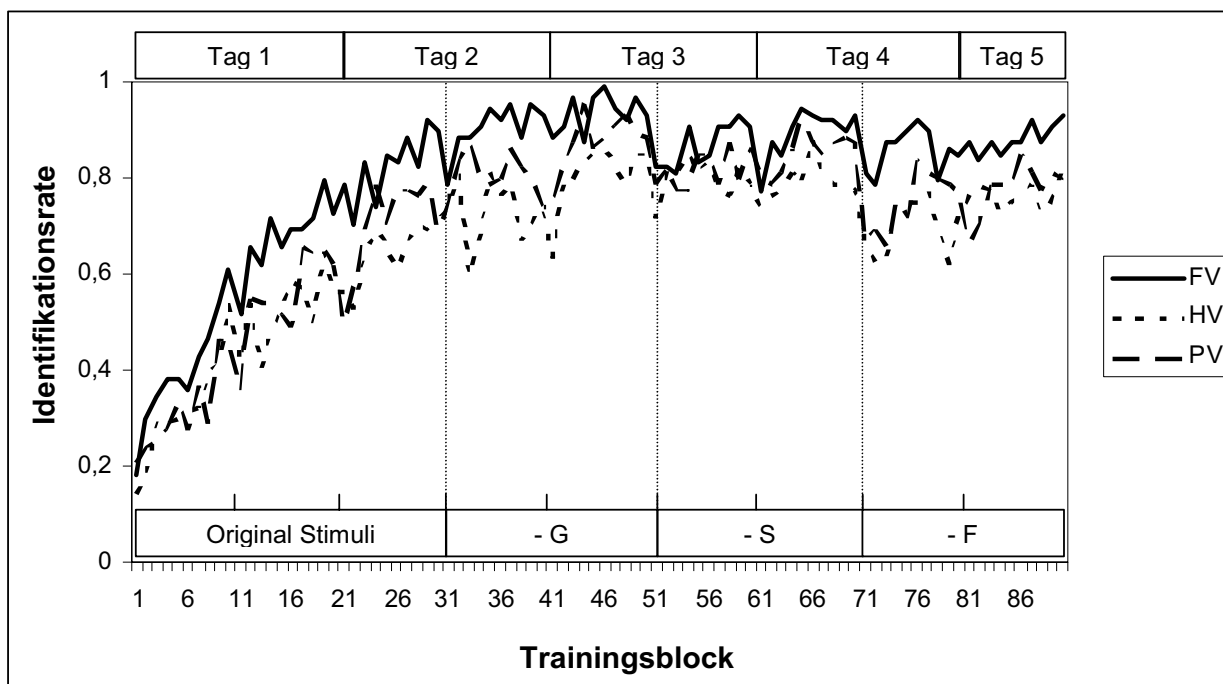


Abb. 1: Lernkurven in Studie 1. – G: Größe norm.; – S: Größe u. Struktur norm.; – F: Größe, Struktur u. Frequenz norm.; FV: Frontalansicht; HV: Halb-Profil Ansicht; PV: Profilansicht.

Alle drei Gruppen (eingeteilt nach Blickwinkel) zeigten eine ansteigende Lernkurve im Laufe der Trainings mit den originalen Stimuli. Nach den ersten drei Sitzungen erreichte die Leistung ein Niveau von 70-90% korrekt erkannter Gangmuster (Abb. 1). Um den Einfluss der drei Faktoren zu berechnen, wurde die mittlere Leistung der jeweils letzten drei Trainingsblöcke vor der Normalisierung mit der mittleren Leistung in den ersten drei Blöcken nach der Normalisierung verglichen. Nach der Normalisierung von Struktur und Schrittfrequenz zeigte sich eine signifikante Verschlechterung der Wiedererkennung, nicht jedoch nach der Normalisierung der Größe der PLDs (Abb. 2). Die Überprüfung der Generalisierungsleistung zeigte in allen Gruppen eine signifikante Verschlechterung der Leistung, wenn der Blickwinkel im Testdurchgang von demjenigen im Training abwich (Abb. 3). Alle Ergebnisse blieben jedoch signifikant über Zufallsniveau (14%).

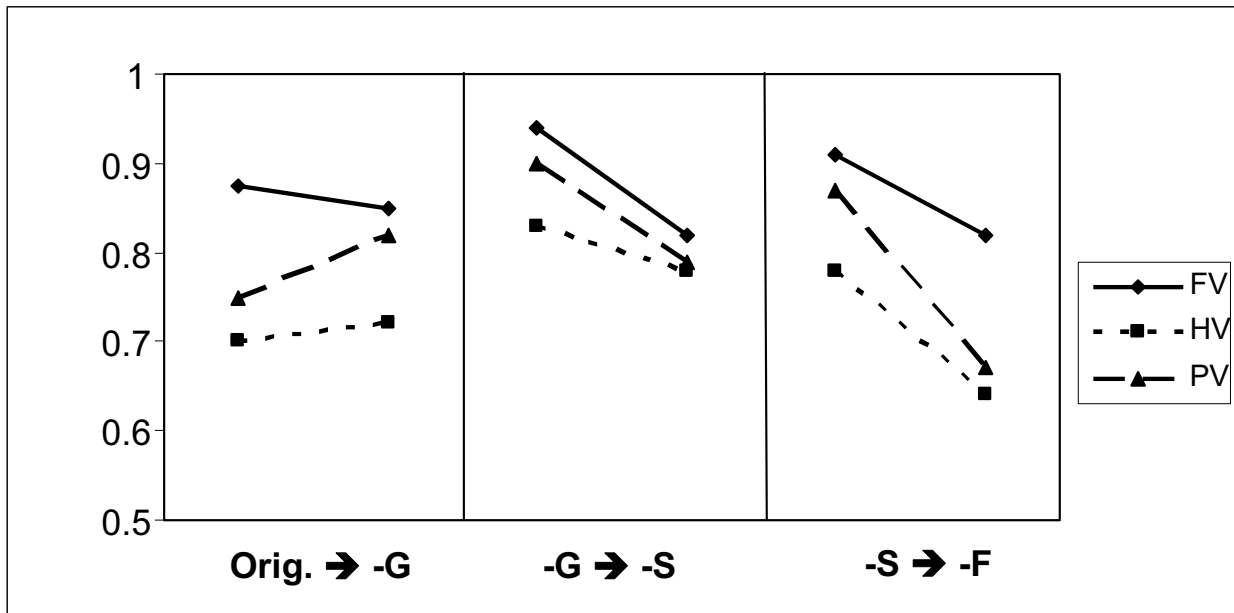


Abb. 2: Effekte der Normalisierung in Studie 1. FV: Frontalansicht; HV: Halb-Profil Ansicht; PV: Profilsansicht.

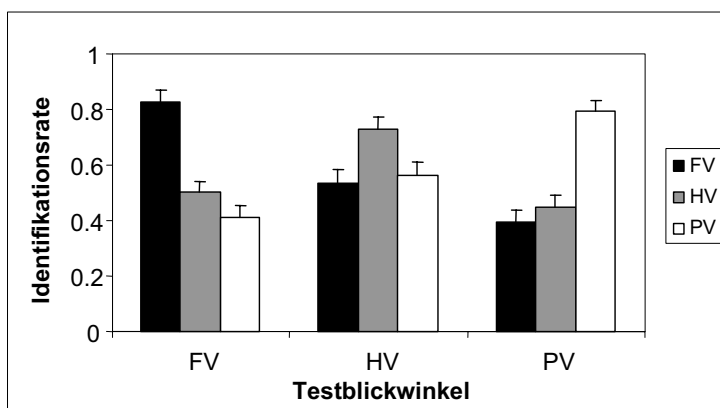


Abb. 3: Generalisierungsleistung der drei Gruppen (Balken) in Studie 1 über alle Blickwinkel (x-Achse). FV: Frontalansicht; HV: Halb-Profil Ansicht; PV: Profilsansicht.

Trotz Normalisierung der drei Parameter betrug die Wiedererkennungsrates am Ende des Experiments ca. 80-90% (vgl. Abb. 1). Die übrigen kinematischen Parameter, die in den Fourier-Komponenten codiert werden, müssen also noch viel individuelle Information enthalten, die zur Identifikation individueller Gangmuster genutzt werden kann. In der zweiten Studie wurden diese kinematischen Faktoren daher genauer untersucht.

Studie 2

Die Stimuli, die in der zweiten Studie verwendet wurden, waren bereits hinsichtlich Größe, Struktur und Schrittfrequenz normalisiert, um die individuelle Information auf die kinematischen Parameter zu beschränken. In zwei Experimenten wurde die Rolle der Harmonischen der Fourier-Analyse bzw. des Phasen- und Amplitudenspektrums der Gangmuster überprüft. Hierzu wurde zunächst ein mittlerer Läufer aus den individuellen Bewegungen erstellt. Für die Teststimuli wurde im ersten Experiment entweder die erste, die zweite oder die dritte bis fünfte Harmonische dieses Läufers durch die individuellen Harmonischen der originalen Läufer ersetzt, so dass diese nur anhand der jeweiligen Harmonischen identifizierbar waren. Im zweiten Experiment wurde entweder das komplette Phasen-Spektrum des mittleren Läufers, welches die zeitliche Relation der Punkte zueinander kodiert, oder das

Amplituden-Spektrum, das die räumliche Auslenkung der Punkte enthält, durch das jeweilige Spektrum der individuellen Läufer ersetzt.

In sechs Trainingssitzungen an zwei aufeinander folgenden Tagen lernten die VPn in beiden Experimenten (jeweils N = 18) sieben männliche Gangmuster aus einem der drei Blickwinkel, die schon im ersten Experiment benutzt wurden, zu benennen. In einem anschließenden Testdurchgang wurde die Wiedererkennung der Stimuli mit den individuellen Harmonischen bzw. mit dem Phasen- oder Amplituden-Spektrum unter allen drei Blickwinkeln geprüft.

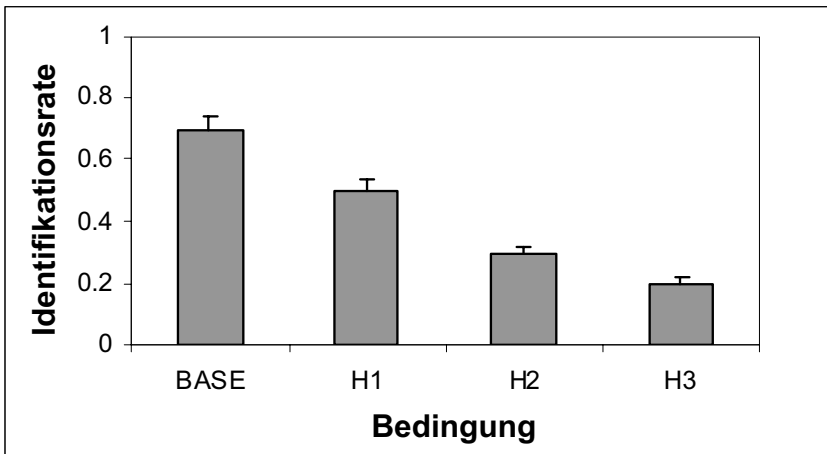


Abb. 4: Identifikationsleistung in Experiment 1 von Studie 2. BASE: Trainingsstimuli; H1-H3: 1./ 2./ 3.-5. Harmonische.

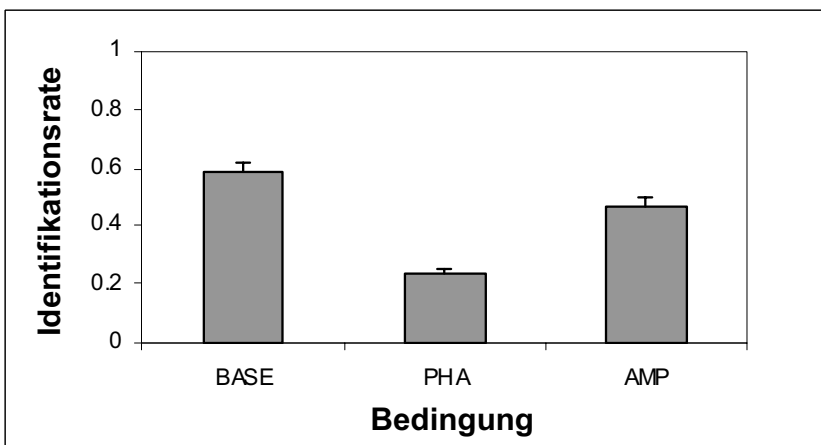


Abb. 5: Identifikationsleistung in Experiment 2 von Studie 2. BASE: Trainingsstimuli; PHA: Phasenspektrum; AMP: Amplitudenspektrum.

Die Ergebnisse des ersten Experiments zeigten, dass die erste Harmonische der Fourier-Analyse den größten Beitrag zur Wiedererkennung der Gangmuster aufweist (Abb. 4).

Allerdings war auch die Identifikationsrate für diejenigen Läufer, die die individuelle Information nur in der zweiten Harmonischen enthielten, signifikant über Zufallsniveau, wohingegen die dritte bis fünfte Harmonische keinen signifikanten Beitrag leistet. Im zweiten Experiment konnte gezeigt werden, dass sowohl das Amplituden- wie auch das Phasen-Spektrum zur Identifikation beitragen, ersteres jedoch einen deutlich größeren Einfluss hat (Abb. 5).

Die zeitliche Relation der Punkte spielte nur unter einem frontalen Blickwinkel eine Rolle. Die Befunde hinsichtlich der Generalisierung über verschiedene Blickwinkel bestätigten die Ergebnisse der ersten Studie.

Zusammenfassung

Im Rahmen von zwei Studien konnte gezeigt werden, dass Beobachter in der Lage sind, verschiedene Personen anhand ihrer Gangmuster, die als Point-Light Displays gezeigt werden, zu differenzieren. Sowohl strukturelle (Körperstruktur) wie

auch kinematische Parameter haben Einfluss auf die Wiedererkennung dieser Bewegungen. Im Bereich der Kinematik tragen sowohl die erste wie auch die zweite Harmonische einen wesentlichen Teil der Information. Außerdem leistet das Amplituden-Spektrum einen gegenüber dem Phasen-Spektrum signifikant größeren Beitrag zur Identifikation individueller Gangmuster. Auf Grund der Generalisierungsleistung kann darauf geschlossen werden, dass die Beobachter nicht nur die lokalen Trajektorien der Lichtpunkte wiedererkennen, sondern dass eine genauere Repräsentation der zu Grunde liegenden Struktur vorliegt.

Literatur

- Beardsworth, T., & Buckner, T. (1981). The ability to recognize oneself from a video recording of one's movements without seeing one's body. *Bulletin of the Psychonomic Society*, 18, 19-22.
- Cutting, J. E., & Kozlowski, L. T. (1977). Recognizing friends by their walk: Gait perception without familiarity cues. *Bulletin of the Psychonomic Society*, 9, 353-356.
- Dittrich, W. H. (1993). Action categories and the perception of biological motion. *Perception*, 22, 15-22.
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, 14, 201-211.
- Johansson, G. (1976). Spatio-temporal differentiation and integration in visual motion perception. *Psychological Research*, 38, 379-393.
- Pollick, F. E., Lestou, V., Ryu, J., & Sung-Bae, C. (2002). Estimating the efficiency of recognizing gender and affect from biological motion. *Vision Research*, 42, 2345-2355.
- Stevenage, S. V., Nixon, M. S., & Vince, K. (1999). Visual analysis of gait as a cue to identity. *Applied Cognitive Psychology*, 13, 513-526.
- Troje, N. F. (2002a). Decomposing biological motion: a framework for analysis and synthesis of human gait patterns. *Journal of Vision*, 2, 371-387.
- Troje, N. F. (2002b). The little difference: Fourier based synthesis of gender-specific biological motion. In R. P. Würtl & M. Lappe (Eds.), *Dynamic Perception* (pp. 115-120). Berlin: Aka Verlag.
- Troje, N. F., Westhoff, C., & Lavrov, M. (2005) Person identification from biological motion: Effects of structural and kinematic cues. *Perception & Psychophysics*, 67, 667-675.

MICHAEL CLAUSEN, MEINARD MÜLLER & TIDO RÖDER

Informatische Aspekte der Bewegungsanalyse

Einführung

Die Analyse menschlicher Bewegungen im Sport, in der Medizin und in der Biometrie basierte bislang häufig auf der halbautomatischen Auswertung von Fotoreihen oder Videomaterial. Vorangetrieben durch Entwicklungen im Bereich der Computeranimation sind in jüngster Zeit computergestützte Aufnahmetechniken zur dreidimensionalen Erfassung von Bewegungen entstanden, die unter dem Oberbegriff *Motion Capturing* (MoCap) zusammengefasst werden.

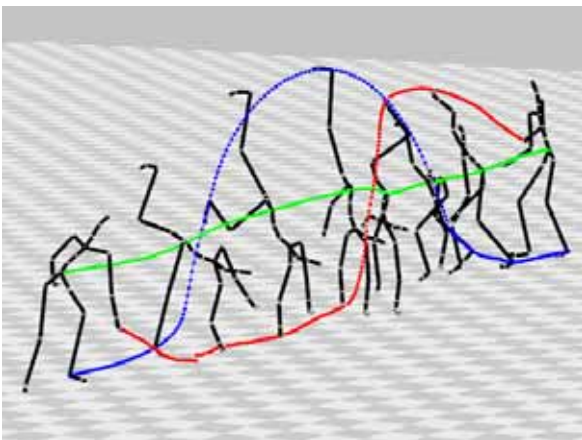


Abb. 1: Einzelposen aus einer MoCap-Aufnahme eines Radschlags. Die 3D-Trajektorien einiger Gelenke sind eingezeichnet.

Dieser Beitrag skizziert einige Grundtechniken zur Bewegungsanalyse basierend auf MoCap-Daten. Insbesondere geht es um die fundamentalen Fragen, wann zwei MoCap-Aufnahmen als *ähnlich* anzusehen sind und wie ähnliche Bewegungen auf automatische Weise in zeitliche Korrespondenz gebracht werden können. Als mögliche Herangehensweise an diese Problemstellungen wird das generische Konzept des *Dynamic Time Warping* vorgestellt.

Optisches Motion Capturing und Datenformate

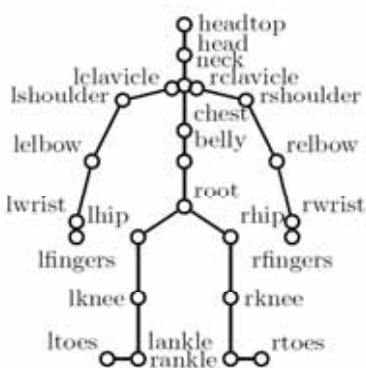


Abb. 2: Skelettmodell.

Bevor wir uns dem Problem der Ähnlichkeitsbewertung für Bewegungen widmen, soll es zunächst um die Aufnahmetechnik und das Datenformat der zugrundeliegenden MoCap-Daten gehen. Beim weitverbreiteten *optischen* Motion Capturing wird eine Anordnung von mehreren Hochgeschwindigkeitskameras verwendet, um durch wiederholte Positionsbestimmung die dreidimensionalen Trajektorien von reflektierenden Markern an ausgewählten Punkten des menschlichen Körpers zu erfassen. Der resultierende MoCap-Datenstrom ist eine fortlaufende Folge von *Frames*, die für jeden Aufnahmezeitpunkt die räumlichen Positionen der ausgewählten Körperpunkte mit einer Auflösung im Millimeterbereich beschreiben, siehe Abbildung 1.

Das gebräuchlichste Modell zur Beschreibung menschlicher Bewegungen ist das der kinematischen Kette, einem vereinfachten Abbild des menschlichen Skelettes, siehe Abbildung 2.

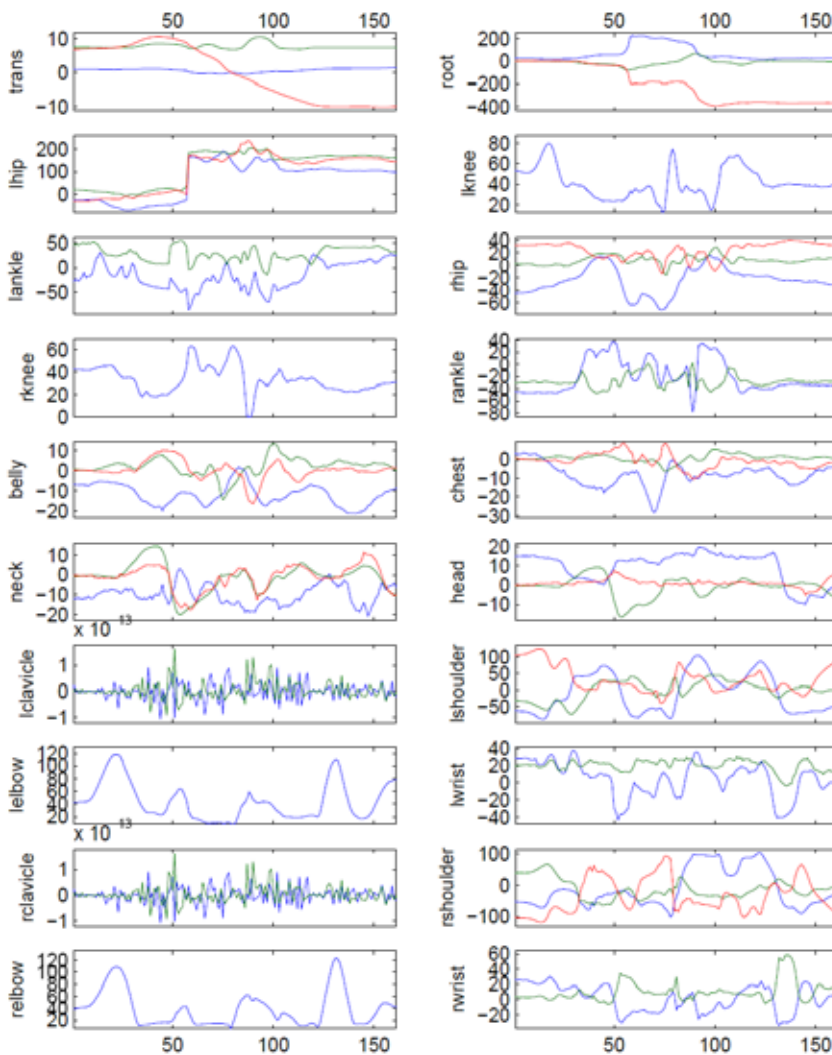


Abb. 3: Gelenkwinkeltrajektorien für den Radschlag aus Abb. 1. Die horizontalen Achsen repräsentieren die Zeit, die vertikalen Achsen die Eulerwinkel in Grad.

verlauf von Gelenken oder den 3D-Ortskoordinatenverlauf von ausgezeichneten Körperpunkten beschreiben (siehe Abbildung 1 und 3).

Aspekte der Ähnlichkeit

Beim Vergleich zweier Bewegungen gibt es Aspekte, die typischerweise keinen Einfluss auf die Ähnlichkeitsbewertung haben sollten. So entspricht es meist der menschlichen Wahrnehmung, globale räumliche Transformationen wie Translationen der Gesamtbewegungen (räumliche Position der Bewegung), räumliche Skalierungen (Größe des Menschen) oder Rotationen um die Hochachse (Richtung der Bewegung) beim Vergleich unberücksichtigt zu lassen, siehe Abbildung 4. In einem gewissen Rahmen sind oft auch globale zeitliche Transformationen (Geschwindigkeit der Bewegung) irrelevant.

Schwieriger wird es für automatische Ähnlichkeitsbewertungen, wenn Bewegungen auch dann noch als ähnlich angesehen werden sollen, wenn sie zwar derselben Bewegungsklasse entstammen, aber in verschiedenen Stilen ausgeführt werden. In Abbildung 5 sieht man z. B. neun Posen aus verschiedenen Gehbewegungen, die

Wie auch das menschliche Skelett besteht eine kinematische Kette aus verschiedenen Typen von beweglichen Gelenken und diese aus verbindenden Segmenten, welche den Knochen entsprechen. Mittels inverser Kinematik kann jeder Frame eines MoCap-Datenstroms in eine Pose der zugrundegelegten kinematischen Kette umgerechnet werden. Eine Pose besteht aus einem Parametersatz, der den Zustand der kinematischen Kette zu einem festen Zeitpunkt beschreibt. Hierzu werden meistens 3D-Ortskoordinaten oder Gelenkwinkel verwendet. Durch Verfolgung solcher Parameterwerte über die Zeit erhält man sogenannte *Trajektorien*, die z. B. den Winkel-

den Stilen „Marschieren“, „Schleichen“, „Humpeln“, „Schlurfen“, „trauriges Gehen“ usw. entsprechen. Die Übergänge zwischen solchen Bewegungsstilen sind fließend, so dass es selbst für den Menschen häufig schwierig ist, Bewegungen hinsichtlich ihrer Ähnlichkeit zu klassifizieren.

