in: G. Sommer (eds): Mustererkennung 2000, Proceedings 22. DAGM-Symposium, Springer, 2000, erscheint

Diskrete Hidden Markov Modelle zur Analyse von Meßkurven amperometrischer Biosensoren*

Jörg Weitzenberg[1], Stefan Posch[1], Manfred Rost[2]

[1] Martin-Luther-Universität Halle-Wittenberg
Fachbereich Mathematik und Informatik, Institut für Informatik
Kurt-Mothes-Str. 1, D - 06099 Halle/Saale, Germany
{weitzenb,posch}@informatik.uni-halle.de
[2] Martin-Luther-Universität Halle-Wittenberg
Fachbereich Physik
Friedemann-Bach-Platz 6, D - 06099 Halle/Saale, Germany
rost@physik.uni-halle.de

Zusammenfassung In diesem Beitrag wird ein auf diskreten HMM basierendes Verfahren zur Analyse von Biosensor-Meßkurven vorgestellt. Für die Generierung der HMM werden sowohl statistische Trainingsdaten als auch spezielles Expertenwissen genutzt. Das System nutzt die HMM zur Unterscheidung verschiedener Kurvenformen, zur Detektion von für die Messung relevanten Abschnitten innerhalb der Meßkurven sowie zur Erkennung des Endes einer Messung. Dazu wurden die Meßkurven in einzelne Abschnitte unterteilt, denen spezielle Zustände im HMM zugeordnet wurden. Bei ersten Experimenten wurden bereits gute Ergebnisse hinsichtlich Genauigkeit und Erkennungsrate erzielt. Es zeigte sich, daß diskrete HMM für diese Problemstellung gut geeignet sind.

Schlüsselwörter: Hidden Markov Modelle, Signalanalyse, Biosensor

1 Einleitung

Hidden Markov Modelle (HMM) haben sich in der letzten Dekade zur Standardmethode bei der automatischen Spracherkennung entwickelt. Sie kommen in den letzten Jahren aber auch für andere Problemstellungen zunehmend zum Einsatz [1,2].
Das hier behandelte Problem der Analyse und Interpretation von Meßkurven tritt in der Biosensorik auf, ist aber auch in anderen Bereichen der Sensorik anzutreffen. Das Ziel dieser Arbeit ist die Enwicklung eines (automatischen) Meßsystems mit amperometrischen Biosensoren, welches sowohl statistische Trainingsdaten als auch spezielles Expertenwissen in die Meßwertgewinnung einbezieht. Amperometrische Biosensoren werden in den Bereichen Medizin-, Lebensmittel- und Umweltanalytik zur Messung der Konzentration C von Analyten wie Glucose, Saccharose und Phenol benötigt. Ein Biosensor ist gekennzeichnet durch

* Gefördert mit Forschungsmitteln des Kultusministeriums des Landes Sachsen-Anhalt.

die Integration einer biologisch sensitiven Komponente (z.B. ein Enzym) mit einem Signalwandler (Transducer) [5]. Der zu bestimmende Analyt reagiert mit der Biokomponente, was zu einem vom Transducer abgebildeten elektrischen Signal führt. Amperometrische Biosensoren liefern als Meßsignal einen zeitlich veränderlichen Strom $I(t, C)$, der sich aus einem Grundstromanteil I_0 und dem konzentrationsabhängigen Signalstromanteil I_D zusammensetzt. Daher müssen bei jeder Messung I_0 und I_D getrennt ermittelt werden. Hierzu wird der Sensor zuerst in eine Pufferlösung (I_0) und dann in die eigentliche Meßlösung (I_D) getaucht. Dabei zeigt sich in der Pufferlösung eine typische Einlaufkurve, an deren flachster Stelle I_0 ermittelt wird. In der Meßlösung I_D zeigen sich charakteristische Signalkurven, deren Form abhängig von Sensortyp, Meßverfahren und der zu messenden Konzentration des Analyten ist. Die Dauer einer Kurve liegt im Bereich von ca. 3-10 min. Anhand des Kurvenverlaufes wird entschieden, wann die Messung beendet wird. Der Wert für I_D wird dann an einer vom jeweiligen Kurventyp abhängenden Position innerhalb der gemessenen Kurve bestimmt. Die Zeitpunkte zum Beenden der Messung sowie zur eigentlichen Meßwertgewinnung innerhalb der Kurve können sich von Kurventyp zu Kurventyp erheblich unterscheiden (siehe Abbildung 1). Aus der Menge der unterschiedlichen Kurvenformen kann erst während des Meßvorganges der für die Meßwertbildung und Endeerkennung zugrundezulegende Kurventyp detektiert werden. Die Übertragungskennlinie von Biosensoren unterliegt gewissen Parameterschwankungen, die einerseits herstellungsbedingt sind (preisgünstige Massenware) oder andererseits durch Alterung bzw. Vergiftung des Sensors verursacht werden. Diese Schwankungen im Meßkurvenverlauf sollen durch das hier vorgestellte Verfahren erkannt bzw. toleriert werden. Schwankungen der Übertragungskennlinie von Biosensoren gleichen Typs werden durch eine Kalibrierung bzw. Rekalibrierung berücksichtigt. Mit Hilfe des Kalibrierwertes kann dann aus I_0 und I_D die gesuchte Konzentration berechnet werden.

In diesem Beitrag wird ein auf diskreten HMM basierendes lernfähiges Meßwertverarbeitungssystem zur Analyse von Biosensor-Meßkurven vorgestellt. Da sich mit HMM zeitabhängige Signale gut modellieren lassen, wurde dieses Konzept zur Verarbeitung der Kurven gewählt. Dabei werden die HMM einerseits zur Unterscheidung verschiedener Kurvenformen, als auch zur Detektion des für die Messung relevanten Abschnittes der Meßkurven und zur Erkennung des Messungsendes genutzt.

Im Rahmen einer früheren Arbeit [11] wurde bereits ein einfaches, auch in Mikrocontrollern implementierbares Erkennungsverfahren für die Anwendung in einem Biosensor-Handmeßgerät vorgestellt. Dieses Verfahren arbeitet mit einer Fuzzy-Logik-Mustererkennung zur Detektion von Anstiegsverläufen innerhalb von Meßkurvenabschnitten. Die Analyse der Meßkurven erfolgt jedoch nur lokal und ohne Erkennung des Endes der Messung, außerdem ist das Verfahren nicht lernfähig. In diesen und in den anderen bereits genannten Punkten liefert das hier vorgestellte Verfahren eine Verbesserung.

Zur Meßwertverarbeitung bei Chemo- und Biosensoren existieren neben diesen weitere Ansätze mit unterschiedlichen Zielstellungen. Beispielsweise werden in [3]