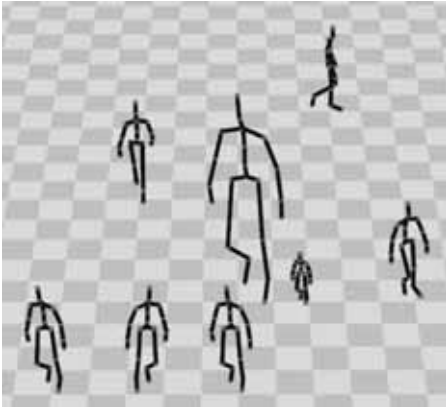


Abb. 4: Transformationen einer Gehbewegung.

Neben globalen Transformationen und Variationen des Bewegungsstils können *lokale* raum-zeitliche Deformationen zu erheblichen Unterschieden zwischen Bewegungen führen. So unterscheiden sich z. B. die in Abbildung 6 gezeigten Gehbewegungen außer in der globalen Fortbewegungsgeschwindigkeit auch hinsichtlich zeitlicher Phasenbeziehungen innerhalb der Gehzyklen sowie hinsichtlich räumlicher Abweichungen, wie z. B. Pendelbewegungen der Arme.

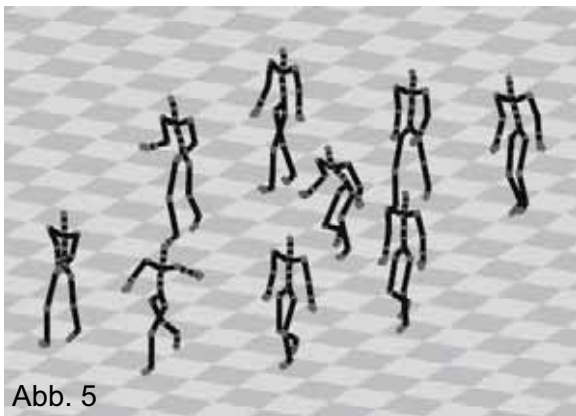


Abb. 5

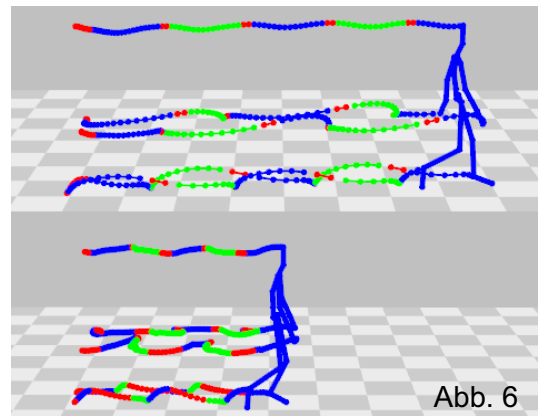


Abb. 6

Abb. 5: Verschiedene Stile beim Gehen. Abb. 6: Schnelles, dynamisches Gehen (oben) gegenüber langsamem Gehen einer alten Frau (unten). Korrespondierende Phasen der Gehzyklen sind in gleichen Trajektorienfarben dargestellt.

Im Folgenden wird eine generische Technik zur Synchronisation von Bewegungen, die sich durch lokale raum-zeitliche Deformationen voneinander unterscheiden, vorgestellt.

Lokale Kostenmaße

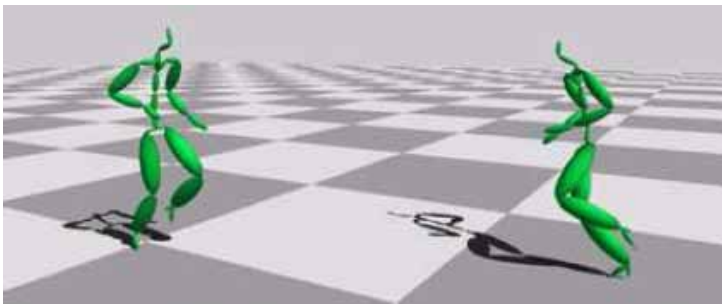
Wie oben erwähnt kann ein MoCap-Datenstrom als Folge von Einzelposen der zugrundeliegenden kinematischen Kette aufgefasst werden. Es liegt daher nahe, beim Vergleich von Bewegungen eine Bottom-Up-Sicht einzunehmen und *lokale* Ähnlichkeitsbewertungen zwischen Posen in geeigneter Weise zu *globalen* Ähnlichkeitsbewertungen zwischen Gesamtbewegungen zusammenzufügen. Wir gehen im Folgenden auf zwei solche lokale Ähnlichkeitsmaße ein. Bei Maßen, die die *Unähnlichkeit* von Bewegungen messen, spricht man auch von *Kostenmaßen*.

Kovar und Gleicher (2004) beschreiben jede Pose durch die Familie der Ortskoordinaten der Skelettpunktmenge, welche mit J bezeichnet sei. Der Vergleich der Posen zu den Zeitpunkten m und n erfolgt im Kontext einer festen Anzahl 2ρ von

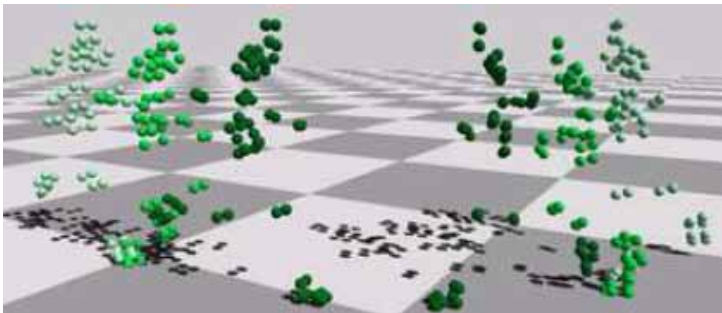
vorhergehenden und nachfolgenden Posen, und die resultierenden Punktwolken werden mittels einer Transformation T so verschoben (Translationsparameter x_0, z_0) und gedreht (Rotationsparameter θ), dass die gewichtete Summe der euklidischen Abstände einander entsprechender Punkte minimiert wird. Dies sind die lokalen Kosten $d^p(m, n)$, oder auch der *Abstand* zwischen der m -ten und der n -ten Pose:

$$d^p(m, n) = \min_{\theta, x_0, z_0} \sum_{t \in [-\rho; \rho]} \sum_{j \in J} w_j \|p_{m+t}^j - T_{\theta, x_0, z_0} p_{n+t}^j\|^2$$

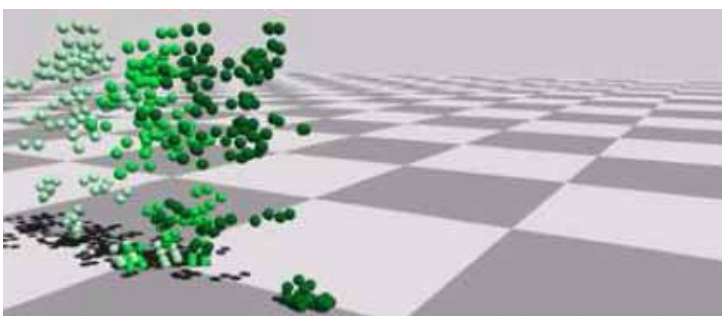
Hierbei bezeichnet p_k^j den Ortskoordinatenvektor des Skelettpunkts $j \in J$ zum Zeitpunkt k . Die Gewichte $w_j \in \mathbf{R}$ für $j \in J$ ermöglichen eine unterschiedlich starke Berücksichtigung der verschiedenen Körperteile.



(a)



(b)



(c)

Ein anderes Kostenmaß basiert auf einer Gelenkwinkeldarstellung von Posen. Mittels sogenannter Einheitsquaternionen q lassen sich die Rotationen an den Gelenken beschreiben. Man bildet nun analog zum vorherigen Kostenmaß eine gewichtete Summe der Beträge von Skalarprodukten solcher Einheitsquaternionen, wobei q_k^j die Quaternionendarstellung der Gelenkwinkel am Gelenk $j \in J$ zum Zeitpunkt k bezeichnet:

$$d^q(m, n) = 1 - \sum_{j \in J} w_j |\langle q_m^j, q_n^j \rangle|$$

Beide vorgestellten Kostenmaße sind invariant gegenüber globalen Translationen und Rotationen. Weiterhin ist d^q invariant gegenüber räumlichen Skalierungen. Im Folgenden wird exemplarisch auf das Kostenmaß d^q zurückgegriffen.

Abb. 7: Zum lokalen Kostenmaß nach Kovar und Gleicher: (a) zwei Posen; (b) Umwandlung in Punktwolken unter Berücksichtigung eines zeitlichen Kontextes; (c) Ausrichtung der beiden Punktwolken durch Rotation und Translation.

Kostenmatrizen

Die lokalen Kosten fasst man systematisch zu sogenannten Kostenmatrizen zusammen. Die Kostenmatrix für zwei MoCap-Bewegungen bestehend aus M bzw. N Posen ist eine Matrix mit M Zeilen und N Spalten, wobei der Eintrag in Zeile m und

Spalte n die lokalen Kosten $d^p(m,n)$ beschreibt. Zur Veranschaulichung dieses Konzepts betrachten wir zunächst den Spezialfall, dass eine Bewegung mit sich selbst verglichen wird.

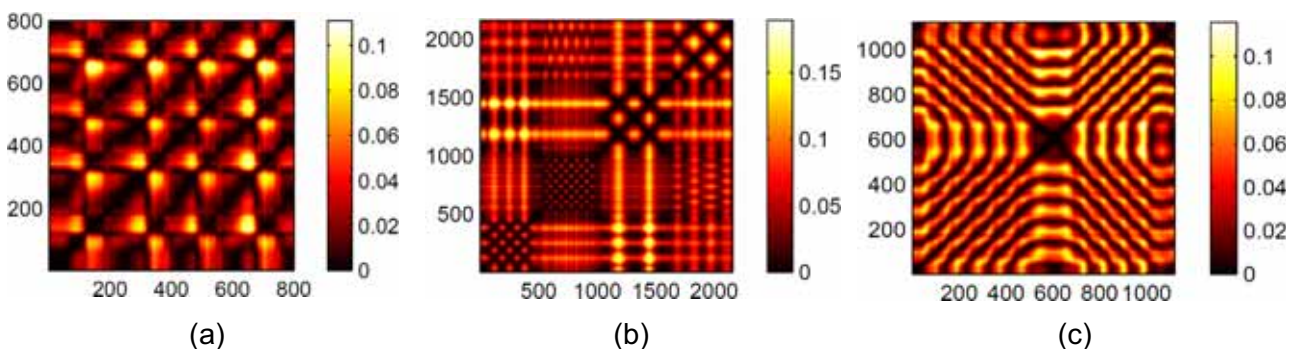


Abb. 8: Kostenmatrizen für den Vergleich diverser Bewegungen mit sich selbst; dunklere Farben stehen für niedrigere Kosten (a) Hüpfbewegung bestehend aus vier Sprüngen; (b) Gymnastiksequenz; (c) Armkreisen, fünf mal vorwärts, vier mal rückwärts.

Abbildung 8 (a) zeigt die Kostenmatrix einer Sequenz von vier aufeinanderfolgenden Hüpfbewegungen. In diesem Fall ist die Kostenmatrix quadratisch und symmetrisch bezüglich der Hauptdiagonalen (diese verläuft von links unten nach rechts oben), entlang derer nur Nullkosten auftreten, da hier die Unähnlichkeit jeder Pose zu sich selbst gemessen wird. Zusätzlich erkennt man verschiedene Nebendiagonalen niedriger Kosten. Diese zeigen an, dass in der Bewegung manche Teile wiederholt vorkommen. So besagt z. B. die Nebendiagonale, die an der vertikalen Achse bei Pose 400 ansetzt, dass die zwei ersten Sprünge (Posen 1-400) ähnlich zu den zwei letzten Sprüngen (Posen 400-800) sind. Man beachte, dass die beiden Dreiecke oberhalb und unterhalb der Hauptdiagonalen aufgrund der Symmetrie des lokalen Kostenmaßes dieselbe Information enthalten.

Als weiteres Beispiel soll eine Gymnastiksequenz dienen, siehe Abbildung 8 (b): zwei Hampelmänner, zehn Schritte auf der Stelle, zwei Kniebeugen, vier Wiederholungen von „Knie berührt Ellenbogen“. Auch hier spiegelt sich die Repetitivität der einzelnen Übungen in der Nebendiagonalenstruktur wider. Weiterhin fallen die zwei Streifen hoher Kosten auf, die sich bei den Posen 120 und 145 durch die gesamte Matrix ziehen. Diese Streifen rühren daher, dass die beiden Kniebeugen sehr unähnlich zu sämtlichen anderen Bewegungen sind. Abb. 8 (c) zeigt die Kostenmatrix für fünfmaliges Armkreisen vorwärts (Posen 1-500) gefolgt von viermaligem Armkreisen rückwärts (Posen 700-1000), jeweils synchron mit beiden Armen ausgeführt. Parallel zur Hauptdiagonalen entstehen auch hier Nebendiagonalen niedriger Kosten als Ausdruck der Bewegungswiederholungen. Zusätzlich ergeben sich auffällige Strukturen in der oberen linken bzw. unteren rechten Ecke der Matrix – man erkennt Bereiche niedriger Kosten, die *senkrecht* zur Hauptdiagonalen verlaufen. Betrachten wir z. B. die Diagonale, die bei Pose 1050 an der vertikalen Achse ansetzt und nach rechts unten bis Pose 450 auf der horizontalen Achse verläuft. Diese Diagonale lässt sich wie folgt interpretieren: Denkt man sich die ersten vier Wiederholungen des Vorwärts-Armkreisens rückwärts abgespielt, dann sind sie ähnlich zu den vier Wiederholungen des Rückwärts-Armkreisens.

Dynamic Time Warping

Berechnet man die Kostenmatrix für zwei verschiedene Bewegungen, wie etwa in Abb. 9 für zwei verschieden schnelle Gehbewegungen, lassen sich analoge Informationen ablesen. Die schnellere Gehbewegung bezeichnen wir hier informell als die *vertikale* Bewegung, weil ihre Zeitachse in der Matrix vertikal verläuft. Entsprechend nennen wir die langsamere Gehbewegung die *horizontale* Bewegung. Pfade von der linken unteren Ecke in die rechte obere Ecke der Kostenmatrix entsprechen hier Zuordnungen von Posen zwischen den Bewegungen: verläuft solch ein Pfad in einem gewissen Bereich in horizontaler Richtung, entspricht das einer Zuordnung mehrerer Posen der horizontalen Bewegung zu einer Pose der vertikalen Bewegung, die Zeit bleibt also für die vertikale Bewegung stehen. Umgekehrt entsprechen vertikale Pfadabschnitte einem Stehenbleiben der Zeit für die horizontale Bewegung. Bei einem diagonalen Verlauf läuft in beiden Bewegungen die Zeit weiter. Von besonderem Interesse sind *kostenoptimale* Pfade in der Kostenmatrix, das sind solche Pfade, entlang derer die Summe der Kosten ein Minimum annimmt.

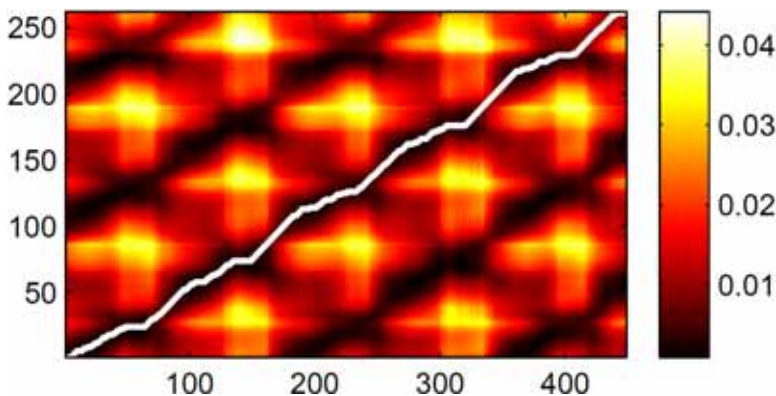


Abb. 9: Kostenmatrix mit weiß markiertem kostengünstigstem Pfad. Es wird eine schnelle Gehbewegung (vertikale Achse) mit einer langsamen Gehbewegung (horizontale Achse) verglichen; jeweils fünf Schritte.

Solche Pfade codieren eine optimale zeitliche Verzerrung, um die beiden Bewegungen bezüglich des gewählten lokalen Kostenmaßes zu *synchronisieren*. Die räumlichen Verzerrungen zwischen den Bewegungen werden dabei durch das lokale Kostenmaß erfasst. Im Beispiel von Abb. 9 sieht man, dass der eingezeichnete Pfad eher horizontal als vertikal ausgerichtet ist. Die Zeit

verläuft also für die vertikale Bewegung langsamer, um sie an die langsamere horizontale Bewegung anzupassen. *Dynamic Time Warping* (DTW) ist eine Technik zur automatischen Berechnung solcher kostenoptimaler Pfade. Der Hauptnachteil dieser Technik ist die hohe Berechnungskomplexität: Zur Synchronisation zweier Bewegungen mit M bzw. N Frames werden größenordnungsmäßig $M \cdot N$ Rechenoperationen benötigt, was für große M und N in der Praxis zu viel ist. Ein völlig anderer, alternativer Ansatz ohne derartige Einschränkungen findet sich in Müller, Röder und Clausen (2005).

Literatur

- Kovar, L., & Gleicher, M. (2004). Automated extraction and parameterization of motions in large data sets. *ACM Transactions on Graphics*, 23(3), 559-568.
- Müller, M., Röder, T., & Clausen, M. (2005). Efficient content-based retrieval of motion capture data. *ACM Transactions on Graphics*, 24(3), 677-685.

TIDO RÖDER & MEINARD MÜLLER

Geometrische Relationen für die Bewegungsanalyse

Einführung

Die Erforschung und Entwicklung von grundlegenden Techniken und Werkzeugen für den Vergleich, die Klassifikation und das Retrieval menschlicher Bewegungsdaten stellt in der Informatik ein aktuelles und anwendungsnahe Forschungsgebiet mit zahlreichen offenen wissenschaftlichen Fragestellungen dar. Die Erfassung solcher Daten erfolgt heutzutage oft unter Verwendung sogenannter optischer, mechanischer oder magnetischer Motion-Capture-Systeme. Motion-Capture-Daten, im folgenden auch MoCap-Daten genannt, kamen in einfacher Form bereits um 1980 in biomechanischen Anwendungen wie der Analyse von Anomalien des menschlichen Gangs zum Einsatz.

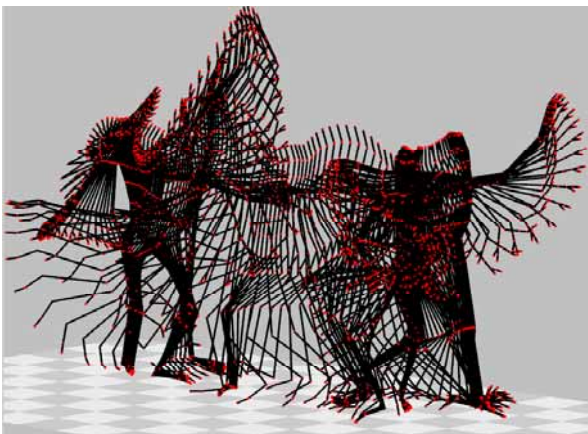


Abb. 1: MoCap-Aufnahme einer Ballettbewegung als Folge von Einzelposen.

In den 90er Jahren erfolgte dann der kommerzielle Durchbruch für MoCap-Anwendungen im Bereich der Computeranimation. Durch den Einsatz von automatisch gewonnenen MoCap-Daten gelang es hier erstmals im großen Stil, künstlichen Charakteren natürlich wirkende Bewegungen aufzuprägen. Neben Anwendungen in der Computeranimation und der Biomechanik stellen MoCap-Techniken mittlerweile wichtige Hilfsmittel in der Biometrie (Personenidentifikation), der Medizin und der Sportwissenschaft dar.

Im Gegensatz zum reinen Bild- oder Videodatenformat liegt dem MoCap-Datenformat ein Modell in Form einer kinematischen Kette zugrunde (siehe Abbildung 2).

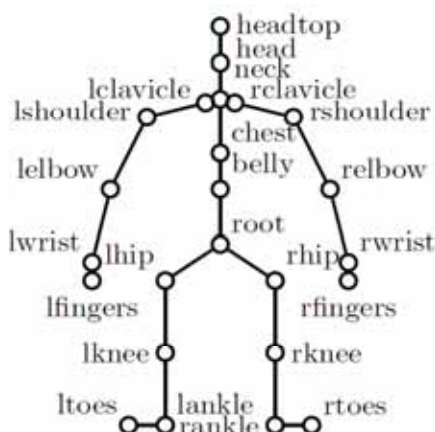


Abb. 2: Skelettmodell.

Analog zum menschlichen Skelett besteht eine solche kinematische Kette aus *Segmenten* (den „Knochen“), die durch *Gelenke* unterschiedlicher Typen verbunden sind. Da bei MoCap-Daten die 3D-Ortskoordinaten der Gelenke explizit und in hoher Genauigkeit vorliegen, spielen typische Computer-Vision-Probleme wie Tracking oder 3D-Rekonstruktion keine Rolle mehr. Durch die strukturelle Modellinformation, also das Wissen, welcher 3D-Punkt zu welchem Körperteil korrespondiert, eröffnen MoCap-Daten neue Möglichkeiten zur Bewegungsanalyse, -synthese und -klassifikation auf hohem semantischen Niveau. So geht es in aktuellen Forschungsprojekten z. B.

um die Untersuchung von Bewegungsdetails, die natürlich aussehende von künstlichen Bewegungen unterscheiden (Ren, Patrick, Efros, Hodgins, & Rehg, 2005). Ebenso steht man bei der Erforschung von Parametern zur Beschreibung bestimmter Bewegungsstile vor zahlreichen offenen Fragestellungen (Liu, Hertzmann, & Popović, 2005; Hsu, Pulli, & Popović, 2005). Große Probleme bereitet auch die Identifikation „logisch ähnlicher“ Bewegungsmuster – bisherige *quantitative* Ansätze wie der von Kovar und Gleicher (2004) verfolgte Zugang sind nur bedingt geeignet, die starken raum-zeitlichen Variabilitäten zu erfassen, die bei verschiedenen Ausführungen perzeptuell ähnlicher Bewegungsabläufe auftreten.

In diesem Beitrag skizzieren wir einen *qualitativen*, auf geometrischen Relationen basierenden, Ansatz zur Erfassung von großmotorischen Bewegungscharakteristika. Davon ausgehend zeigen wir, wie sich auf natürliche Weise semantisch sinnvolle Segmentierungen von Bewegungen in Bewegungsfragmente ergeben. Als Anwendungsbeispiel beschreiben wir abschließend eine Technik zur inhaltsbasierten Suche in Bewegungsdaten, siehe auch Müller, Röder und Clausen (2005).

Geometrische Relationen

Aufbauend auf dem Skelettmodell, das den MoCap-Daten zugrunde liegt, führen wir nun informell das Konzept der geometrischen Relationen ein. Betrachten wir dazu beispielhaft die beiden Skelettposen in Abbildung 3(a).

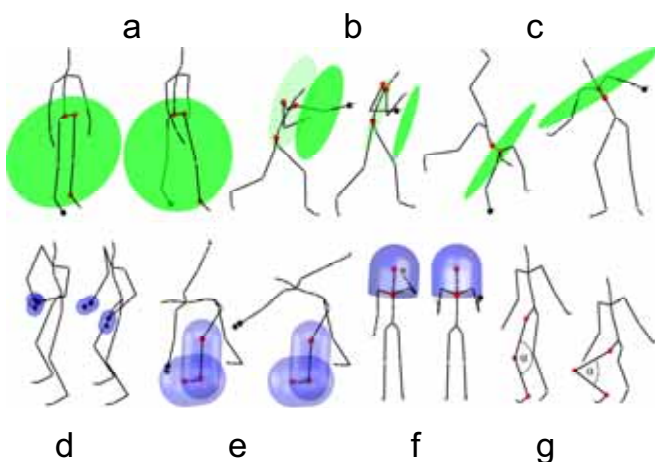


Abb. 3: Beispiele geometrischer Relationen.

Die grünen Scheiben stellen die Ebene dar, die vom linken Sprunggelenk, vom linken Hüftgelenk und vom Hüftzentrum aufgespannt werden. Je nachdem ob nun das *rechte* Sprunggelenk vor oder hinter dieser Ebene liegt, ordnen wir der jeweiligen Pose den Wahrheitswert null (Fuß vor der Ebene) oder eins (Fuß hinter der Ebene) zu. Vereinfacht gesagt handelt es sich hierbei um einen Test, ob sich der rechte Fuß gerade vor oder hinter dem Körper befindet. Durch Vertauschung

von links und rechts bei der Definition der Testebene ergibt sich die komplementäre Relation für den linken Fuß. Da diese Testebene am Körper fixiert ist, erhält man automatisch die Invarianz der Relationen gegenüber Translationen, Rotationen, Skalierungen, Änderungen der Skelettgeometrie und zu großen Teilen auch gegenüber Änderungen der Körperhaltung. Analog stellt Abbildung 3(b) einen Test auf eine vorwärtsgestreckte Hand dar und Abbildung 3(c) einen Test auf eine angehobene Hand (Testkörperteile jeweils mit verdickten Gelenkpunkten markiert). Die Abbildungen 3(d-f) zeigen Tests auf Berührungen verschiedener Körperteile, während in 3(g) auf Anwinkelung eines Knies geprüft wird. Viele weitere derartige Tests sind denkbar.

Einzelne Skelettposen und Bewegungen, also zeitliche Abfolgen von Skelettposen, können nun auf einfache Weise qualitativ erfasst werden. Anstelle der quantitativen Spezifikation der Posen, z. B. mittels 3D-Trajektorien, wird jede Pose qualitativ durch einen *Merkmalsvektor* beschrieben, der die Wahrheitswerte einer geeigneten Menge geometrischer Relationen kodiert.

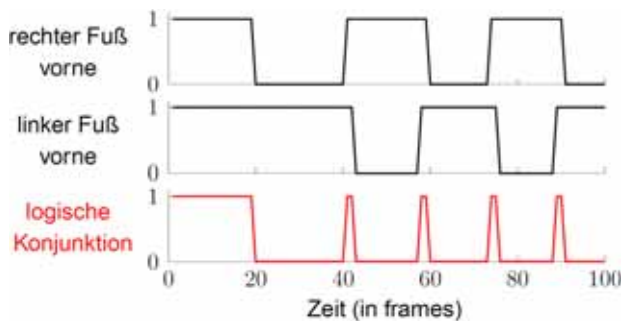


Abb. 4: Logische Verknüpfung geometrischer Relationen für eine Gehbewegung (Schrittfolge links-rechts-links-rechts-links).

Als Beispiel betrachten wir einen Übergang vom Stehen ins Gehen und konzentrieren uns dabei auf die grobe Bewegung der Beine. Dazu wählen wir die aus Abbildung 3(a) bekannte Relation „rechter Fuß vorne“ und zusätzlich die komplementäre Relation „linker Fuß vorne“. Abbildung 4 (oben) zeigt die resultierenden zeitlichen Verläufe der Wahrheitswerte. Beim anfänglichen Stehen sind per Konstruktion der Ebenen beide Beine vor

dem Körper, also ergibt sich der Merkmalsvektor $(1,1)$. Bei Schritten mit dem linken Fuß ergibt sich $(0,1)$, bei Schritten mit dem rechten Fuß dagegen $(1,0)$.

Boolesche Ausdrücke in solchen Merkmalssignalen erlauben zusätzliche Interpretationen. So nimmt z. B. die logische Konjunktion der beiden Signale in Abbildung 4 genau dann den Wert eins an, wenn *beide* Füße vorne sind – die Spitzen im Signal treten also immer dann auf, wenn beim Gehen die Beine gerade aneinander vorbeiziehen. Anhand des Abstandes aufeinanderfolgender Signalspitzen kann man leicht die Schrittgeschwindigkeit bestimmen, einen semantisch sehr hochwertigen Parameter.

Als Fazit halten wir fest, dass der Einsatz geometrischer Relationen eine semantisch interpretierbare Vergrößerung von Bewegungen erlaubt, bei der durch geeignete Auswahl der Relationen gewisse Bewegungsaspekte betont und andere Aspekte ausgeblendet werden können.

Adaptive zeitliche Segmentierung

Bis jetzt haben wir einzelne Posen nur *räumlich* vergrößert – die *zeitliche* Dichte der Bewegungsrepräsentation wurde nicht angetastet. Es zeigt sich, dass eine noch kompaktere Beschreibung möglich ist, wenn man aufeinanderfolgende Posen, deren Merkmalsvektoren sich nicht unterscheiden, zu *einem Segment* zusammenfasst. Indem man nur noch die Abfolge der Merkmale der einzelnen Segmente betrachtet, ohne die Segmentlängen zu berücksichtigen, erhält man so etwas wie einen „Fingerabdruck“ der Bewegung, der invariant unter räumlichen und zeitlichen Deformationen ist. Diese Invarianz rührt daher, dass die Segmentgrenzen per Konstruktion durch bewegungsdefinierte Ereignisse bestimmt sind, deren *Reihenfolge* sich bei unterschiedlichen Bewegungsausführungen nicht ändert. Langsames und schnelles Gehen, müdes oder fröhliches Gehen ergeben auf diese Weise alle denselben Fingerabdruck.

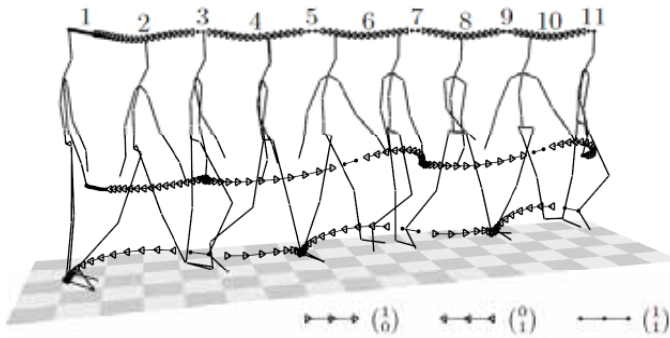


Abb. 5: Adaptive Segmentierung einer Gehbewegung mit den Relationen „rechter Fuß vorne“ und „linker Fuß vorne“.

vorliegt. Dadurch ergeben sich elf Segmente mit der zugehörigen Merkmalsabfolge $(1,1)$, $(0,1)$, $(1,1)$, $(1,0)$, $(1,1)$, $(0,1)$, $(1,1)$, $(1,0)$, $(1,1)$, $(0,1)$, $(1,1)$, die den Prototypen einer Gehbewegung charakterisiert.

Wir zeigen nun eine mögliche Anwendung dieser Techniken zur Gewinnung eines Fingerabdrucks von Bewegungen auf, bei der die beschriebenen Invarianzen eine wichtige Rolle spielen.

Anwendung bei der inhaltsbasierten Suche in Bewegungsdaten

Das Problem der inhaltsbasierten Suche in Bewegungsdaten verdeutlicht man sich am besten aus Sicht eines Benutzers, der eine große Kollektionen von MoCap-Aufnahmen, wie sie z. B. bei der Produktion von Computeranimationsfilmen anfallen, durchsuchen möchte. Bei der manuellen Suche nach einer bestimmten Aufnahme oder einem bestimmten Bewegungstypen steht als einziger Hinweis auf den Inhalt einer Aufnahme häufig nur ein kryptischer Dateiname zur Verfügung. In diesem Fall scheint die Lösung zu sein, jede Aufnahme von Hand zu überprüfen, was aber aus Zeitgründen bei großen Datenmengen nicht machbar ist.

Hier können inhaltsbasierte Suchtechniken weiterhelfen. Vereinfacht ausgedrückt kann ein Computer die Aufgabe übernehmen, sich alle Aufnahmen vorab einmal anzusehen und wichtige Charakteristika der Bewegungen abzuspeichern, um anschließend Suchanfragen schnell beantworten zu können. Man nennt diesen einmaligen Vorgang *inhaltsbasierte Indexierung*. Um mit dem resultierenden Index suchen zu können, müssen Suchanfragen auf geeignete Weise formuliert werden. Häufig bedient man sich hier des „Query-By-Example“-Paradigmas, das es vorsieht, eine Suchanfrage in Form einer kurzen Beispielbewegung zu stellen. Aus dieser Beispielbewegung werden analog zum Indexierungsschritt wichtige Bewegungscharakteristika extrahiert. Anschließend wird der vorab erstellte Index mit einem geeigneten Suchverfahren nach diesen Charakteristika durchsucht. Treffer im Index korrespondieren zu Treffern in der ursprünglichen Datenkollektion und werden als Suchergebnis zurückgeliefert.

Betrachten wir zum Beispiel die Gehbewegung in Abbildung 5. Wir beschränken uns wie im vorhergehenden Beispiel auf die beiden geometrischen Relationen für die Beine. Die 3D-Trajektorien von Kopf, rechter Hand und rechtem Sprunggelenk wurden hier beispielhaft eingezeichnet und mit drei verschiedenen Markierungstypen versehen, je nachdem welche Merkmalsausprägung in der dazugehörigen Pose



Abb. 6: 15 Treffer bei der Suche nach einem Strecksprung (Anfrage ist eingekreist).

Entscheidend für die Qualität der Suchergebnisse ist die Beschaffenheit der Bewegungscharakteristika, die bei der Indexierung und bei der Anfragespezifikation extrahiert werden. Es zeigt sich, dass die oben eingeführten „Fingerabdrücke“ sehr gut für die inhaltsbasierte Suche geeignet sind. Zum Beispiel wurde in Abbildung 6 nach Strecksprüngen gesucht, wozu ein MoCap-Fragment von ca. 2

Sekunden als Anfrage diente. Die resultierenden Treffer enthielten neben Strecksprüngen auch noch weitere, ähnliche Bewegungen. Allerdings ließen sich die Treffer auf automatische Weise derart sortieren, dass die 15 besten Treffer auch wirklich alle in der Datenkollektion enthaltenen Strecksprünge waren. Typische Suchzeiten für derartige Anfragen betragen nur wenige zehntel Sekunden.

Literatur

- Ren, L., Patrick, A., Efros, A. A., Hodgins, J., & Rehg, J. M. (2005). A data-driven approach to quantifying natural human motion. *ACM Transactions on Graphics* 24(3), 1090-1097.
- Liu, C. K., Hertzmann, A., & Popovic, Z. (2005). Learning physics-based motion style with nonlinear inverse optimisation. *ACM Transactions on Graphics* 24(3), 1071-1081.
- Hsu, E., Pulli, K., & Popovic, J. (2005). Style translation for human motion. *ACM Transactions on Graphics* 24(3), 1082-1089.
- Kovar, L., & Gleicher, M. (2004). Automated extraction and parameterization of motions in large data sets. *ACM Transactions on Graphics* 23(3), 559-568.
- Müller, M., Röder, T., & Clausen, M. (2005). Efficient content-based retrieval of motion capture data. *ACM Transactions on Graphics* 24(3), 677-685.

JÜRGEN PERL

Einsatz künstlicher neuronaler Netze zur Mustererkennung im Sport

Kurzfassung

Grundlagen

Künstliche Neuronale Netze vom Typ KFM (Kohonen Feature Map) können Typen, Häufigkeiten und Verteilungen von Mustern lernen sowie, in der von uns entwickelten dynamischen Variante DyCoN (Dynamically Controlled Network), ihre zeitabhängigen Veränderungsprozesse erkennen. Das einzelne Muster kann dabei eine statische Struktur von Item-Werten sein (z. B. bei der Diagnose von Zuständen) oder eine Zeitreihe von Item-Werten (z. B. bei Prozessen aus den Bereichen Strategie, Motorik oder Rehabilitation).

Ein KFM verfügt über sog. Neuronen, die mit Information trainiert werden und dabei Cluster ähnlicher Information bilden können. Die weiterführende Idee von DyCoN ist, dass jedes einzelne Neuron individuell und zeitlich unbegrenzt lernen und Information bereitstellen kann.

Anwendungsbeispiele

Bei der Analyse von Prozessen im Sport, z. B. biomechanische Prozesse im Fall von Bewegungen und strategische Prozesse im Fall von Spielverhalten, besteht häufig das Problem sehr vieler und zum Teil hochpräziser Daten. Hier hilft der Netzansatz, aus einer unüberschaubaren Menge komplexer Daten die relevante Information herauszufiltern und dem Beurteilungs- und Entscheidungsprozess zugänglich zu machen. Insbesondere können so u. a. auch die z. T. subtilen Differenzen motorischer Muster in inter- und intraindividuellen vergleichenden Analysen verdeutlicht werden. Auf diese Weise gelingt es, Bewegungsmuster oder Verhaltensmuster zu erkennen, ihre Wirksamkeit besser einzuschätzen und die so erhaltene Information optimierend in das Training zurückzuführen.

Im vorliegenden Beitrag werden auf diese Weise speziell Lernverläufe analysiert und mit prototypischen Verläufen verglichen. Es zeigt sich, dass es in den Lernverläufen starke individuelle Unterschiede gibt, die aber über Simulation mit Hilfe des physiologischen Adaptationsmodells PerPot parametrisch auf prototypische Grundmuster zurückgeführt werden können.

Neue Paradigmen der Modellbildung

Einführung

Für komplexe Interaktionssysteme, wie sie u. a. in sozialen Gruppen und speziell im Sport auftreten, sind konventionelle Paradigmen wie Determinismus, Präzision und Vollständigkeit für die Formulierung der erforderlichen Handlungsmodelle häufig nicht adäquat. Darüber hinaus sind die aus diesen Paradigmen abgeleiteten

Anforderungen an Berechnungs- und Auswertungsverfahren an einem eher mathematischen Algorithmusverständnis orientiert, dessen Wirksamkeit in komplexen Interaktionssystemen stark eingeschränkt ist.

Dem stehen die aus der Biologie bekannten und wegen ihrer hohen Wirksamkeit auch in technische Wissenschaftsbereiche übernommenen sog. neuen oder unkonventionellen Paradigmen wie etwa Zufälligkeit, Unschärfe und Flexibilität gegenüber. Die daraus abgeleiteten algorithmischen Konzepte ersetzen bzw. ergänzen determiniertes Berechnen durch zielorientiertes Suchen, bestimmen unscharfe Bereiche statt exakte Punkte und versuchen, fehlerhafte Lösungen anzupassen statt sie zu verwerfen. Stellvertretend seien hier mit Genetischen *Algorithmen*, *Fuzzy Logik* und *Neuronalen Netzen* die drei wohl bekanntesten unkonventionellen Algorithmen-Konzepte genannt.

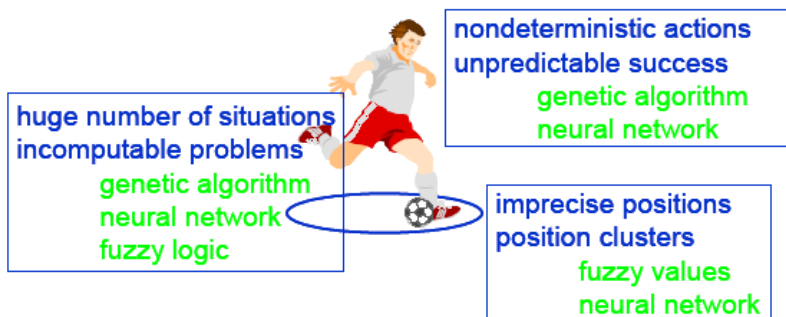


Abb. 1: Probleme der konventionellen Modellierung von Verhaltensprozessen und Einsatzmöglichkeiten unkonventioneller algorithmischer Konzepte.

In Abbildung 1 werden am Beispiel eines Fußballspielers die für eine konventionelle Modellierung problematischen Punkte sowie die jeweils einsetzbaren unkonventionellen Algorithmuskonzepte angeführt.

Von den angesprochenen Konzepten wird im folgenden schwerpunktmäßig das der Neuronalen Netze vorgestellt.

Neuronale Netze

Bei den Neuronalen Netzen kann man ganz grob zwischen den überwachtem lernenden Feed Forward- oder Back Propagation-Netzen und den unüberwachtem lernenden Self Organizing Maps (SOM) oder Kohonen Feature Maps (KFM) unterscheiden:

Beim überwachtem Lernen (s. Abbildung 2, links) wird im wesentlichen eine Input-Output-Relation vorgegebener Trainingsdaten gelernt, wobei die entsprechenden Berechnungen in den Verknüpfungen der inneren Neuronen des Netzes (blau) mit den Eingabeneuronen (rot) und den Ausgabeneuronen (grün) codiert sind. Das Lernen erfolgt i. d. R. über eine bewertete Rückkopplung (blauer Pfeil), die die Übereinstimmung zwischen tatsächlichem und berechnetem Output entsprechend der Güte belohnt (daher auch die Bezeichnung *Reinforcement Learning*).

Beim unüberwachtem Lernen (s. Abbildung 2, rechts) werden die vorgegebenen Datensätze vom Netz selbständig bezüglich ihrer Ähnlichkeit gruppiert und entsprechenden Clustern von Neuronen zugeordnet. Das Ergebnis ist weniger die Berechnung eines Input-Output-Verhaltens als vielmehr die Erkennung struktureller Muster auf der Menge der Datensätze.

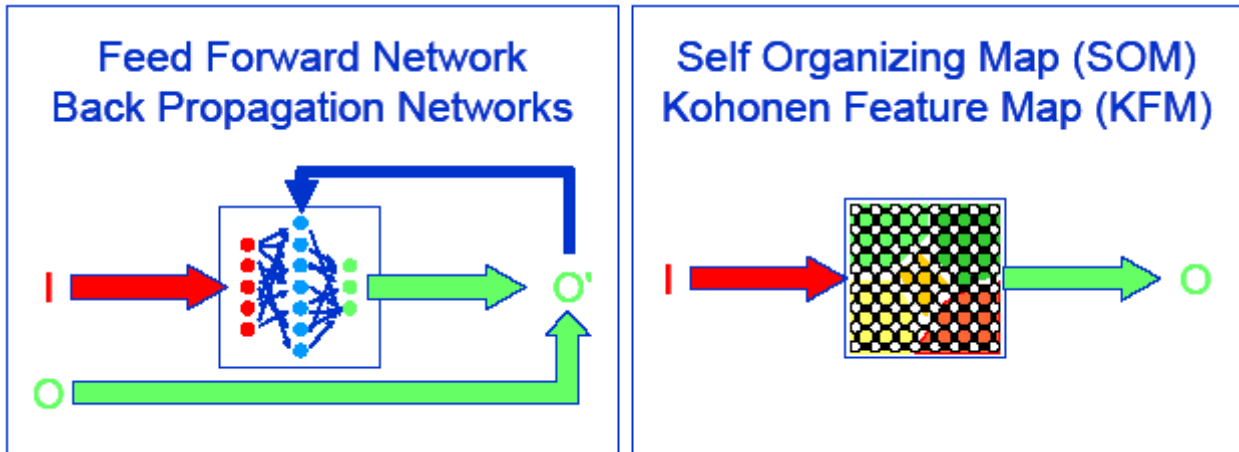


Abb. 2: Netz-Konzepte für überwachtes Lernen (links) und unüberwachtes Lernen (rechts).

KFM: Muster und Trajektorien

Im Beispiel in Abbildung 3 sind Prozesse (z. B. Ballverläufe oder Spielerbewegungen) auf dem Spielfeld als (Daten-)Muster erfasst und dem Neuronalen Netz auftrainiert worden. Das Ergebnis ist, dass Klassen ähnlicher Muster auf Cluster benachbarter Neuronen abgebildet werden. Darüber hinaus repräsentiert jedes einzelne Neuron i. d. R. nicht nur ein einzelnes Muster, sondern einen Muster-Typ, der eine Menge nur leicht variierender Ausprägungen umfasst. Erst dadurch wird es möglich, eine sehr große Zahl von Originalmustern (i. d. R. einige zehntausend) auf eine kleine Zahl von Neuronen (z. B. 400 wie in den weiteren Beispielen) abzubilden.

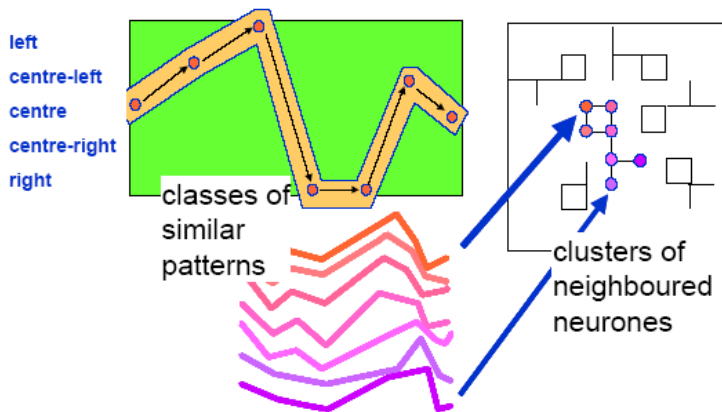


Abb. 3: Zusammenhang zwischen Klassen ähnlicher Muster und den korrespondierenden Clustern benachbarter Neuronen eines KFM.

Präsentieren die einzelnen Datensätze nicht isolierte Zustände, sondern sind Bestandteile eines zeitabhängigen Prozesses, dann ermöglicht die korrespondierende Netz-Repräsentation eine dimensionsreduzierte Visualisierung dieses Prozesses: Jeder Prozesszustand markiert ein korrespondierendes Neuron, die, im entsprechenden zeitlichen Ablauf miteinander verbunden, als Trajektorie den Prozess repräsentieren. Dabei werden

hochdimensionale Zustands-Datensätze (z. B. für Gelenkkoordinaten, Winkel, Geschwindigkeiten usw.) auf die zweidimensionalen Koordinaten der Neuronen reduziert.

Das in Abbildung 4 vorgestellte Ergebnis zeigt für das Beispiel Rudern, wie Bewegungsabläufe durch Trajektorien auf dem Netz repräsentiert und so als zeitabhängige Bewegungsmuster miteinander verglichen werden können: Der generelle

Vergleich der oberen mit der unteren Musterreihe zeigt, dass die Bewegungsmuster von Ruderer A und Ruderer B interindividuell gut übereinstimmen. Die intraindividuelle Analyse zeigt dagegen, dass die Muster von A deutlich stabiler sind als die von B.

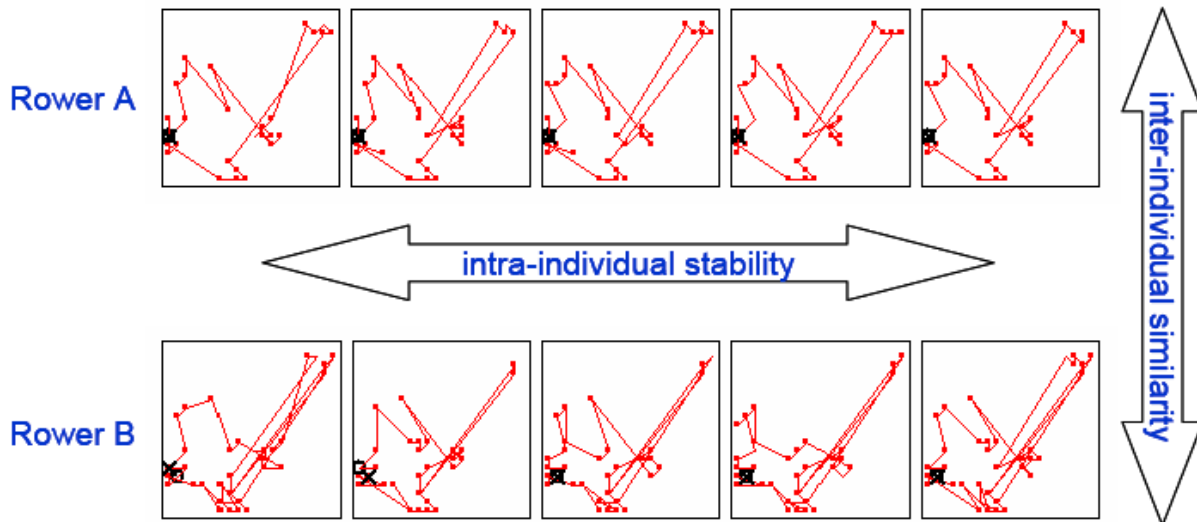


Abb. 4: Abbildung von Bewegungsmustern auf Trajektorien zur Analyse von inter- und intraindividuellen Ähnlichkeiten.

KFM vs. DyCoN

Ein Nachteil des KFM-Ansatzes besteht darin, dass für das Training sehr viele Trainingsdaten benötigt werden (i. d. R. einige zehntausend), die aus der Praxis häufig nicht zur Verfügung stehen. Ein weiterer Nachteil besteht darin, dass sich ein einmal trainiertes Netz in einem "eingefrorenen", durch weiteres Training nicht mehr adaptierbaren Zustand befindet und daher an sich ändernde Musterlandschaften nicht angepasst werden kann.

Das Konzept "DyCoN" (Dynamically Controlled Network) ergänzt das KFM-Konzept durch individuell und dynamisch lernende Neuronen, deren Lernmodell sich an einem Modell physiologischer Adaptation orientiert. Dadurch wird es zum einen möglich, den Lernvorgang beliebig zu unterbrechen und fortzusetzen und so u. a. auch auf Änderungen der Daten- oder Musterlandschaft durch Adaptation zu reagieren. Zum anderen ermöglicht dies, das Netz in zwei Phasen zu trainieren, wobei in der ersten Phase mit stochastisch generierten Daten eine Grundverteilung auf das Netz trainiert wird, die dann in der zweiten Phase mit spezifischen Originaldaten an die reale Musterverteilung angepasst wird. Auf diese Weise wird es mit DyCoN insbesondere möglich, auch auf sehr geringen Datenbeständen netzgestützte Musteranalysen durchzuführen, indem die wenigen Originaldaten in der ersten Phase zur Datengenerierung dienen und erst in der zweiten Phase selbst für die spezifische Prägung verwendet werden. Abbildung 5 zeigt hierzu zwei Beispiele aus den Bereichen Reha und postoperative Erholung, die normalerweise wegen der geringen Zahl an Datensätzen für eine netzgestützte Analyse nicht geeignet sind:

Die Daten aus dem Reha-Bereich (Abbildung 5, links) wurden wöchentlich über 5 bis 10 Wochen erhoben und enthielten jeweils ca. 30 Zustandsdaten. Eine statistische Analyse war auf dieser Datenkonstellation genauso wenig möglich wie eine konventionelle KFM-Analyse. Dagegen zeigt die DyCoN-Analyse Trajektorien, die die zugehörigen Reha-Prozesse transparent und nachvollziehbar machen. Das Beispiel postoperative Erholungsphase (Abbildung 5, rechts) ist hinsichtlich der Datenkonstellation genauso problematisch wie das Reha-Beispiel.

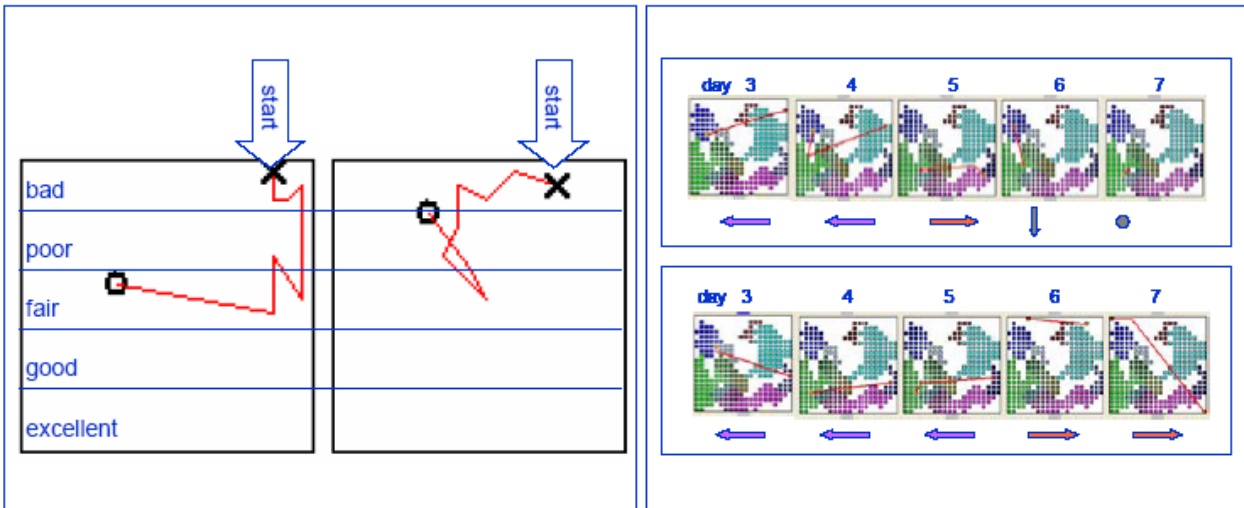


Abb. 5: Reha (links) und postoperative Erholung (rechts) als Beispiele für DyCoN-Analysen von Prozessen mit problematischen Datenkonstellationen.

Zusätzlich besteht das Problem, dass die Messungen täglich zu je drei Zeitpunkten vorgenommen wurden, so dass über 5 Tage je eine Trajektorie vorlag. Hier charakterisieren die Änderungen der Trajektorien über die Tage den jeweiligen Entwicklungsprozess, der im oberen Fall in einen stabilen Zustand mündet, während im unteren Fall am sechsten postoperativen Tag eine massive Destabilisierung eintritt. In beiden Fällen sind auf Grund der charakteristischen Phasenübergänge systematische Prognosen bezüglich des weiteren Prozessverlaufs und damit gegebenenfalls präventive Maßnahmen denkbar.

Fallstudie: DyCoN-Analyse von Lernprozessen

Eine ähnlich problematische Datenkonstellation wie in den oben behandelten Beispielen Reha und postoperative Erholung liegt im Fall der Analyse von Lernprozessen vor, die zusammen mit Daniel Memmert am Institut für Sportwissenschaft an der Uni Heidelberg durchgeführt wird (vgl. den Beitrag von D. Memmert): An drei mehrere Monate auseinander liegenden Messzeitpunkten wurden für jeden Probanden von 3 Ratern jeweils insgesamt 6 Leistungsbewertungen auf einer Skala von 1 bis 10 erhoben. Aus diesen 3 Datensätzen á 6 Bewertungsziffern sollten Aussagen über Lernprozess und Lernerfolg abgeleitet werden. Konventionelle statistische Methoden wie die Berechnung von Mittelwert und Streuung konnten allenfalls inadäquat vergrößerte Ergebnisse liefern, während ein vollständiger Verzicht auf Informationsreduktion einen Vergleich und damit eine Anordnung der 6-dimensionalen Wertetupel verhindert hätte.

Die Fähigkeit des Neuronalen Netzes zur selbständigen ähnlichkeitsorientierten Clusterung erlaubt wie in den oben behandelten Fällen eine Abbildung der Wertetupel auf die Neuronen und damit eine Reduktion der Dimension auf den Wert "2".

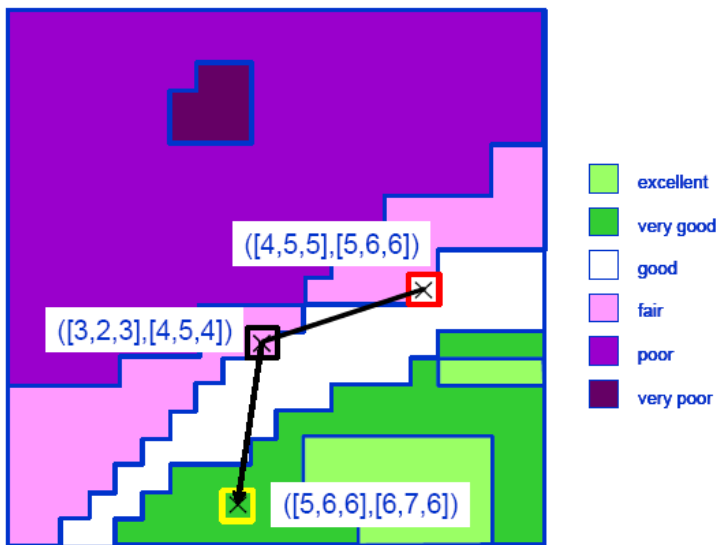


Abb. 6: Kalibriertes Netz vom DyCoN-Typ mit eingetragener Beispiel-Trajektorie, beginnend im rot markierten Neuron und endend im gelb markierten Neuron.

und dritten (gelb) Messzeitpunkt korrespondierten. Die verbindende Trajektorie repräsentiert so den individuellen Lernprozess des Probanden.

Zunächst wurde das Netz vom DyCoN-Typ mit stochastisch generierten Daten vortrainiert, bevor ihm die reale Verteilung mit den Originaldaten aufgeprägt wurde. Anschließend wurde das Netz mit verifizierten Daten kalibriert, so dass Bereiche von Neuronen ähnlicher Bewertung markiert werden konnten (vgl. farbig markierte Netz-Bereiche in Abbildung 6). Schließlich konnten so für jeden Probanden diejenigen Neuronen auf dem Netz markiert werden, die mit dem ersten (rot), zweiten (schwarz)

und dritten (gelb) Messzeitpunkt korrespondierten. Die verbindende Trajektorie repräsentiert so den individuellen Lernprozess des Probanden. Bei einer vergleichenden Analyse über die Probanden und unterschiedliche Lernsituationen stellte sich heraus, dass die Lernprozesse höchst unterschiedlich verliefen und dabei überwiegend nicht dem Schema einer linearen Verbesserung entsprachen. Im Gegenteil bestand eine sehr häufig zu beobachtende Prozessform darin, dass zunächst ein Lernzuwachs erfolgte, der später wieder reduziert wurde, oder dass zunächst ein Verlust auftrat, der später kompensiert wurde (vgl. die winkelförmigen Trajektorien in Abbildung 7).

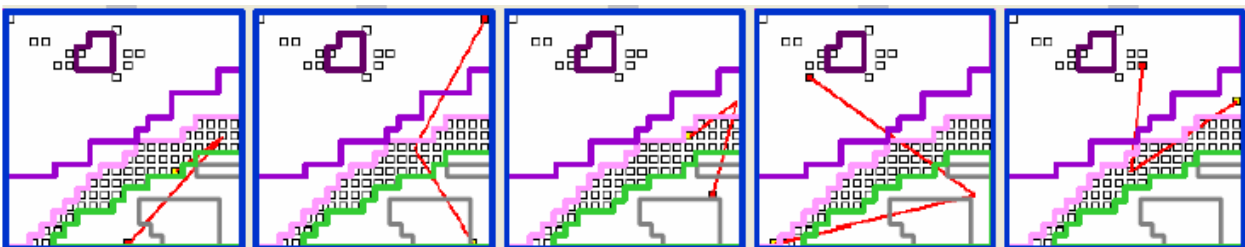


Abb. 7: Beispiel-Trajektorien von Lernprozessen mit überwiegend winkelförmigem Verlauf.

Dieses zunächst so nicht erwartete Ergebnis findet eine mögliche Erklärung im Vergleich mit dem Verlauf physiologischer Adaptationsprozesse: Wie mit PerPot, einem Modell zur Analyse antagonistischer Adaptationsphänomene, per Simulation gezeigt werden konnte, existieren spezifische Formen der reizgesteuerten Anpassung, die im wesentlichen durch das Größenverhältnis zweier zentraler Wirkungsverzögerungen – der Vergess-Verzögerung DS und der Lern-Verzögerung DR in Abbildung 8 – charakterisiert sind.






		DS delay of unlearning	DR delay of learning
$a=b=c$		1.5	1.2
ac		7.0	3.5
$a>b<c$		4.7	7.5
$a\leq b\leq c$		10.5	6.5
$a\geq b\geq c$		3.3	3.5

Abb. 8: Lernverläufe in Abhängigkeit von Vergess- und Lern-Verzögerungen.

Die zweite Spalte in Abbildung 8 symbolisiert den Lernverlauf über die drei Messzeitpunkte; in Spalte 1 sind die jeweiligen Messwerte "a", "b" und "c" in ihrer Größenrelation beschrieben. Die dritte und die vierte Spalte enthalten die zugehörigen Verzögerungswerte für das Vergessen und das Lernen (bezogen auf das Trainings-Zeitraster der Untersuchung), wie sie mit Hilfe von PerPot durch Simulation der Lernverläufe gefunden wurden.

Die abschließende Abbildung 9 zeigt, korrespondierend zu Abbildung 8, links die mit PerPot simulierte Adaptationsverläufe und rechts, zusätzlich überlagert, die entsprechenden Prototypen der Lernverläufe. Dabei bedeuten die roten Punktfolgen jeweils eine konstante Belastungseingabe in das PerPot-Modell, die im vorliegenden Kontext als Lernintensität interpretiert werden kann. Die hellblauen Punktfolgen sind die zeitabhängigen Leistungsausgaben, hier in der Bedeutung von Lernleistung. Die verschiedenen Kurvenverläufe der Lernleistungen repräsentieren unterschiedliche Belegungen mit Verzögerungswerten, die einerseits die Dynamik des PerPot-Modells steuern, und die andererseits im Ergebnis Lerntypen charakterisieren. Diese simulierten Lernverläufe lassen sich nun mit den prototypischen, aus den Tests erhaltenen Lernverlauf-Schemata vergleichen.

Dabei erkennt man einerseits eine recht gute Übereinstimmung zwischen den simulierten Verläufen (links) und den prototypischen Schemata (rechts, dunkelblaue Linien). Daraus leitet sich die Möglichkeit ab, das Ergebnis eines Lernprozesses über eine Analyse der entsprechenden Verzögerungsparameter des Probanden den Lernverlauf individuell vorherzusagen bzw. optimierend zu begleiten.

Es wird aber auch deutlich, dass mehr als drei Messzeitpunkte – das sind in den dunkelblau eingetragenen Schemata jeweils der Startpunkt, der Endpunkt und der "Knick" nach dem ersten Drittel der Zeitskala – für eine solche Prozessanalyse und -optimierung erforderlich sind, wenn individuellen Fluktuationen im Lernverlauf Rechnung getragen werden soll.

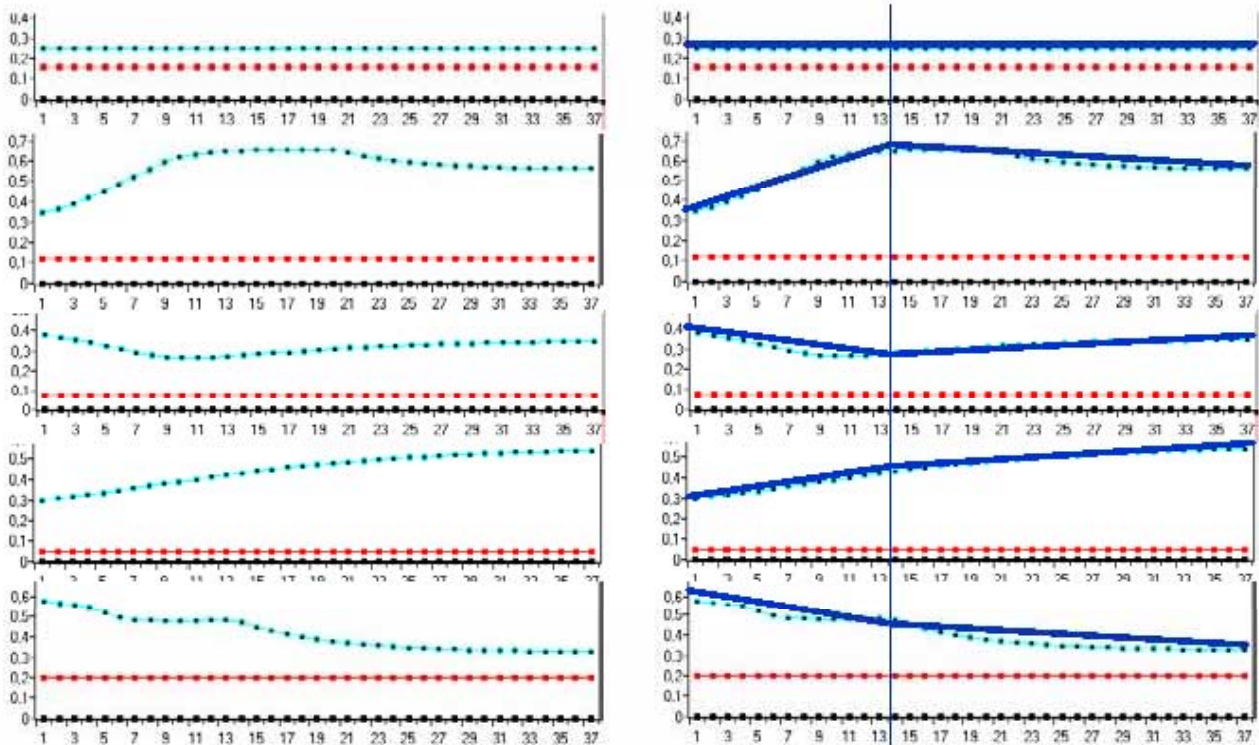


Abb. 9: Vergleich von simulierten und prototypischen Lernverläufen (die Skalenbeschriftungen sind modellbezogen normiert und für das Verständnis ohne Bedeutung).

Zusammenfassung

Es konnte gezeigt werden, dass mit Hilfe von Neuronalen Netzen und modellbasierter Simulation komplexe Adaptationsphänomene, wie sie etwa beim Lernen auftreten, qualitativ analysiert und so gegebenenfalls auch inhaltlich besser verstanden werden können. Insbesondere zeigt sich (nicht ganz unerwartet), dass Lernverläufe

- * über längere Zeiträume nicht linear sind,
- * also auch nicht durch wenige Messzeitpunkte charakterisierbar und prognostizierbar sind,
- * individuell sehr unterschiedlich sein können.

(Für detailliertere Ergebnisse siehe den Beitrag von D. Memmert.)

Neben einer Verbesserung im Hinblick auf ein individuell angepasstes Training könnten diese Ergebnisse nicht zuletzt auch Konsequenzen für Talentsichtung, Talentauswahl und Talentförderung haben.

Weiterführende Literatur

Zum Thema Einsatz von Neuronalen Netzen, DyCoN und PerPot im Sport siehe www.informatik.uni-mainz.de/dycon/

DANIEL MEMMERT

Kennzeichnung inter- und intraindividuelle taktischer Unterschiede mit Hilfe Neuronaler Netze

Einleitung

In diesem Beitrag wird die Re-Analyse eines Datensatzes von Roth und Memmert (2003) hinsichtlich der Charakterisierung inter- und intraindividuelle taktischer Unterschiede im Sportspiel mittels Neuronaler Netze vorgestellt. In dem vom BISP geförderten Feldexperiment wurden dabei in einem längsschnittlich strukturierten Design mit einem vierfachgestuften Treatmentfaktor die beiden abhängigen Variablen *konvergente* und *divergente* taktische Spielfähigkeit generiert. Unter der erstgenannten Denkfähigkeit, der Spielintelligenz, versteht man das optimale Lösen von taktischen Spielkonstellationen im Sportspiel. Mit der zweitgenannten Denkfähigkeit, der Kreativität, ist die überraschende, originelle und flexible Produktion von taktischen Antwortmustern gemeint. Im Folgenden wird zunächst kurz das zugrunde liegende Datenmaterial beschrieben. Danach werden die Ergebnisse zur Kennzeichnung inter- und intraindividuelle taktischer Unterschiede im Sportspiel mit Hilfe Neuronaler Netze vorgestellt. Diese berücksichtigen zunächst die taktische Spielintelligenz und dann die spielerische Kreativität.

Taktische Kreativitätsschulung im Sportspiel

Zielstellung

Im Rahmen eines vom Bundesinstitut für Sportwissenschaft geförderten Projektes (VF 0407/06/12/2001-2002) erfolgte eine Präzisierung und Überprüfung sportartübergreifender vs. sportspielspezifischer Trainingskonzepte, mit einer in dieser Form noch nicht realisierten Akzentsetzung auf die Schulung der *taktischen Spielintelligenz und Kreativität*. In die Auswertung wurden nahezu 100 komplette Datensätze aus elf Gruppen der Ballschule Heidelberg, sechs Sportvereinen sowie zwei Grundschulklassen einbezogen. Das Einstiegsalter wurde relativ konstant gewählt. Im Mittel liegt es bei 7.2 Jahren (SD: 0.9).

Unabhängige Variablen – Treatmentinhalte

Gegenstände der Output-Evaluationen sind verschiedene Treatmentprogramme. Neben dem allgemeinen sportartenunabhängigen Grundlagentraining der Ballschule Heidelberg wird eine fußball- und hockeyspezifische spezifische Schulungskonzeption einbezogen (vgl. ausführlicher Roth & Memmert, 2003). Die Kontrollgruppe besteht aus Schülern der ersten Klassen einer Grundschule. Diese sind nicht Mitglied der Ballschule oder eines Sportvereins und beteiligen sich lediglich an dem „normalen“ Sportunterricht.

Abhängige Variablen

Die abhängigen Variablen bilden die sportspielübergreifende Kreativität und Spielintelligenz. Sie werden an drei Messzeitpunkten bestimmt. Dazu werden mit konzeptorientierten Expertenratings konvergente und divergente Kennziffern in den Spieltestsituationen LÜCKEN AUSNUTZEN und ANBIETEN & ORIENTIEREN erhoben. Unter Spieltestsituationen sind Aufgabenstellungen zu verstehen, bei denen durch die Vorgabe der Spielidee, der Spielerzahlen sowie von Regel- und Umgebungsbedingungen bestimmte taktische Verhaltensweisen zuverlässig und regelmäßig provoziert werden (vgl. Memmert & Roth, 2003). Wesentlich ist, dass durch jeweils zwei Spielerrotationen pro Vp die Positionsverteilungen systematisch verändert werden. Dabei wird nacheinander mit Hand, Fuß und Hockeyschläger agiert. Zur Bewertung der aufgezeichneten taktischen Verhaltensweisen wird ein konzeptorientiertes Expertenrating eingesetzt, das u. a. durch festgelegte Merkmalsdefinitionen und Skalierungen sowie ein vierstufiges Trainingsprogramm gekennzeichnet ist. In umfangreichen Vorstudien wurde dieses Diagnoseinventar hinsichtlich der Klassischen Gütekriterien abgesichert (vgl. Memmert, 2004).

Qualitative Analyse mit DyCoN

Interindividuelle konvergente Unterschiede im Sportspiel

Für die Spielintelligenz in der Basistaktik ANBIETEN & ORIENTIEREN mit der motorischen Ausführungsform Hockeyschläger werden die Ergebnisse mittels Neuronaler Netze des Typs DyCoN in Abbildung 1 dargestellt.

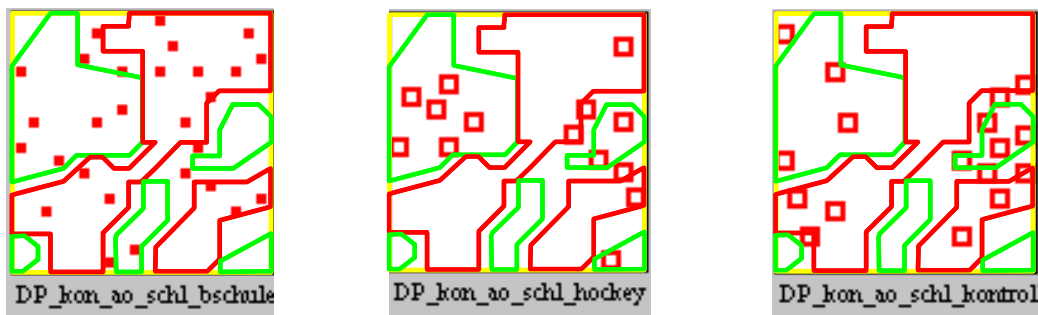


Abb. 1: Kennzeichnung des Lernverhaltens im Basis-Netz (links: Ballschule; Mitte: Hockey; rechts: Kontrollgruppe) bzgl. der konvergenten taktischen Leistung ANBIETEN & ORIENTIEREN mit der motorischen Ausführungsform Hockeyschläger. Die Größe der Neuronen in den Netzdarstellungen entspricht jeweils der Häufigkeit, mit der die korrespondierenden Konstellationen in den Testdaten auftraten.

Dabei wird die Verteilung der Lernerfolge als Farblandschaft auf dem Basis-Netz beschrieben. Jeder einzelne Datensatz wird durch den Test auf diesem Netz einem korrespondierenden Neuron zugeordnet, das seinerseits eine Position in einem der Areale hat und dadurch seine Lage eine qualitative Bewertung erhält. Wie die Resultatsmuster mittels dynamisch-adaptiven Neuronaler Netze zeigen, führt beispielsweise das sportspielübergreifende und hockeyspezifische Trainingsprogramm zur größten Ausbildung der Spielintelligenz, da vermehrte Häufigkeiten in den grünen Bereichen auftreten (vgl. Memmert & Perl, 2005). Im Gegensatz dazu weist die Kontrollgruppe leichte Rückgänge in den konvergenten taktischen Kennziffern auf

(vgl. Abb. 1 rechts). Interessanterweise – trotz gleicher prozentualer Verbesserungen – stellen sich jedoch die taktischen Entwicklungen der Ballschule-Kinder als sehr viel heterogener heraus als die der Hockeykinder (vgl. Abb. 1 links und Mitte). Dies verdeutlicht die Landschaftsstruktur der Neuronalen Netze, die viele kleine Häufigkeitszentren aufweist.

Intraindividuelle divergente Unterschiede im Sportspiel

Abbildung 2 zeigt die Qualitätsbereiche des Neuronalen Netzes für die Kreativität, wobei die Ausprägungen „sehr gut“ (hellgrün), „gut“ (dunkelgrün), „mittel“ (weiß), „schlecht“ (pink), „sehr schlecht“ (violett) und „extrem schlecht“ (dunkelviolet) unterschieden werden.

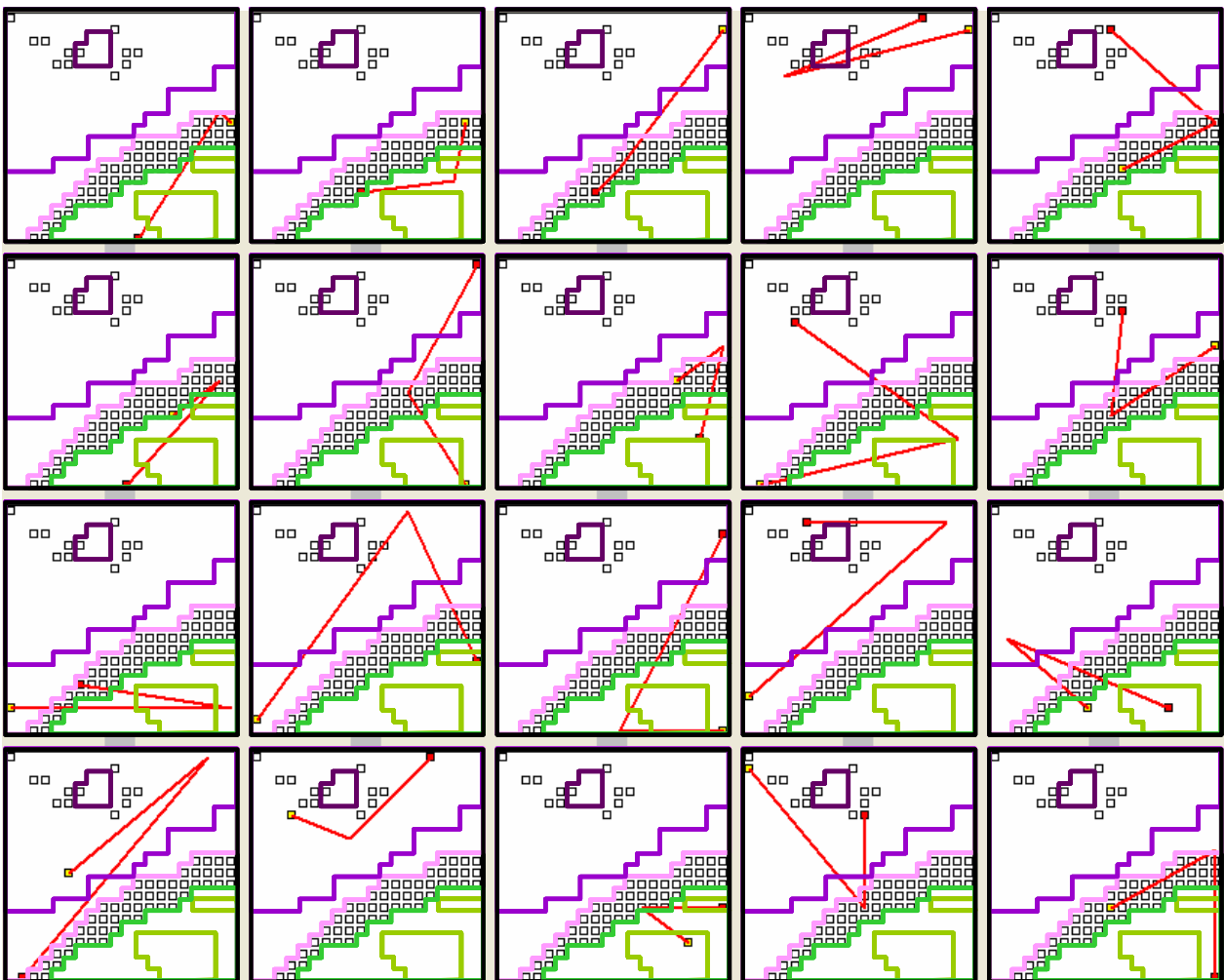


Abb. 2: Repräsentation der intraindividuellen Trajektorien des Fußballtrainings (n=20). Der Lernprozess beginnt im roten und endet im gelben Quadrat (vgl. Memmert & Perl, under review).

In Abbildung 2 sind beispielhaft die Trajektorien dargestellt, die den individuellen Trainingsprozess der Kinder mit dem fußballspezifischen Treatment in ihren divergenten Leistungen in der Basistaktik LÜCKEN AUSNUTZEN mit der motorischen Ausführungsform Hand repräsentieren. Beginnend mit dem roten und endend mit dem gelben Quadrat sind die Zeitschritte des jeweiligen Prozesses als rote Kanten auf dem Netz dargestellt. In den drei Schritten des Prozesses durchläuft die Trajektorie dabei die entsprechenden, farbig markierten Qualitätsbereiche des

Netzes (hellgrün – sehr gut bis dunkelviolett – extrem schlecht). Dabei zeigt die Entwicklung der handballspezifischen Spielkreativität der 20 in Abbildung 2 repräsentierten Fußballspieler über die 15 Trainingsmonate sehr unterschiedliche Ausprägungen (vgl. Memmert & Perl, under review): In 5 von 20 Fällen (25%) steigt die Leistung anfangs, ist aber am Ende dann schlechter als in der Mitte des Trainingsprozesses (up-down-Fluktuations-Prozess). Das umgekehrte Verhalten war bei 30% der Kinder zu beobachten (down-up-Fluktuations-Prozess). In 25% der Fälle stieg die Leistung monoton, während sie in 10% der Fälle monoton fiel. In 10% der Fälle blieb die Leistung (fast) völlig unverändert.

Tab. 1: Zusammenfassung der Ergebnisse über alle Trajektorien der drei Treatmentgruppen. Die fünf verschiedenen Typen von Lernverhalten sind in der zweiten Spalte skizziert. Die Anzahlen der einzelnen Typen für die jeweiligen Ausführungen (Hand, Fuß, Hockeyschläger) sind in Klammern angegeben (vgl. Memmert & Perl, under review).

		Hockeygruppe	Fußballgruppe	Kontrollgruppe
a=b=c		3 (1/0/2)	5 (2/3/0)	4 (1/3/0)
a≤b≤c		10 (1/5/4)	11 (5/2/4)	11 (7/2/2)
a≥b≥c		12 (5/4/3)	8 (2/1/5)	17 (4/6/7)
ac		24 (9/7/8)	11 (5/0/6)	11 (4/4/3)
a>b<c		1 (1/0/0)	17 (6/9/2)	6 (2/3/1)

Tabelle 1 repräsentiert die Gesamtergebnisse der beiden spezifisch trainierten Gruppen und der Kontrollgruppe jeweils in den drei motorischen Ausführungsformen. Interessanterweise zeigt die Hockeygruppe unabhängig von der Motorik jeweils andere Ergebnismuster als die Fußballgruppe. Mittels Chi²-Statistik kann ein signifikanter Effekt der Fluktuations-Prozesse induziert werden ($\chi^2(1) = 18.942$; $p < .001$, $n = 52$). Die Kinder der Hockeygruppe zeigen stärkere up-down Fluktuationen als die der Fußballgruppe, und umgekehrt. Zwischen der Fußball- und der Hockeygruppe gab es im Vergleich zur Kontrollgruppe zu keiner der Fluktuationen signifikante Unterschiede. Zu den restlichen drei Typen des Lernprozesses gab es keine wesentlichen Differenzen zwischen den Gruppen.

Diskussion

Das primäre Ziel dieser Re-Analyse war es gewesen, herauszufinden, ob und wie – bezogen auf das gruppen- und personenbezogene Lernverhalten – Spielintelligenz und Kreativität durch ein standardisiertes Trainingsprogramm verbessert werden kann. Ein derartiger prozessorientierter Vergleich mittels Neuronaler Netze kann helfen, Problembereiche aufzudecken und Gründe für spezifische Auffälligkeiten zu finden. Dabei werden nicht nur grundlagenorientierte Erkenntnisgewinne sondern vor allem der praxisrelevante Nutzen deutlich, der durch den Einsatz Neuronaler Netze entstehen kann.

Im speziellen Beispiel der intraindividuellen Unterschiede in den divergenten taktischen Denkfähigkeiten könnten die unterschiedlichen Lernmuster beispielsweise

dadurch erklärt werden, dass die Trainingsprozesse jeweils durch unterschiedliche Lernverläufe in den jeweiligen Trainingsgruppen geprägt waren. Dabei konnten überraschenderweise fünf verschiedene Lerntypen unterschieden werden. Nur zwei davon zeigten die erwartete unveränderte oder konstant verbesserte Leistung (vgl. Tab. 1 erste und zweite Zeile). Ein dritter Typ zeigte eine konstante Verschlechterung (Zeile 3); die beiden letzten Typen zeigten die oben bereits beschriebene up-down- und down-up-Fluktuation (Zeilen 4 und 5). Dies bedeutet, dass praxisorientierte Trainingsprogramme zu ganz verschiedenen Typen von individuellem Lernverhalten führen können. Mögliche Erklärungen für die down-up- und up-down-Fluktuationen können mit Hilfe des Performance Potential Metamodel PerPot interpretiert werden (vgl. Perl, 2004, Perl in diesem Band), das ursprünglich für die Analyse physiologischer Adaptationsprozesse entwickelt wurde. Der Einsatz von PerPot basiert auf der Idee, dass kreatives Verhalten auch als dynamischer Adaptationsprozess verstanden werden kann (z. B. Boden, 2003; Simonton, 2003).

Literatur

- Boden, M. A. (2003). Computer Models of Creativity. In R. J. Sternberg (Ed.), *Handbook of Creativity* (pp. 351–372). Cambridge: Cambridge University Press.
- Memmert, D. (2004). *Kognitionen im Sportspiel. Einsatz, Transfer und Entwicklung sportspielübergreifender Basistaktiken*. Köln: Sport & Buch Strauß.
- Memmert, D., & Perl, J. (2005, in press). Game Intelligence Analysis by Means of a Combination of Variance-Analysis and Neural Networks International. *Journal of Computer Science in Sport*.
- Memmert, D., & Roth, K. (2003). Individualtaktische Leistungsdiagnostik im Sportspiel. *Spektrum der Sportwissenschaft*, 15(1), 44–70.
- Memmert, D., & Perl, J. (under review). Game Creativity Analysis by Means of Neural Networks.
- Perl, J. (2004). A Neural Network Approach to Movement Pattern Analysis. *Human Movement Science*, 23(5), 605–620.
- Roth, K., & Memmert, D. (2003). Sportspielübergreifende Talentförderung – Taktische Kreativitätsschulung im Sportspiel. In Bundesinstitut für Sportwissenschaft (Hrsg.), *BISp Jahrbuch 2002* (S. 181–190). Bonn: BISp.
- Simonton, D. K. (2003). Scientific creativity as constrained stochastic behavior: The integration of product, person and process. *Psychological Bulletin*, 129(4), 475–494.

ALFRED O. EFFENBERG

Bewegungs-Sonification und multisensorische Integration: Empirische Befunde und zukünftige Perspektiven

Wahrnehmungsseitige Grenzen der Bewegungsregulation

Wodurch ist das motorische Verhaltensrepertoire des Menschen grundsätzlich begrenzt? Zunächst einmal können metabolische und physiologische Faktoren gegenüber motorisch-koordinativen Faktoren abgegrenzt werden – wenn auch damit keine vollständig disjunkte Kategorisierung gelingt. Während limitierende Aspekte von Metabolismus und Physiologie hier weitgehend unberücksichtigt bleiben sollen, liegt das Interesse auf den Charakteristika der informationalen Prozesse, die die Qualität und Vielfalt der Bewegungskoordination begrenzen. Selektivität, Umfang, Genauigkeit und Zeitcharakteristik von perzeptiven, kognitiven und motorischen Prozessen determinieren dabei den Umfang und die Qualität motorischer Repräsentationen bzw. internaler Bewegungskontrollstrukturen und damit auch der Bewegungsregulation. Sollen neue Techniken erlernt werden (motorisches Lernen) oder ist es das Ziel, die Qualität von Bewegungstechniken zu optimieren, stabilisieren oder auch gezielt zu variieren, so ist die Optimierung der perzeptiven und kognitiven Prozesse eine anzustrebende Voraussetzung. Bisher wurden in diesen Zusammenhängen dabei die verschiedenen Sinnesmodalitäten ganz überwiegend separat behandelt, d. h. es wurden Konzepte zur Unterstützung der visuellen oder auch der verbalen Informationsgestaltung weitgehend unabhängig von der auditiven oder kinästhetischen Information realisiert. Ansätze zu einer gezielten Einbindung intermodaler Wahrnehmungssynergien fehlen weitgehend innerhalb der Motorikforschung, ein erster Ansatz dazu wurde von Effenberg (2004) vorgestellt.

Aus verschiedenen Wissenschaftsbereichen kommen Hinweise zur Funktionsweise des Wahrnehmungssystems darauf, dass die einzelnen Sinne keinesfalls isoliert arbeiten. An verschiedenen Orten im ZNS werden die sensorischen Efferenzen der einzelnen Sinnesmodalitäten integriert, und zwar bis hinunter auf die Ebene einzelner Neurone, sogenannter Multisensorischer Konvergenzneurone. Solche verhaltenspsychologischen und neurophysiologischen Erkenntnisse zur multisensorischen Integration¹ im Wahrnehmungssystem des Menschen sollten zukünftig bei Überlegungen zur Optimierung der Bewegungswahrnehmung und der Bewegungsregulation generell Berücksichtigung finden: Denn während auf der einen Seite positive Effekte auf die Verhaltensregulation bei zueinander passenden - konvergenten - visuellen und auditiven Stimuli beschrieben werden, werden ebenso störende Effekte durch divergente Stimuluskonstellationen erzeugt. Auch wenn gegenwärtig noch nicht alle Konvergenz- bzw. Divergenzkriterien bekannt sind, so sollte bei der Entwicklung von Informationskonzepten für die Bewegungspraxis darauf geachtet werden, dass 'Stimuli' bzw. Information unterschiedlicher Modalität

¹ Überblick der Forschung zur Multisensorischen Integration in Calvert, Spence und Stein (2004).

zeitlich, räumlich und auch hinsichtlich der Intensitätsverläufe zueinander passend gestaltet werden. Das gilt vor allem dann, wenn die Information in direkter Weise in die entsprechenden Wahrnehmungs- und Steuerungsprozesse integriert werden soll.

Bewegungs-Sonification

Hier erfolgt die Integration solcher verhaltenspsychologisch-neurophysiologischen Erkenntnisse mit dem Ansatz der datenbasierten Bewegungsakustik, der Bewegungs-Sonification. Die Bewegungs-Sonification zielt nicht nur auf eine umfassendere Nutzung der auditiven Wahrnehmung durch bewegungsakustische Informationen, sondern ebenso auf die zusätzliche Einbindung audiovisueller bzw. multimodaler Integrationsinstanzen². Diese Mechanismen sollen durch neuartige audiovisuelle Instruktions- wie auch Feedbackformen für die Optimierung von Bewegungswahrnehmung, Bewegungsnachvollzug und motorischem Lernen in verschiedenen Bereichen von Sport und motorischer Rehabilitation genutzt werden.

Die Dominanz von visueller und verbaler Information beim Bewegungslernen im Sport und in der Motorikforschung ist offensichtlich: 'Visuomotorik' und 'Visuomotorisches Lernen' sind längst fest etablierte Forschungsbereiche. Doch besitzt auch das Gehör besondere Wahrnehmungseigenschaften, die für die Bewegungswahrnehmung und den Lernzusammenhang interessant sind: Eine besonders hohe zeitliche Auflösungsfähigkeit etwa, die u. a. die Bestimmung der Raumrichtung wahrgenommener Schallereignisse ermöglicht; oder die horizontale Integrationsfähigkeit, die bei der Takt-, Rhythmus- und Melodiewahrnehmung wirksam wird. Allerdings wurde die auditive Wahrnehmung bisher wenig bearbeitet, Beiträge zu 'Audiomotorik' und 'Audiomotorischem Lernen' liegen lediglich in geringer Anzahl vor. Dieser Sachverhalt spiegelt sich auch beim Einsatz unterstützender Informationsverfahren in der Sportpraxis wider: Weder ist gegenwärtig ein auditives Analogon zur vielfach eingesetzten Videotechnik bzw. Bewegungsvisualisierung zu erkennen, noch wird der Bewegungsakustik beim Videoeinsatz hinreichend Bedeutung zugemessen. Die akustische Umsetzung von Bewegungsdaten stellt einen ersten Schritt in diese Richtung dar, sie kann zudem bestehende Messplätze im Leistungssport ergänzen (Effenberg & Mechling, 2004).

Dabei ermöglicht die datenbasierte Bewegungsakustik eine objektivierende und komplexe Abbildung ausgewählter Bewegungsmerkmale in ihrem zeitlichen Verlauf. Es entstehen mehrspurige bewegungsdefinierte Klangsequenzen, die technikspezifische Koordinationsstrukturen repräsentieren. Für das Bewegungslernen im Sport verspricht der Einsatz bewegungsakustischer Zusatzinformationen im Feedback- wie auch im Instruktionzusammenhang eine Präzisierung der Bewegungswahrnehmung und daran anschließender Steuerungs-, Repräsentations- und Lernfunktionen.

² Audiovisuelle und multisensorische Integrationsinstanzen werden unterschieden, da auch trimodale Neurone bei Affen nachgewiesen wurden. Diese erhalten zusätzlich somatosensorischen Input (Stein & Meredith, 1993).

Abbildungen 1-3 zeigen Bewegungs-Sonificationen des Brustschwimmens, die auf der Basis kinematischer Parameter realisiert wurde. Die resultierenden Klangstrukturen sind typisch für das Brustschwimmen, sie repräsentieren z. B. die zeitliche Kopplung von Armzug, Beinstoß und dem Aufrichten des Oberkörpers. Auch individuelle Stilmerkmale werden hörbar, hier z. B. die hohe Zyklusfrequenz und das Fehlen einer Gleitphase. Eine solche Bewegungs-Sonification kann lernunterstützend eingesetzt werden, beispielsweise als audiovisuelles Modell. Die Sonificationen wurden in drei Stufen differenziert; sie können unter dem Link <http://www.sonification-online.com/> unter "Beispiele" heruntergeladen werden.

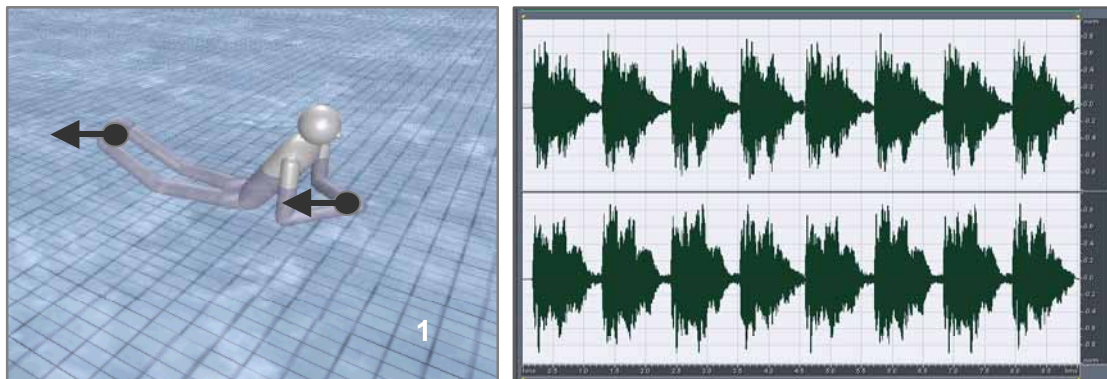


Abb. 1: Horizontale Komponenten der relativen Hand- und Sprunggelenkgeschwindigkeit in der Durchzug- und Stoßphase: Die Parameter werden computerbasiert berechnet und modulieren Klangfrequenz und -amplitude (Amplitudenpegeldiagramm, rechts). 8 Zyklen in 9 sec verweisen auf eine hohe Zyklusfrequenz.

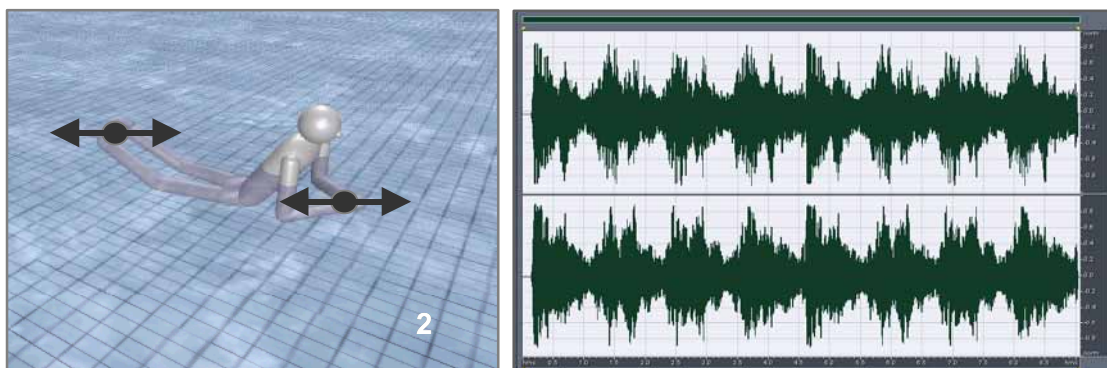


Abb. 2: Horizontale Komponenten der relativen Hand- und Sprunggelenkgeschwindigkeit während des gesamten Zyklus.

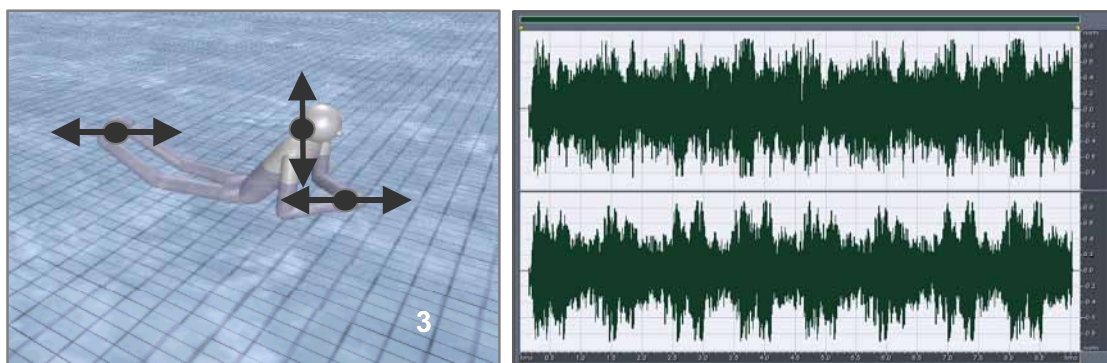


Abb. 3: Horizontale Komponenten der relativen Handgelenk- und Sprunggelenkgeschwindigkeit und vertikale Komponente der relativen Nackengeschwindigkeit modulieren Frequenz und Amplitude von 3 elektronischen Sounds, die sich im Amplitudenpegeldiagramm überlagern (Abbildung modifiziert nach Effenberg, i.p.).

Beim Schwimmen bzw. im Wasser ist eine bewegungssimultane Rückmeldung über die eigene Technikausprägung technisch aufwändig, beim Techniktraining in anderen Sportarten, z. B. beim Turnen, ist die bewegungsakustische Information in Echtzeit technologisch einfacher zu realisieren. Sie stellt auch eine interessante auditive Ergänzung beim Einsatz visueller Modellbewegungen, die instruktiv präsentiert werden, dar.

Empirische Hinweise zur Wirksamkeit bewegungsakustischer Information

Zu den den Lernuntersuchungen vorgelagerten Bereichen 'Wahrnehmung', 'Beurteilung' und 'Nachvollzug' von komplexen Sportbewegungen konnten verschiedene Untersuchungen abgeschlossen werden; die Ergebnisse sind publiziert (Effenberg, 2004; 2005a). Dabei wurden nicht nur deutliche Indizien zur auditiven wie auch audiovisuell-integrativen Wirksamkeit akustischer und audiovisueller Informationen in Form einer deutlich höheren Genauigkeit bei der Beurteilung unmittelbar zuvor wahrgenommener großmotorischer Sportbewegungen gewonnen. Mit einer extern validen Untersuchungsmethode konnten ebenso deutliche Effekte beim direkten Nachvollzug einer unmittelbar zuvor wahrgenommenen Modellbewegung (Hockstrecksprünge unbekannter Höhen) belegt werden. Zudem wurden erste Untersuchungen zur Wirkungsweise dieser Informationsarten realisiert. So konnte gezeigt werden, dass zusätzliche aufgabendienliche Kognitionen die Nachvollzugsgenauigkeit durchaus negativ beeinflussen können (Effenberg, 2005b) oder auch, dass divergente audiovisuelle Stimuluskonstellationen das Nachvollzugsverhalten unbewusst – aber systematisch – modifizieren.

Diese verhaltenspsychologischen Indizien befinden sich in Einklang mit Untersuchungsergebnissen aus dem Bereich der bimodalen Sprachwahrnehmung, die wiederum über hirnhysiologische Messungen zusätzlich unterstützt werden konnten (Calvert, Brammer, & Iversen, 1998; Calvert & Thesen, 2004). Auch im Rahmen unserer bisherigen Untersuchungen wurden erste Messungen im Rahmen einer begleitenden fMRT-Untersuchung abgeschlossen, für die in Kooperation mit der Radiologie der Universität Bonn eine neuartige Messmethode entwickelt werden konnte (siehe dazu den Beitrag von Effenberg, Fehse und Mechling in diesem Band). Eine erste Auswertung der gewonnenen Daten weist u. a. auf eine audiovisuell-integrative Funktion des superioren Temporalsulcus (STSp) hin. So ist der aktuelle Forschungsstand dadurch gekennzeichnet, dass die zeitnahe Wirksamkeit bewegungsakustischer wie auch multisensorisch integrativer Information belegt werden konnte und dass zudem erste Hinweise zur Wirkungsweise der Bewegungs-Sonification vorliegen. Mit nachfolgenden Lernuntersuchungen soll nun die zeitüberdauernde Wirksamkeit dieser Informationsarten überprüft werden, die sich im Rahmen sportlichen wie rehabilitativen Bewegungslernens, Um- und Überlernens wie auch bei der Technikstabilisierung und -optimierung zukünftig vielfältig einsetzen ließe.

Zukünftige Perspektive von Forschung und Interventionspraxis

Generell soll mit der Bewegungs-Sonification das Wahrnehmungssystem bei der Verhaltensregulation in einer erweiterten Form genutzt werden. Mit dem Einsatz zusätzlicher bewegungsakustischer Informationen wird eine Effizienzsteigerung motorischer Lernprozesse und eine Optimierung koordinativ-motorischer Prozesse angezielt. Es gibt allgemeine Hinweise darauf, dass audiomotorische Bewegungsnachvollzüge gar ein Maß an Präzision erreichen können, das bewusst nicht einmal intendierbar ist (Thaut, Tian, & Azimi-Sadjadi, 1998). Diese bewusstseinsunabhängigen direkten audiomotorischen und audiovisuomotorischen Verbindungen könnten zukünftig mit der Bewegungs-Sonification im Leistungssport für eine Unterstützung des motorischen Lernens, des Umlernens, der Technikoptimierung und auch der gezielten Technikvariation genutzt werden, etwa indem die zeitliche Präzision der Bewegungskoordination durch die zusätzliche Einbindung dieser Funktionen gesteigert wird.

Für den Rehabilitationszusammenhang sind – neben diesen präzisierenden Effekten – zwei weitere Aspekte von besonderem Interesse: (1) Die Unmittelbarkeit audiomotorischer und audiovisuomotorischer Verbindungen, die eine Wirksamkeit audiovisueller Instruktions- und Feedbackformen auch unterhalb der Bewusstseinschwelle ermöglichen. Insofern könnten bewegungsakustische Informationen bereits zu einem frühen Interventionszeitpunkt eingesetzt werden und einen sehr direkten Einfluss auf die Bewegungsregulation erzielen. (2) Die Transformation von Daten bzw. Information aus einer anderen Modalität ermöglicht deren partielle auditive Substitution: Es können Informationen zwischen verschiedenen Modalitäten übertragen werden. Eine Perspektive, die bei der Beeinträchtigung einzelner Sinnesbereiche oder auch für den Blindensport in besonderer Weise interessant sein könnte. Technologisch ist die hinreichend genaue Erfassung der Bewegungsparameter in Echtzeit allerdings mit einem gewissen Aufwand verbunden, der gegenwärtig noch einem breiten Einsatz der Bewegungs-Sonification entgegensteht. Erste Pilot-Untersuchungen in der Rehabilitation befinden sich jedoch bereits in der Planungsphase.

Literatur

- Calvert, G. A., Brammer, M. J., & Iversen, S. D. (1998). Crossmodal Identification. *Trends in Cognitive Sciences*, 2(7), 247-253.
- Calvert, G. A., Spence, C., & Stein, B. E. (Eds.). (2004). *The handbook of multisensory processes*. Cambridge, Massachusetts: MIT Press.
- Calvert, G. A., & Thesen, T. (2004). Multisensory integration: methodological approaches and emerging principles in the human brain. *Journal of Physiology - Paris*, 98(1-3), 191-205.
- Effenberg, A. O. (2004). *Synergien der Sinne für die Bewegungsregulation*. Frankfurt a. M.: Peter Lang.
- Effenberg, A. O. (2005a). Movement Sonification: Effects on Perception and Action. *IEEE Multimedia, Special Issue on Interactive Sonification*, 12(2), 53-59.

- Effenberg, A. O. (2005b). Denken oder Tun? Effekte visuell-akustischer Bewegungsdemonstrationen und bewusster Reflexionen auf die Nachvollzugsgenauigkeit von Sportbewegungen. In H. Gabler, U. Göhner & F. Schiebl (Hrsg.), *Zur Vernetzung von Forschung und Lehre in Biomechanik, Sportmotorik und Trainingswissenschaft*. Hamburg: Czwalina, 106-112.
- Effenberg, A. O. (i. p.): Enhancing Motor Control and Learning by Additional Movement Sonification. In Hunt, A. & Leman, M. (eds.), *Applications of Gestural Audio Systems, Part C: Medical & Social Applications*.
- Effenberg, A. O., & Mechling, H. (2004). Perspektiven der direkten bewegungsbezogenen Vermittlung komplexer Bewegungsdaten. In J. Krug & H.-J. Minow (Hrsg.), *Messplatztraining*. Sankt Augustin: Academia, 122-127.
- Stein, B. E., & Meredith, M. A. (1993). *The Merging of the Senses*. Cambridge: MIT Press.
- Thaut, M. H., Tian, B., & Azimi-Sadjadi, M. R. (1998). Rhythmic finger tapping to cosine-wave modulated metronome sequences: Evidence of subliminal entrainment. *Human Movement Science*, 17(6), 839-863.

ANDREAS WEBER

Rekonstruktion dynamischer Daten aus kinematischen Daten für die Sonification: Schwerpunktstrajektorien, Summen einwirkender Kräfte und Inverse-Dynamik-Berechnungen

Schwerpunktstrajektorien und Summen einwirkender Kräfte

Die Erfassung von Bewegungsdaten mit optischen Motion-Capture-Systemen ist ein Standard-Verfahren zur Erfassung von Bewegungsdaten in der Computeranimation, für die Spiele-Industrie und in den Sportwissenschaften. Es existieren Verfahren, um aus den zunächst erfassten 3D-Markerdaten Skelettrepräsentationen zu extrahieren, die die Grundlage für Visualisierungen der Bewegungsdaten in Animationssystemen liefern.

Diese Bewegungsdaten liegen zunächst rein in einer kinematischen Form vor, die gleichen Bewegungsdaten könnten zu verschiedenen Personen mit gleicher Skelettgeometrie aber unterschiedlichen Körpergewichtes oder Masseverteilung gehören. Mittels anthropometrischer Tabellen lassen sich jedoch die Masseverteilungen der einzelnen Körpersegmente aus Gesamtgewicht (und Körpergröße) eines Menschen abschätzen (Miller & Morrison, 1975; Robbins & Wu, 2003).

Bei bekannter Masseverteilung der Körpersegmente lässt sich aus den kinematischen Motion-Capture-Daten die Trajektorie des Körperschwerpunktes leicht berechnen, und aus dieser die *Summe* der auf alle Segmente einwirkenden *äußeren Kräfte*, da ganz allgemein bei einem physikalischen System die Schwerpunktstrajektorie durch die Summe der externen Kräfte bestimmt ist.

Ein Nachweis der Wirksamkeit der Sonification zur Effizienzsteigerung von Bewegungswahrnehmung, -beurteilung und -nachvollzug auf Basis *gemessener externer Kräfte* (Bodenreaktionskräfte beim Hock-Streck-Sprung) wurde u. a. von Effenberg erbracht (Effenberg, 2005), so dass eine Rekonstruktion der Summe der einwirkenden externen Kräfte relevant ist.

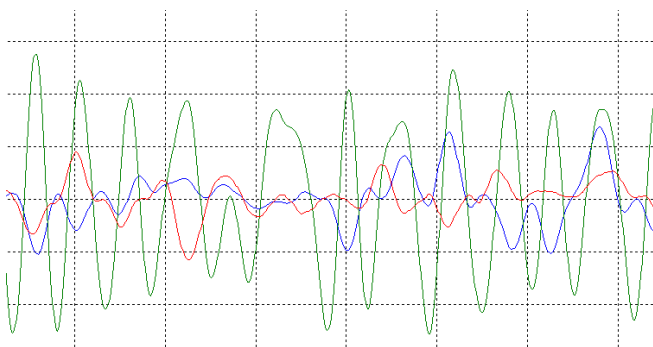


Abb. 1: Rekonstruierte Beschleunigungen des Schwerpunkts bei Boxbewegung (Bewegung 17_05; grün: vertikal, rot und blau: horizontal): Die Phasen, in denen beide Füße kurz den Boden verlassen, sind gut von den Phasen unterscheidbar, bei denen noch ein Bodenkontakt besteht.

(Bodenreaktionskräfte beim Hock-Streck-Sprung) wurde u. a. von Effenberg erbracht (Effenberg, 2005), so dass eine Rekonstruktion der Summe der einwirkenden externen Kräfte relevant ist.

Um die Robustheit der rekonstruierten externen Kräfte gegenüber dem Einfluss der Masseverteilungen abschätzen zu können, wurden einige Simulationen durchgeführt, zum Teil mit Extrembeispielen. Die Beispielbewegungen wurden der CMU-Mocap-Datenbank entnommen: mocap.cs.cmu.edu/.

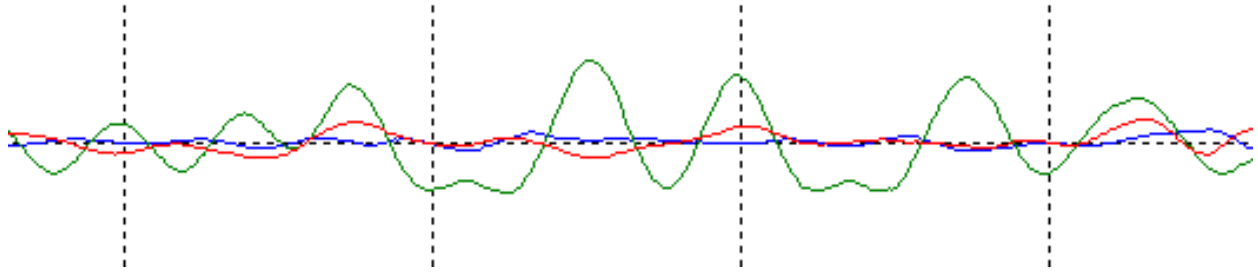


Abb. 2: Rekonstruierte Beschleunigungen des Schwerpunkts bei Tanzbewegung (05_16; grün: vertikal, rot und blau: horizontal) mit 3 Sprüngen.

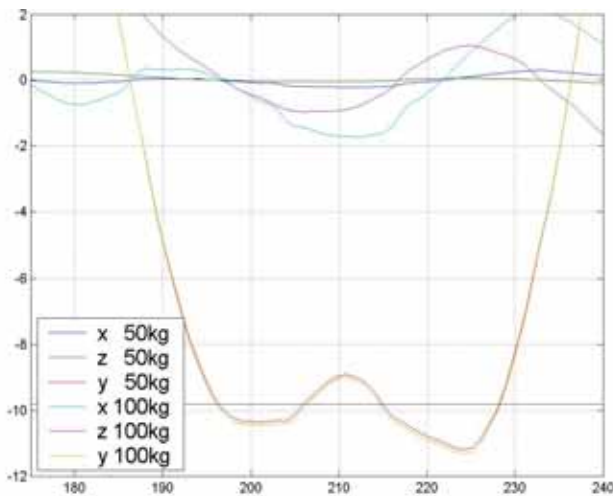


Abb. 3: Ausschnitt aus rekonstruierten Beschleunigungen des Schwerpunkts bei Tanzbewegung (05_16) bei stark variierendem Körpergewicht (50 kg und 100 kg).

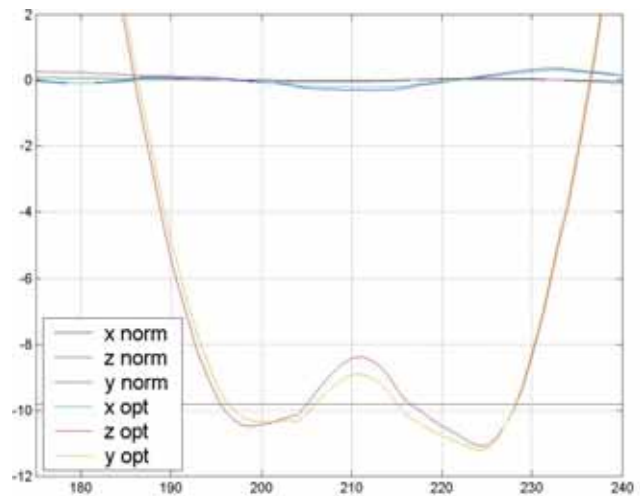


Abb. 4: Ausschnitt aus rekonstruierten Beschleunigungen des Schwerpunkts bei Tanzbewegung (05_16): Vergleich mit „optimierter Masseverteilung“ bei einem Sprung.

Abbildung 4 zeigt das Ergebnis einer Optimierung der Masseverteilung bezüglich Konstanz (und Anpassung der y-Achse an Senkrechte) während eines Sprunges.

Durch Änderung der Masseverteilung beim Mehrkörpermodell kann die Beschleunigung des Schwerpunktes also nur ungenau an die theoretische Vorhersage angepasst werden. Um eine größere Genauigkeit zu erreichen, ist der Rückgriff auf ein Modell für Gewebe-Deformation notwendig, zumindest in einer rudimentären Form.

Inverse Dynamik-Berechnungen

Auch Dynamik-Daten der einzelnen Segmente sind von Interesse. Diese können bei Mehrkörpersystemen grundsätzlich effizient berechnet werden, wenn die externen Kräfte bekannt sind und die Segmentmassen, auf die die externen Kräfte einwirken, und das Problem geschlossener Segmentschleifen nicht vorkommt oder durch eine gute Heuristik gelöst werden kann. In diesem Fall involvieren inverse Dynamik-Berechnungen Einwärts- und Auswärtsrekursionen von Endeffektoren zu einem Wurzelsegment, die linear in der Anzahl der Segmente sind und nur wenige arithmetische Operationen pro Segment benötigen, so dass Berechnungen in

Realzeit problemlos möglich sind (Effenberg, Melzer, Weber, & Zinke, 2005; Stepanenko & Vukobratovic, 1976).

Beispiele von Sonificationen von Momenten, die durch inverse Dynamik-Berechnungen aus kinematischen Daten rekonstruiert wurden, finden sich unter

<http://cg.cs.uni-bonn.de/project-pages/sonification/>.

Diese Algorithmen zur inversen Dynamik berechnen Momente auf Segmenten an abstrahierten Ansatzpunkten und nur die Gesamtmomente an Segmenten. Eine Rekonstruktion „realer“ Muskelkräfte ist wesentlich aufwändiger, und es besteht auch das Problem, dass mehrere Muskeln gemeinsam an einem Segment angreifen können.

Rechenzeiten für eindeutige Rekonstruktionen von Muskelkräften sind wesentlich höher als für die grundlegenden inversen Dynamik-Berechnungen. Silva & Ambrosio (2003) geben etwa 45 sec Rechenzeit an, um zu einer eindeutigen Rekonstruktion von Muskelkräften bei einer Bewegung von 66 Frames Länge zu gelangen. (Diese Rekonstruktion basiert auf einem Mehrkörpersystem.)

Schlussfolgerungen

Eine Rekonstruktion dynamischer Daten aus kinematischen Daten ist grundsätzlich möglich. Auf der Basis von Mehrkörpersystemen kann dies auch jetzt schon in Echtzeit erfolgen, bei komplexeren und verfeinerten Modellen ist dies aber bislang noch nicht möglich.

Da sowohl eine Verfeinerung von Modellen von Mehrkörpersystemen für Sonifications-Zwecke wohl notwendig ist als auch ein echtzeitfähiges System für viele Anwendungen, ist eine enge Zusammenarbeit zwischen der Informatik und den Sportwissenschaften notwendig, um diesen teilweise komplementären Zielen gerecht werden zu können. Dabei ergibt sich als zentrale Fragestellung für die Informatik:

- * Welche Verfeinerungen des Modells sind bei Echtzeitsystemen möglich? (Insbesondere bei Verallgemeinerungen von Mehrkörpersystemen, Berücksichtigung von „aktiven Muskelkräften“ vs. „passiven Spannungs Kräften“, etc.)

Diese Fragestellung muss aber zusammen mit der folgenden betrachtet werden, die insbesondere durch empirische Untersuchungen aus dem Bereich der Sportwissenschaften behandelt werden sollte:

- * Welche Verfeinerungen des Modells sind zur Unterstützung des motorischen Lernens notwendig?

Danksagung

Mit besonderem Dank an meine Diplomanden Björn Krüger und Joachim Melzer, sowie an meinen Mitarbeiter Arno Zinke.

Referenzen

- Effenberg, A. O. (2005). Movement Sonification. Effects on Perception and Action. *IEEE Multimedia, Special Issue on Interactive Sonification*, 12(2), 53-59.
- Effenberg, A. O., Melzer, J., Weber, A., & Zinke, A. (2005). MotionLab Sonify: A Framework for the Sonification of Human Motion Data. *Proceedings of the 2nd International Conference on Non-visual & Multimodal Visualization*. London, UK: IEEE Press.
- Miller, D. I., & Morrison W. (1975). Prediction of segmental parameters using the Hanavan human body model. *Med. Sci. Sports*, 7(3), 207-212.
- Robbins, K. L., & Wu, Q. (2003). Development of a computer tool for anthropometric analyses. In F. Valafar & H. Valafar (Hrsg.), *Proceedings of the International Conference on Mathematics and Engineering Techniques in Medicine and Biological Sciences (METMBS'03)* (pp. 347-353). Las Vegas, USA: CSREA Press.
- Silva, M. P. T., & Ambrosio, J. A. C. (2003). Solution of redundant muscle forces in human locomotion with multibody dynamics and optimization tools. *Mechanics Based Design of Structures and Machines*, 31(1), 381-411.
- Stepanenko, Y., & Vukobratovic, M. (1976). Dynamics of Articulated Open-chain Active Mechanisms. *Math. Biosciences*, 28, 137-170.

FRANK KURTH

Sonificationssysteme

Zusammenfassung

Die computergestützte Generierung akustischer (Musik-) Signale stellt eine der ältesten Teildisziplinen auf dem Gebiet der Computermusik dar. Ein klassisches Ziel ist dabei die künstliche Nachbildung instrumenten- und damit realitätsnaher musikalischer Noten und deren spezifischer Klänge. In diesem Beitrag sollen zunächst einige Basismethoden zur Realisierung solcher künstlichen, in verschiedenen Dimensionen parametrisierbaren Tonerzeugung vorgestellt werden. Darauf aufbauend werden verschiedene Ansätze diskutiert, wie unter Verwendung dieser Basismethoden zur Klang- und Tonerzeugung komplexe Datenströme in akustische Signale überführt, also sonifiziert, werden können. Eine Herausforderung ist dabei, dass die betrachteten Datenströme, wie im Beispiel der akustischen Bewegungstransformation, im allgemeinen nicht musikalischen Ursprungs sein müssen, was gewisse Anforderungen an die Sonification stellt. Zum Abschluss werden einige ausgewählte, in Projekten der Arbeitsgruppe Multimedia-Signalverarbeitung entstandene Softwaremodule genannt und deren Anwendbarkeit im Bereich Sonification skizziert.

Sonification

Historisch gesehen stellt die Arbeitsdefinition *'Sonification is a mapping of numerically represented relations in some domain under study to relations in an acoustic domain for purposes of interpreting, understanding, or communicating relations in the domain under study'* von Carla Scaletti aus dem Jahre 1992 einen der frühen Spezifikationsversuche der Sonifikationsaufgabe dar. Aus informatischer Sicht geht es bei der Sonification darum, gegebene zeitabhängige Eingangsdaten in akustische Ausgangsdaten umzuwandeln. Die Eingangsdaten können dabei zeitlich kontinuierlich oder diskret gegeben sein. Im diskreten Fall unterscheidet man weiterhin zwischen äquidistant abgetasteten, also in gleichen zeitlichen Abständen am System eintreffenden, und zeitlich nicht äquidistanten, also nur zu bestimmten Ereigniszeitpunkten eintreffenden Messdaten. Aus Sicht der Signalverarbeitung lässt sich die Sonifikationsaufgabe dann im allgemeinsten Fall in die beiden Schritte der Signalaufbereitung oder -vorverarbeitung und der eigentlichen Klangerzeugung unterteilen. Ein klassischer Versuch zur Spezifikation unterschiedlicher Sonifikationsarten unterteilt weiterhin den Schritt der Klangerzeugung in eine Folge von Klanggeneratoren. Hierbei legt die Anzahl der hintereinandergeschalteten Klanggeneratoren die Komplexität der Sonification fest. Werden die eventuell aufbereiteten Eingangsdaten unmittelbar akustisch wiedergegeben, spricht man von Audifikation. Schaltet man hingegen einen Klanggenerator vor die Ausgabe, spricht man von Sonification (erster Ordnung). Schaltet man hinter den ersten Klanggenerator einen zweiten, so dass die Ausgabe des ersten Klanggenerators die Klangerzeugung mittels des zweiten Generators steuert, spricht man von Sonification zweiter

Ordnung. Eine Reihenschaltung von Klanggeneratoren liefert entsprechend die Sonification N -ter Ordnung.

Diese sehr allgemeine Fassung einer Sonificationsaufgabe führt zu einer Vielzahl praxisrelevanter Fragestellungen im Zusammenhang mit dem Anforderungsprofil eines Sonificationssystems:

- * Wieviele Eingangsdatenströme sind simultan zu sonifizieren?
- * Welche Extraktions- und (Vor-) verarbeitungsmethoden werden benötigt?
- * Welche Methoden der Audio- und Klangerzeugung werden benötigt und welche stehen zur Verfügung?
- * Wie wichtig ist eine perzeptuell ansprechende Qualität des Ausgangssignals?
- * Wie wichtig ist die Echtzeitfähigkeit des Sonificationssystems?

Während die Fragen nach der Anzahl zu verarbeitender Eingangsdaten, geeigneten Datenverarbeitungsmethoden, der perzeptuellen Qualität des Ausgangssignals sowie der Echtzeitfähigkeit des Sonificationssystems stark von der jeweiligen Anwendung abhängen, gibt es aus dem Bereich der Audiosignalverarbeitung allgemeine Methoden und Prinzipien zur Audio- und Klangerzeugung, die im weiteren etwas genauer dargestellt werden sollen.

Signalverarbeitungsmethoden zur Klangerzeugung

Eines der wichtigsten klassischen Ziele der Klangerzeugung aus Signalverarbeitungssicht ist die realistische Modellierung und künstliche Erzeugung von Instrumentenklängen. Dies betrifft sowohl die Generierung isolierter Töne, simultaner Töne, Töne in bestimmter Artikulation, sowie den Zusammenklang mehrerer musikalischer Stimmen und Instrumente. Drei der wichtigsten Paradigmen der Klangerzeugung, die im folgenden im Hinblick auf deren Anwendbarkeit in der Sonification skizziert werden, sind *parametrische Modelle*, *samplebasierte Synthese* und *Physical Modelling*.

Parametrische Modelle

Die Idee der bereits seit einigen Jahrzehnten bekannten parametrischen Modelle besteht darin, dass Instrumentenklänge oft schon als Überlagerung nur weniger einfach strukturierter Basistöne gut approximiert werden können. Dies gilt etwa für Blasinstrumente, die einen stark harmonischen Klangcharakter besitzen. Abb. 1 zeigt als Beispiel die in einem auf einer Flöte gespielten, wenige Sekunden andauernden Ton enthaltenen Frequenzkomponenten, die bei Vielfachen von ca. 261 Hertz liegen. Diese Zeit-Frequenzsicht legt eine Modellierung von Instrumentenklängen durch die additive Überlagerung von K Basistönen bestehend aus parametrisierten Sinusoid-Funktionen nahe.

Im sinusoiden Modell ist der k -te solche Basiston von der Form $a_k(t)\sin(2\pi f_k t + p_k(t)) + r_k(t)$, wobei $a_k(t)$ und $p_k(t)$ die (zeitabhängigen) Amplituden- und Phasenverläufe, f_k die Frequenz des Tons und $r_k(t)$ ein (ebenfalls zeitabhängiger) Rauschanteil ist.

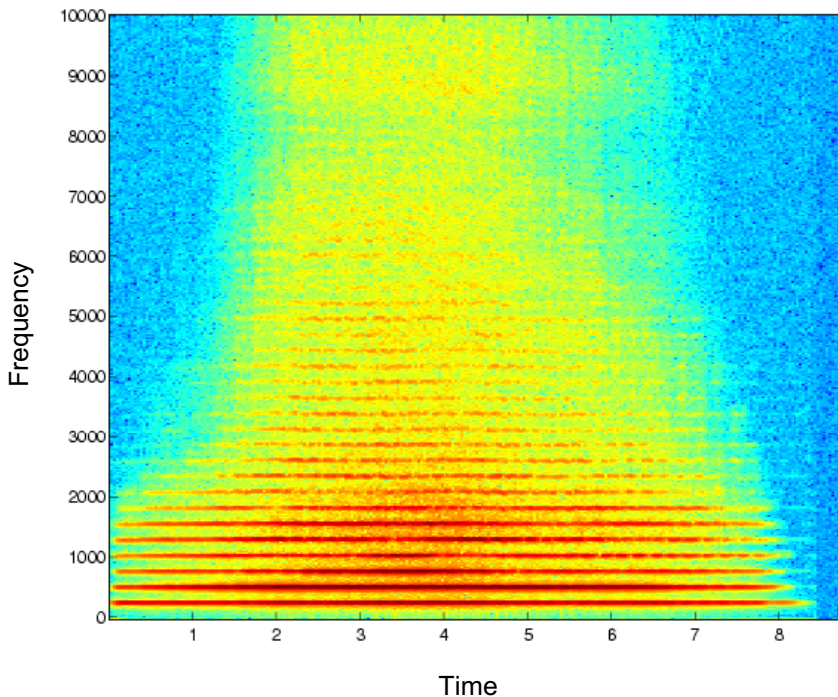


Abb. 1: Zeit-Frequenz-Darstellung des Tons c1 gespielt auf einer Flöte.

Die Vorteile der sinusoiden Modellierung sind die geringe Anzahl der Parameter und die damit verbundene einfache Realisierbarkeit, die effiziente Implementierbarkeit und die einfache Möglichkeit zur Modulation von Klangparametern wie Tonhöhe oder Tondauer. Die wichtigsten Nachteile dieser Art der Klangerzeugung sind die Einschränkung, dass ohne weiteres nur Klänge mit tonalem Charakter modelliert werden können und dass die erzielte Klangqualität oft nicht wirklichkeitsgetreu ist.

Samplebasierte Synthese

Der Ansatz der samplebasierten Synthese besteht darin, bereits auf realen Instrumenten erzeugte Töne zur Klangerzeugung zu verwenden. Hier wird vorab eine (Sample-) Bibliothek aller möglichen Instrumententöne aufgenommen. Zur Klangerzeugung werden dann die benötigten Töne anhand einer Klangerzeugungsvorschrift – im Musikbereich im allgemeinen eine Partitur – der Bibliothek entnommen und in einer aus der Vorschrift abgeleiteten zeitlichen Reihenfolge überlagert. Als Beispiel einer kommerziellen Samplebibliothek sei die *Vienna Symphonic Library* (<http://www.vsl.co.at/>) genannt, die mit über 1.5 Mio. Einzelaufnahmen von Orchesterinstrumenten eine der weltgrößten Sammlungen von Samples darstellt. Diese ca. 1 Terabyte große Datensammlung umfasst unter anderem Aufnahmen aller Orchesterinstrumente für Noten aller Tonhöhen, alle musikalisch relevanten Intervallübergänge, eine Vielzahl unterschiedlicher Lautstärken, unterschiedliche Tempi und unterschiedliche Artikulationen. Zur klavierunterstützten Synthese steht ein ‚Performance-Tool‘ zur Verfügung. Weiterhin werden auf Mac- und PC-Seite gängige Synthesizer-Softwaresysteme unterstützt.

Die Vorteile der samplebasierten Synthese sind die sehr hohe Realitätstreue und Qualität der erzeugbaren Klänge und eine sehr hohe Effizienz. Nachteilig ist, dass eine Modulation von Parametern wie Tonhöhe und Tondauer oft algorithmisch schwierig oder nicht in sehr guter Qualität möglich ist. Ein Ausweg besteht hier

darin, die Töne bei der Erzeugung der Bibliothek in entsprechend vielen verschiedenen Spielweisen aufzunehmen. Hierin besteht dann ein Trade-off zwischen der Anzahl der zu erfassenden Töne und den verbrauchten Speicher- und Zeitressourcen. In heutzutage üblichen PC-Soundkarten wird allgemein die samplebasierte Synthese verwendet, wobei eine für die meisten Anwendungen ausreichende Klangqualität erzielt wird.

Physical Modelling

Der Ansatz des Physical Modelling (Smith, 2004) besteht darin, ein mathematisches Modell der physikalischen Schallerzeugung im Instrument zu bilden. Dies führt im allgemeinen auf ein komplexes System mathematischer Gleichungen, dass oft nur numerisch gelöst werden kann. Neben der Erzeugung einzelner Instrumententöne in unterschiedlichen Artikulationen können mit Methoden des Physical Modelling auch komplexe akustische Szenarien modelliert werden. Ein Beispiel ist ein Feedback-Szenario, wo der Klang einer E-Gitarre über einen Lautsprecher wiedergegeben wird und so schließlich wiederum die vor dem Lautsprecher platzierte E-Gitarre zur Schwingung anregt.

Das Physical Modelling ermöglicht die Erzielung sehr guter, realitätsnaher Klangqualität, erlaubt die Modellierung auch exotischer Instrumente und ermöglicht zudem eine Modulation der Tonparameter durch Triggerung der Modellparameter. Zu den Nachteilen dieses Ansatzes zählen der aus theoretischer Sicht hohe Anspruch an die Modellierung, der Expertenwissen aus den verschiedenen involvierten Bereichen erfordert, sowie der eventuell sehr hohe Zeitaufwand für anfallende numerische Berechnungen.

Sonificationstypen

In diesem Abschnitt soll kurz skizziert werden, wie in Abhängigkeit von den Eingangsdaten eine Sonification mit Hilfe der Klangerzeugungsmethoden erfolgen kann. Die Eingangsdaten sind, nach etwaiger Vorverarbeitung, häufig entweder als kontinuierlicher Messdatenstrom oder als Folge isolierter Ereignisse gegeben. Ein kontinuierlicher Messdatenstrom könnte etwa die Ortskoordinaten einer mit Motion-Capture-Methoden erfassten menschlichen Gliedmaße beinhalten, während ein Strom isolierter Ereignisse etwa Auskunft darüber geben könnte, wann und an welchem Ort die Füße eines Probanden bei einer menschlichen Laufbewegung auf dem Boden auftreffen. Entsprechend dieser beiden charakteristischen Arten von Eingangsdaten unterscheiden wir zwischen *Sonification durch Modulation* und *ereignisgesteuerter Sonification*.

Sonification durch Modulation

In parametrischen Ansätzen zur Klangerzeugung besteht ein kanonischer Ansatz zur Klangerzeugung darin, einzelne Modellparameter direkt durch die eventuell vorverarbeiteten Eingangsdaten zu ersetzen. Genauer gesagt werden hier die zeitlich veränderlichen Parameter ersetzt.

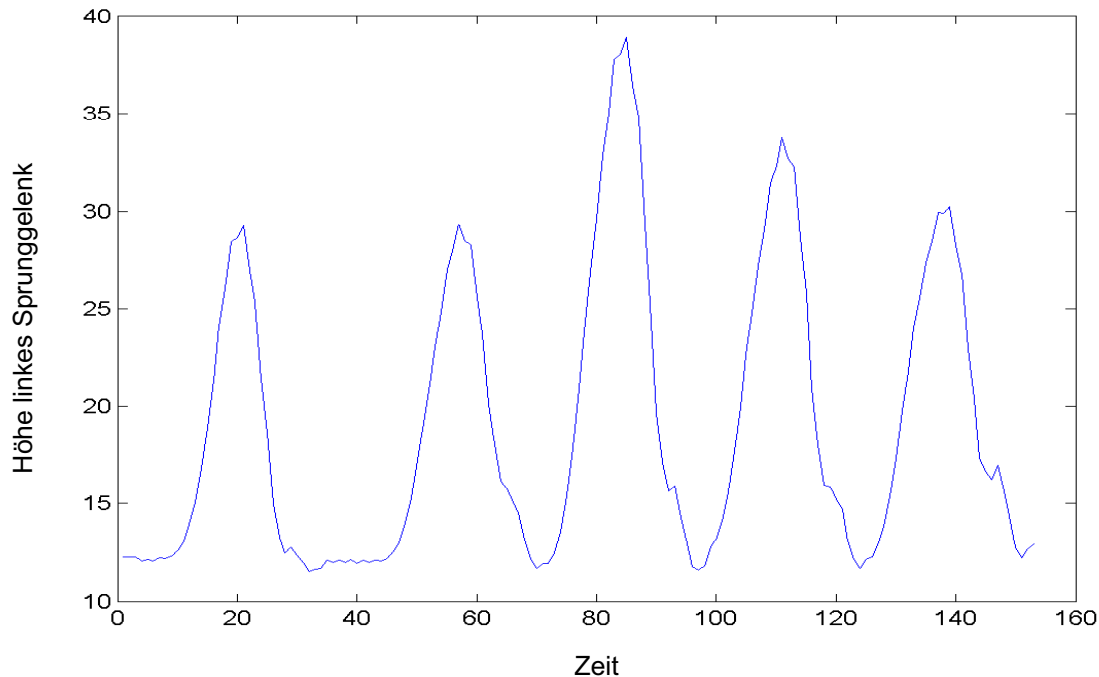


Abb. 2: Höhe des linken Sprunggelenks bei einer Gehbewegung.

Betrachtet man das einfache sinusoidale Modell $f(t) := \sin(2\pi 200 + p(t))$, bei dem das Ausgangssignal f aus einer einzelnen Sinusschwingung besteht, so könnte man beispielsweise den Phasenparameter p durch die Höhe eines der menschlichen Sprunggelenke bei einer Gehbewegung ersetzen, vgl. Abb. 2.

Ereignisgesteuerte Sonification

Im Falle der samplebasierten Klangerzeugung besteht ein kanonischer Ansatz zur ereignisgesteuerten Sonification darin, einen bestimmten Audioklang genau dann abzuspielen, wenn ein bestimmtes Ereignis im Datenstrom auftritt. Bestimmte Eigenschaften des Ereignisses können weitere Sonificationsparameter, wie die Dauer des wiederzugebenden Klanges oder dessen Lautstärke, steuern.

Abb. 3 zeigt die aus einer ereignisgesteuerten Sonification der Gehbewegung aus Abb. 2 resultierende akustische Wellenform. Hierbei sind einzelne Ereignisse durch die lokalen Maxima der Sprunggelenkhöhe gegeben. Bei jedem der fünf Ereignisse wird ein Instrumententon mit Lautstärke proportional zur Sprunggelenkhöhe wiedergegeben.

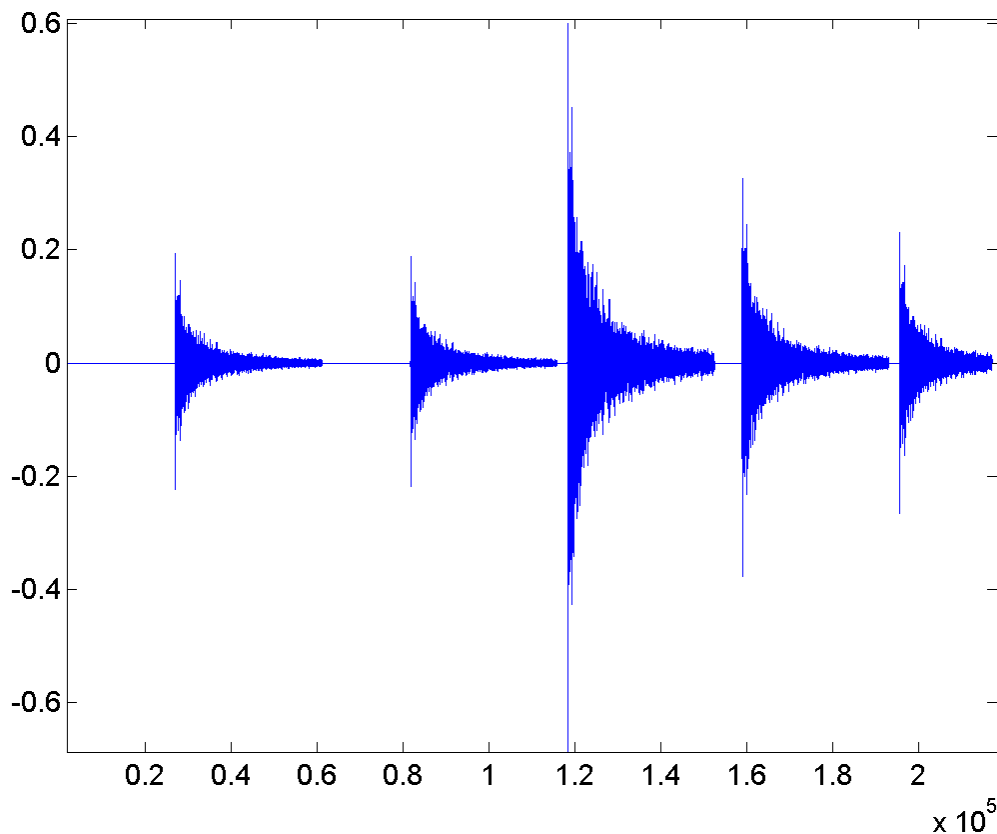


Abb. 3: Ereignisgesteuerte Sonifikation der Sprunggelenkshöhe aus Abb. 2.

Softwaresysteme

In diesem Abschnitt werden einige - im Sonifikationskontext direkt oder indirekt relevante - Softwaresysteme vorgestellt, die in der Arbeitsgruppe Multimedia-Signalverarbeitung am Institut für Informatik III der Universität Bonn in den letzten Jahren entstanden sind.

In Kooperation mit dem Institut für Sportwissenschaft und Sport entstand in den Jahren 1998/1999 ein generisches System zur akustischen Bewegungstransformation. Die Klangerzeugung erfolgt in diesem System durch Modulation einzelner, durch einen MIDI-Synthesizer erzeugter Töne. Die Eingangsdaten sind 2D- oder 3D-Bewegungsdaten, die nicht in Echtzeit anfallen. Das System wurde in einer Folgearbeit (Becker, 1999) so erweitert, dass in Echtzeit mittels einer Kraftmessplatte erfasste Daten sonifiziert werden können.

Die automatische Extraktion wesentlicher Ereignisse, wie etwa bestimmter Bewegungsabläufe oder Körperhaltungen, aus den Eingangsdaten ist eine in der Praxis häufig anzutreffende Aufgabenstellung. Dies bedeutet, dass der Vorverarbeitungsschritt der Sonifikation im Echtzeitszenario eine bestimmte Art des *Monitorings* ist. Beim Monitoring geht es allgemein darum, zeitliche Datenströme zu überwachen und signifikante Ereignisse hieraus abzuleiten. Generische Monitoringsysteme, wie das in einem weiteren Projekt (Wild, 2005) entwickelte GenMAD-System, erlauben

hier eine beliebige Verschaltung entsprechender Module zur Signalvorverarbeitung und Ereignisextraktion.

Schließlich sei noch der Bezug zu der in gewisser Weise komplementären Aufgabe der *Musiksynchronisation* hergestellt, bei der es darum geht, ein akustisches Signal und einen symbolischen Datenstrom (etwa eine CD-Aufnahme eines Musikstücks und die dazu gehörige Partitur) zeitlich zu synchronisieren. Synchronisation bezeichnet hier die Aufgabe, die Noten der Partitur denjenigen Zeitpositionen der CD-Aufnahme zuzuordnen, an denen sie tatsächlich gespielt werden. Ein entsprechendes System zur simultanen akustischen und visuellen Wiedergabe solch einer Synchronisation ist der SyncPlayer (Kurth et al., 2005). Im Gegensatz zur Synchronisation, wo das akustische Signal in Form der CD-Aufnahme gegeben ist und eine Beziehung zu den symbolischen Daten erst noch hergestellt werden muss, sind im Falle der Sonification die nicht-akustischen Daten (in Form der Eingangsdaten) gegeben und die akustischen Daten müssen erzeugt werden.

Fazit

Die seitens der Audiosignalverarbeitung verfügbaren Methoden der Klangerzeugung bieten sowohl für die ereignisgesteuerte als auch die modulationsbasierte Sonification ein breites Spektrum an Möglichkeiten zur Erzeugung akustischer Signale aus gegebenen Messdaten. Die qualitativ hochwertigsten Resultate lassen sich aktuell mit Methoden der samplebasierten Synthese und des Physical Modelling erzielen, wobei beim Physical Modelling dem enormen Spektrum an Möglichkeiten der Klangerzeugung ein ebenfalls benötigter hohe Interaktionsbedarf seitens Domänenexperten gegenübersteht. Letztlich stellt sich die Sonificationsaufgabe in der hier vorgestellten Form so allgemein dar, dass der bei der Sonification zu erzielende Klangcharakter und dessen perzeptuelle Qualität erst durch menschliche Interaktion festgelegt werden muss. Die dabei erzielbare hohe musikalische Qualität wurde unlängst in einem Vortrag von T. Dolby (Dolby, 2005) demonstriert.

Literatur

- Smith, J. O. III (2004). Physical Audio Signal Processing - Digital Waveguide Modelling of Musical Instruments and Audio Effects, www-Version: <http://ccrma.stanford.edu/~jos/pasp/>
- Becker, A. (1999). Echtzeitverarbeitung dynamischer Bewegungsdaten mit Anwendungen in der Sonification. Diplomarbeit, Institut für Informatik V der Universität Bonn.
- Wild, H. (2005). Ein generisches Monitoringsystem für akustische Datenströme, Diplomarbeit. Institut für Informatik V der Universität Bonn.
- Kurth, F., Müller, M., Damm, D., Fremerey, C., Ribbrock, A., & Clausen, M. (2005). SyncPlayer - An Advanced System for Multimodal Music Access. 6th International Conference on Music Information Retrieval. London, UK.
- Dolby, T. (2005). Sonifications (Invited Talk). 6th International Conference on Music Information Retrieval. London, UK.

ALFRED O. EFFENBERG, URSULA FEHSE & HEINZ MECHLING

Zur Verarbeitung audiovisueller Bewegungsinformation im Sport: eine fMRT-Studie

Einführung

Die Wahrnehmung, die Beobachtung und der Nachvollzug von sportlichen Bewegungen sind Schlüsselfunktionen beim motorischen Lernen. Bislang dominieren dabei in der Sportpraxis visuelle und verbale Informationen. Mit der Bewegungs-Sonification lässt die Bewegungswahrnehmung durch eine zusätzliche Bewegungsakustik unterstützen bzw. präzisieren. Zeitlich synchron zu der Videodarstellung der Bewegung werden bestimmte Bewegungsparameter akustisch dargestellt. Tatsächlich belegen verhaltenspsychologische Untersuchungen die Wirksamkeit bewegungsakustischer Zusatzinformationen bei der Wahrnehmung, Beurteilung und dem Nachvollzug großmotorischer Sportbewegungen. Diese Indizien stehen im Einklang mit Untersuchungsergebnissen zu visuellen, auditiven und audiovisuellen Effekten bei der Wahrnehmung biologischer Bewegungen und der Sprachwahrnehmung. Während aber für diese beiden Forschungsbereiche begleitend neurophysiologische Daten zu den funktionalen Zusammenhängen vorliegen, ist das für die zusätzliche bewegungsakustische Information auf Basis der Sonification bisher nicht der Fall. Erkenntnisse zu den neurophysiologischen Wirkungszusammenhängen könnten aber nicht nur bei der Erklärung der beobachteten Effekte hilfreich sein. Sie könnten zudem bei der Bestimmung zukünftiger Anwendungsbereiche und bei der Weiterentwicklung der bewegungsakustischen Transformationsregeln maßgebliche Hinweise liefern. Ebenso gewannen Einträge der Sonifications-Forschung in die Motorik- und Lerntheorie eine neue Qualität, wenn die bisherige kognitionswissenschaftliche Orientierung durch eine neurokognitive Ausrichtung ergänzt würde. Vor diesem Hintergrund ist eine konsistente Ausrichtung der Methode für die neurophysiologische Untersuchung an die Methodik der bereits realisierten Verhaltensstudien unbedingt erforderlich. Nur dann lassen sich verhaltenspsychologische und die neurophysiologische Daten direkt aufeinander beziehen. Im folgenden wird dazu die Entwicklung einer fMRT-Untersuchungsmethode vorgestellt, die unmittelbar auf der zuvor verwendeten verhaltenspsychologischen Methode aufbaut.

Verhaltensbezogene und neurophysiologische Indizien

Empirische Untersuchungen haben ergeben, dass die Beurteilungs- und Nachvollzugsleistung von visuell präsentierten Hock-Streck-Sprüngen durch eine zusätzliche konvergente akustische Darstellung der vertikalen Komponente der Bodenreaktionskraft verbessert werden kann (Effenberg, 2004; 2005). Offenbar unterstützt die Bewegungs-Sonification in Kombination mit der Bewegungsvisualisierung die Präzision der Wahrnehmungsleistung. Doch wodurch kommt eine solche Steigerung der Wahrnehmungsleistung – die auch motorisch wirksam wird – zustande? Zur Beantwortung dieser Frage ist zu klären, welche Gehirnareale bei der

audiovisuellen Wahrnehmung sportlicher Bewegungen angesprochen werden und wie die audiovisuellen Bewegungsinformationen verarbeitet werden. Hinweise darauf liefern Studien aus benachbarten Forschungsbereichen, wie z. B. aus der Sprachwahrnehmung und der Wahrnehmung biologischer Bewegungen.

Auf den ersten Blick scheint die Sprachwahrnehmung ein auditives Phänomen zu sein. Sumbly und Pollack zeigten jedoch, dass die zusätzliche visuelle Wahrnehmung der entsprechenden Mundbewegungen die Verständlichkeit der Sprachsequenzen verbessert (Sumbly & Pollack, 1954). Mit der Frage, wo der Unterschied in der Gehirnaktivität zwischen unimodaler und audiovisueller Sprachwahrnehmung liegt, haben sich Calvert, Campbell und Brammer (2000) beschäftigt. Via fMRT (funktionelle Magnetresonanztomographie) suchten sie Gehirnareale, die auditive und visuelle Sprachinformationen integrieren. Ein multisensorisches Areal wurde von den Autoren in Anlehnung an die 'single-neuron'-Forschung dahingehend charakterisiert, dass es sowohl bei der rein auditiven, als auch bei der rein visuellen Sprachwahrnehmung aktiviert. Ebenso sollte es bei audiovisuell konvergenten Sprachinformationen eine Aktivität aufweisen, die die Summe der unimodalen Aktivitäten übersteigt. Diese supraadditive Aktivitätscharakteristik ist ein 'hartes' Kriterium für die Lokalisation multisensorischer Integrationsinstanzen im ZNS. Stimmen die auditiv und visuell wahrgenommenen Sprachinformationen offensichtlich nicht überein, sollte die Aktivität hingegen unter der Summe der Einzelaktivitäten liegen. Das Areal sollte also die auditiven und visuellen Informationen nicht nur einfach parallel verarbeiten, sondern 'erkennen', ob sie zueinander passen und folglich als ein Ereignis angesehen werden können. Alle drei Kriterien erfüllte lediglich ein Bereich im oberen Teil des linken Schläfenlappens, der ventraler Bereich im posterioren Teil des superioren Temporal sulcus (STSp) in der linken Gehirnhälfte.

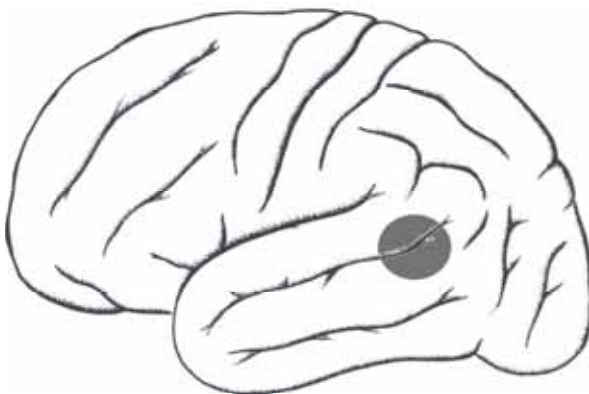


Abb. 1: Der posteriore superiore Temporal-sulcus (STSp).

Die bimodale Aktivierung auf konvergente audiovisuelle Sprache war 30-80% größer als die Summe der auditiven und visuellen Aktivierung. Im Gegensatz dazu reduzierten divergente audiovisuelle Inputs die Aktivität in diesem Areal auf weniger als 50% der summierten unimodalen Reaktionen. Diese Aktivierungscharakteristik ähnelt der der multisensorischen Integrationsneuronen im Kortex von nichtmenschlichen Säugetieren. Der linke STSp integriert offenbar auditive und visuelle Sprachinformationen (Calvert et al., 2000).

Doch lassen sich solche Erkenntnisse zur Sprachwahrnehmung auf andere Bereiche menschlicher Bewegung übertragen? Dazu haben unter anderem Grossmann und Blake eine Studie durchgeführt. Sie wollten herausfinden, ob es ein Gehirnareal gibt, das auf die Verarbeitung menschlicher großmotorischer Bewegungen spezialisiert ist. Um sicherzugehen, dass die kortikale Aktivierung tatsächlich auf die Bewegung zurückzuführen ist und nicht auf die Tatsache, dass ein Mensch – also

lediglich das dargestellte Objekt – wahrgenommen wurde, mussten sie in besonderer Weise reduzierte Stimuli verwenden: Dazu bot sich die von Johansson (1973) entwickelte point-light Technik an. Bei dieser werden menschliche Bewegungen ausschließlich über die Bewegung der Hauptgelenkpunkte dargestellt. Das Bewegungsmuster wird nur durch die relative kinematische Struktur dieser Punkte aktiviert¹. Sobald die Bewegung beginnt, kann sie in der Regel sehr schnell erkannt werden.

Grossman und Blake verwendeten biologische point-light Bewegungssequenzen und reale Bewegungssequenzen als Experimentalstimuli. Unter dem Begriff der 'biologischen Bewegung' werden neben Lokomotionsbewegungen auch eine Vielzahl von großmotorischen Sport- und Alltagsbewegungen wie z. B. Hüpfen, Werfen und Treten subsummiert. Um diese gegen nicht-biologische Bewegungen abzugrenzen, wurden in einer Kontrollbedingung Objektbewegungen und bedeutungslose Punktbewegungen dargestellt. Letztere bestanden aus den gleichen Einzelpunktbewegungen, wie die biologischen point-light Stimuli, die aber anders angeordnet und folglich nicht mehr als menschliche Bewegung zu erkennen waren.

Der posteriore superiore Temporalsulcus (STSp) wies bei den biologischen point-light Bewegungen eine signifikant stärkere Aktivierung auf als bei den Kontrollbewegungen. Durch zusätzliche Kontrollbedingungen konnte ausgeschlossen werden, dass lediglich die Identifikation des Objektes – unabhängig von der Bewegung – zu der Aktivierung geführt hat. Der STSp kann offensichtlich die Bewegung von 12 Punkten in einer spezifischen Anordnung als biologische menschliche Bewegung erkennen, er scheint auf die Wahrnehmung biologischer Bewegungen spezialisiert zu sein (Grossman & Blake, 2002). Zudem weisen aktuelle Studien darauf hin, dass der STSp auch bei der auditiven Wahrnehmung einer menschlichen großmotorischen Bewegung, wie dem Gehen, anspricht (Bidet-Calet et al., 2005). Untersuchungen an Affen haben zudem ergeben, dass viele STSp-Neurone sowohl auditiven als auch visuellen Input erhalten (Bruce et al., 1981) und dass sie sensitiv für die Wahrnehmung biologischer Bewegungen sind (Oram & Perrett, 1994). Die vorgestellten fMRT-Studien belegen, dass der STSp audiovisuelle Sprachinformationen integriert und dass er auf die visuelle und auditive Wahrnehmung biologischer Bewegungen anspricht. Wenn der audiovisuelle Integrationsmechanismus im STSp nicht nur auf die Sprachwahrnehmung begrenzt funktioniert, sollten auch audiovisuelle Bewegungsinformationen ähnliche Integrationseffekte im STSp anregen. Eine solche zusätzliche Adressierung audiovisueller Integrationsinstanzen könnte eine Erklärung für die beobachteten verhaltensseitigen Effekte sein. Mit der fMRT-Untersuchung sollte dies überprüft werden.

Wie funktioniert fMRT?

Mithilfe der fMRT lassen sich Rückschlüsse auf die Gehirnaktivität ziehen. Genauer gesagt können durch sie Unterschiede in der lokalen Gehirnaktivität unter verschiedenen Bedingungen erfasst werden. Die Methode gewährleistet eine

¹ Ein Überblick zur point-light Technik ist im Beitrag von Westhoff und Troje in diesem Band zu finden.

bessere räumliche Auflösung als die meisten anderen Verfahren, die Daten über Gehirnaktivität erzeugen (EEG, MEG, PET und SPECT²) und zumindest eine bessere zeitliche Auflösung als PET und SPECT. Es ist im Gegensatz zu einigen der anderen Methoden weder die Applikation von Kontrastmitteln erforderlich, noch werden Röntgenstrahlen verwendet. Der Magnetresonanztomograph besteht aus einem starken Magneten, dessen Feldstärke der 50.000-fachen Stärke des Erdmagnetfeldes entspricht. Kurze Radiofrequenzimpulse regen bestimmte Atome zum Schwingen (Resonanz) an. Dabei wird Energie freigesetzt und es entsteht ein Magnetfeld, das kurze Zeit später wieder zerfällt. Eine Spule erfasst diesen Zerfall (Magnetresonanz-Signal). So können zum einen verschiedene Gewebearten anhand der unterschiedlichen Wasserstoffkonzentration voneinander unterschieden werden als auch unterschiedliche Funktionszustände identischer Gewebe anhand unterschiedlicher Sauerstoffsättigungsgrade des Hämoglobins.

Eine Erhöhung der Gehirnaktivität erfordert nämlich einen erhöhten Energiebedarf, was wiederum mit der Notwendigkeit einer erhöhten Sauerstoffbereitstellung einher geht. Der Sauerstoff im Blut ist an Hämoglobin gebunden. Im Blut sind zwei Arten von Sauerstoff vorhanden: Oxyhämoglobin (sauerstoffangereichertes Hämoglobin) und Desoxyhämoglobin (Hämoglobin ohne Sauerstoffmolekül). Oxyhämoglobin und Desoxyhämoglobin zeigen unterschiedliches Verhalten in einem Magnetfeld, was zu einem Unterschied in der 'hämodynamischen Antwort' führt. Eine Erhöhung der Gehirnaktivität in einer bestimmten Region führt also letztendlich zu einer Änderung des Magnetresonanzsignals, das aus dieser Region stammt. Es können Aktivierungsbilder erzeugt werden, in denen Regionen, die treatmentbedingt unterschiedlich aktivieren, farblich markiert sind. Die Messungen lassen jedoch keine Rückschlüsse auf die Qualität der Gehirnaktivität bzw. der damit verbundenen Kognitionen zu.

Methodenentwicklung

In der hier vorgestellten Studie sollen Unterschiede in der hämodynamischen Antwort bei unimodalen und audiovisuellen Stimuli erfasst werden. Den Probanden wurden, wie auch in der verhaltenspsychologischen Untersuchung (Effenberg 2004; 2005), Video- und/ oder Tonaufnahmen von Hock-Strecksprüngen unterschiedlicher Höhen präsentiert. Zur Aufmerksamkeitslenkung bekamen sie hier jedoch die Aufgabe, die jeweils dargebotene Sprunghöhe zu beurteilen, da ein Vergleich zweier Sprunghöhen wie bei den verhaltenspsychologischen Untersuchung den Aufenthalt der Probanden im Tomographen zu sehr in die Länge gezogen hätte. Es gab fünf verschiedene Bedingungen: zwei unimodale (rein auditiv bzw. visuell), zwei bimodale und eine Baseline-Bedingung. In der rein auditiven Bedingung wurde die bewegungsdefinierte Klangsequenz dargeboten, welche die vertikale Komponente der Bodenreaktionskraft darstellt. Bei der audiovisuell konvergenten Bedingung wurden Bild und Ton synchron dargeboten. Zusätzlich zu

² EEG: Elektroenzephalographie, MEG: Magnetenzephalographie, PET: Positronenemissionstomographie, SPECT: Single-Photon-Emissions-Computertomographie.

den drei Bedingungen aus den verhaltenspsychologischen Untersuchungen musste eine audiovisuell divergente Bedingung hinzugefügt werden, in der Bild und Ton zeitlich stark versetzt waren (Fehse 2005). In allen vier Treatments waren die Stimuli so gestaltet, dass die Sprunghöhe nur über die relative Kinematik und die Dynamik zu bewerten war, was die Stimuli ökologisch valide machte. Die Kontrollbedingung bestand aus einem Schwarzbild ohne Ton. Die Erhöhung der Versuchsbedingungen von drei auf fünf war nötig um zu gewährleisten, dass Änderungen in der Gehirnaktivität auch tatsächlich auf die spezifischen Unterschiede zwischen den einzelnen Treatments zurückzuführen waren.

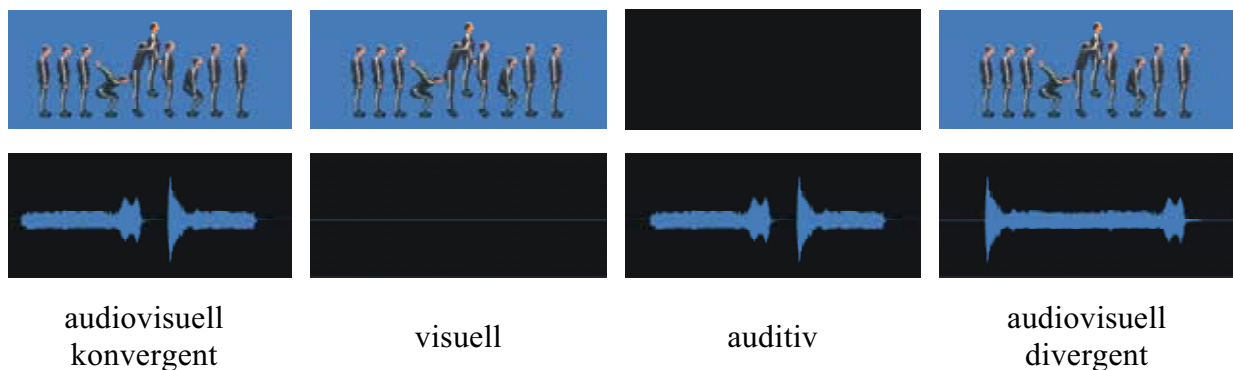


Abb. 2: Die Stimuli aus den vier Treatments. Obere Bildreihe: Einzelbildteile aus dem Video. Untere Bildreihe: Amplitudenpegeldiagramm der Tonspur des Videos.

Bei der Entwicklung der Methode gab es einige Schwierigkeiten zu überwinden. Zum einen musste berücksichtigt werden, dass im Untersuchungsraum elektronische Geräte nicht funktionieren und magnetische Metalle in den Geräten nicht enthalten sein dürfen. Daher wurden pneumatische Kopfhörer für die Präsentation der auditiven Stimuli als auch einen lichtleitergesteuerten Taster als Eingabegerät für die Sprunghöhenbeurteilung gewählt. Ein weiteres Problem war, dass der Tomograph während der Aufnahmen laute Geräusche erzeugt, welche über Maskierungseffekte die Wahrnehmung der auditiven Stimuli beeinträchtigt hätten. Dieses Problem ließ sich durch den Einsatz eines sogenannten SPARSE-Designs lösen, ein Messdesign, bei dem die Messung der hämodynamischen Antwort erst einige Sekunden nach der Stimuluspräsentation erfolgt. Bei diesem Messdesign macht man sich das Phänomen zunutze, dass das Maximum der hämodynamischen Antwort erst einige Sekunden nach der Erhöhung der lokalen Gehirnaktivität erreicht wird.

Während der Untersuchung lagen die Probanden im Tomographen, bekamen über ein Spiegelsystem die auf eine Leinwand projizierten Videosequenzen präsentiert, hörten über den Kopfhörer die Klangsequenzen und bedienten mit der dominanten Hand über zwei Tasten des Tasters eine ebenfalls auf der Leinwand dargestellte Beurteilungsskala.

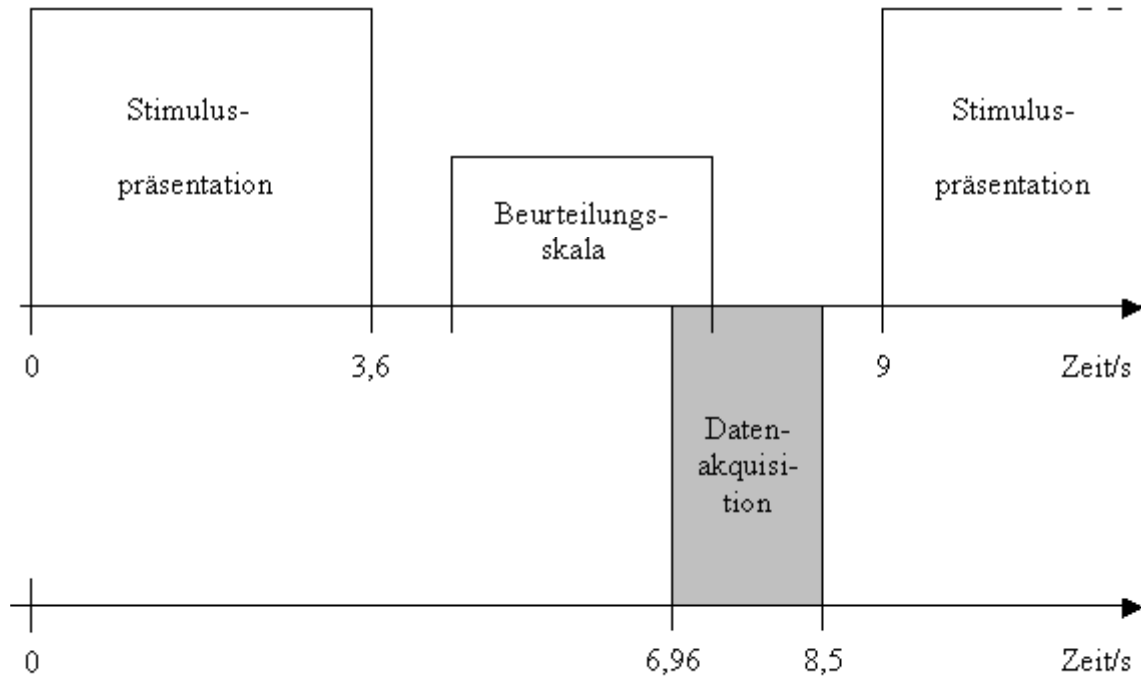


Abb. 3: Das Untersuchungsdesign.

Stand der Auswertung

Nach der Untersuchung mussten sowohl die Verhaltensdaten (Höhenbeurteilung) als auch die fMRT-Daten ausgewertet werden. Die Probanden waren überzufällig gut in der Lage, die Sprunghöhen zu beurteilen. Damit ist sichergestellt, dass die Probanden ihre Aufmerksamkeit während der gesamten Untersuchung auf die Stimuli fokussiert haben (Fehse 2005).

Die fMRT-Messungen lieferten zwei Haupterkennnisse. Erstens, dass die auditiven Stimuli den STSp – insbesondere rechtshemisphärisch – aktivieren. Insofern wurde die gleiche Region, die auch bei der Wahrnehmung von realen bewegungsbegleitenden Geräuschen anspricht, durch die Bewegungs-Sonification aktiviert. Das ist ein erstes Indiz dafür, dass die bewegungsdefinierte Klangsequenz entsprechend den ökologischen bewegungsbegleitenden Geräuschen verarbeitet wurde. Zweitens wurde der STSp in der audiovisuell konvergenten Bedingung stärker aktiviert als in der audiovisuell divergenten Bedingung. Das lässt darauf schließen, dass der STSp die konvergenten und divergenten Stimuluskombinationen unterschiedlich verarbeitet. Diese beiden Ergebnisse können als erste Hinweise darauf interpretiert werden, dass zusätzlich zu den visuellen Bewegungsinformationen eingesetzte akustische Informationen, die mit der Bewegungs-Sonification hergestellt werden, tatsächlich im STSp verarbeitet werden.

Literatur

- Bidet-Caulet, A., Voisin, J., Bertrand, O., & Fonlupt, P. (2005). Listening to a walking human activates the temporal biological motion area. *NeuroImage*, 28(1), 132-139.
- Bruce, C., Desimone, R., & Gross, C.G. (1981). Visual Properties of Neurons in a Polysensory Area in Superior Temporal Sulcus of the Macaque. *Journal of Neurophysiology*, 46(2), 369-384.
- Calvert, G. A., Campbell, R., & Brammer, M. J. (2000). Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Current Biology*, 10(11), 649-657.
- Effenberg, A. O. (2004). Synergien der Sinne für die Bewegungsregulation. Frankfurt a. M.: Peter Lang.
- Effenberg, A. O. (2005). Movement Sonification: Effects on Perception and Action. *IEEE Multimedia, Special Issue on Interactive Sonification*, 12(2), 53-59.
- Fehse, U. (2005). Zur Funktion auditiver, visueller und audiovisueller Mechanismen bei der Wahrnehmung sportlicher Bewegungen. Unveröffentlichte Magisterarbeit, Universität Bonn.
- Grossmann, E. D., & Blake, R. (2002). Brain Areas Active during Visual Perception of Biological Motion. *Neuron*, 35(6), 1167-1175.
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, 14(2), 201-211.
- Oram, M. W., & Perrett, D. I. (1994). Response of Anterior Superior Temporal Polysensory (STPa) Neurons to "Biological motion" Stimuli. *Journal of Cognitive Neurosciences*, 6(2), 99-116.
- Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *Journal of the Acoustical Society of America*, 26(2), 212-215.

Autorenverzeichnis



Clausen, Michael, Prof. Dr., Institut für Informatik III, Universität Bonn,
email: clausen@informatik.uni-bonn.de

Dr. Michael Clausen studierte Mathematik, Physik und Informatik. Promotion in Mathematik, Habilitation in Informatik. Nach Lehrtätigkeiten an den Universitäten Aachen, Bayreuth, Zürich und Karlsruhe seit 1989 Professor für Informatik an der Universität Bonn. Aktuelle Forschungsgebiete: Multimedia Information Retrieval, digitale Musikbibliotheken, wissensbasiertes Monitoring.



Effenberg, Alfred Oliver, PD Dr., Deutsche Sporthochschule Köln / Institut für Sportwissenschaft und Sport, Universität Bonn, email: effenberg@uni-bonn.de

Dr. Alfred O. Effenberg studierte Sportwissenschaft und Journalistik an der Uni Hamburg. 1993 Forschungsaufenthalt an der UCLA, 1995 Promotion an der Universität Hamburg, 2002 Habilitation an der Universität Bonn. Arbeitsschwerpunkte: Motorik- und Bewegungsforschung, Wahrnehmungs- und Handlungspsychologie, Multisensorische Integration, Bewegungs-Sonification. Gegenwärtig Wissenschaftlicher Mitarbeiter an der Deutschen Sporthochschule Köln



Fehse, Ursula, Institut für Sportwissenschaft und Sport, Universität Bonn,
email: UrsulaFehse@buschdorf.de

Ursula Fehse M. A. studierte Sportwissenschaften an der Universität Bonn, wo sie als Wissenschaftliche Hilfskraft beschäftigt ist. Arbeitsschwerpunkte: Wahrnehmungs- und Handlungspsychologie, Multisensorische Integration.



Kurth, Frank, PD Dr., Institut für Informatik III, Universität Bonn,
email: frank@iai.uni-bonn.de

Dr. Frank Kurth studierte Informatik und Mathematik an der Universität Bonn, wo er im Fach Informatik 1999 promovierte (summa cum laude) und sich 2005 habilitierte. Seit 1997 arbeitet er in der Arbeitsgruppe Multimedia-Signalverarbeitung am Institut für Informatik der Universität Bonn mit Forschungsschwerpunkten Multimedia Information Retrieval, Multimedia-Signalverarbeitung und Digitale Bibliotheken. Dr. Kurth ist Mitglied von IEEE und AES.



Mechling, Heinz, Prof. Dr., Deutsche Sporthochschule Köln / Institut für Sportwissenschaft und Sport, Universität Bonn,
email: mechling@uni-bonn.de

Dr. phil. Heinz Mechling ist Universitätsprofessor an der Universität Bonn und Direktor des Instituts für Sportwissenschaft und Sport. Seine Arbeitsschwerpunkte liegen in der Motorik- und Bewegungsforschung im Motorischen Lernen und Techniktraining, der Bewegungsdiagnostik und im Alterssport.



Memmert, Daniel, Dr., Institut für Sport und Sportwissenschaft, Universität Heidelberg, email: daniel.memmert@urz.uni-heidelberg.de

Dr. Daniel Memmert ist derzeit wissenschaftlicher Mitarbeiter am Institut für Sport und Sportwissenschaft der Ruprecht-Karls-Universität Heidelberg. Promoviert hat er über Kognitionen und Taktiken im Sportspiel (2003). Seine aktuellen Forschungsschwerpunkte liegen in der Bewegungs- und Trainingswissenschaft, im Bereich der Sportspiel- und Evaluationsforschung sowie in den Forschungsmethoden.



Müller, Meinard, Dr., Institut für Informatik III; Universität Bonn,
email: meinard@iai.uni-bonn.de

Dr. Meinard Müller studierte Mathematik (Diplom, 1997) und Informatik (Promotion, 2001) an der Universität Bonn. Nach einem Forschungsaufenthalt 2002/2003 am mathematischen Institut der Keio Universität, Japan, forscht er seit 2003 an der Universität Bonn im Bereich der multimedialen Informationsverarbeitung.



Perl, Jürgen, Prof. Dr., Institut für Informatik, Universität Mainz,
email: perl@informatik.uni-mainz.de

Dr. Jürgen Perl, 1971: Promotion in Mathematik, Freie Universität Berlin. 1975: Informatik-Professur an der Universität Osnabrück. 1984: Informatik-Professur an der Universität Mainz. Arbeitsschwerpunkte: Informatik in Sport und Sportwissenschaft und in der Medizin. Präsident der 'Association of Computer Science in Sport': 2003-2007.



Röder, Tido, Dipl.-Inform., Institut für Informatik III, Universität Bonn,
email: roedert@cs.uni-bonn.de

Tido Röder studierte Informatik an der Universität Bonn und an der Washington State University, USA. Im Jahr 2002 schloss er das Diplom in der Arbeitsgruppe von Prof. Michael Clausen mit einer Arbeit über inhaltsbasierte Bildsuche ab. Seit 2003 promoviert er bei Prof. Clausen über Retrieval und Klassifikation von menschlichen Bewegungsdaten.



Troje, Nikolaus Friedrich, Prof. Dr., Department of Psychology and School of Computing, Queen's University, Kingston, email: troje@post.queensu.ca

Dr. Nikolaus F. Troje, Promotion 1994 in Tierphysiologie, Universität Freiburg. Wissenschaftlicher Assistent am Max Planck Institut für Biologische Kybernetik, Tübingen. 1997-1999 Visiting Professor an der Queen's University, Kingston, Ontario. 1999-2005 Universität Bochum, Leitung des BioMotionLab im Department für Psychologie. Seit 2003 ist er 'Associate Professor' am Department für Psychologie und an der 'School of Computing', Queen's University.



Weber, Andreas, Prof. Dr., Institut für Informatik II, Universität Bonn,
email: weber@cs.uni-bonn.de

Dr. Andreas Weber studierte Mathematik an der Universität Tübingen und der University of Colorado, Boulder, USA. Promotion in Informatik an der Universität Tübingen 1993. Postdoktorand an der Fakultät für Informatik der Universität Tübingen, dem Department of Computer Science der Cornell University, Ithaca, NY, USA und am Fraunhofer-Institut für Graphische Datenverarbeitung, Darmstadt. Ab 2001 Professor für Praktische Informatik, Multimedia an der Universität Bonn.



Westhoff, Cord, Dr., Institut für Kognitive Neurowissenschaften, Universität Bochum, email: cord.westhoff@rub.de

Dr. Cord Westhoff arbeitet als wissenschaftlicher Mitarbeiter am Institut für Arbeitsphysiologie an der Universität Dortmund. Im Juli 2005 promovierte er an der Ruhr-Universität Bochum über das Thema Personenidentifikation anhand von biologischer Bewegung. Die aktuellen Forschungsschwerpunkte liegen im Bereich des Bewegungslernens und der Bewegungsausführung.

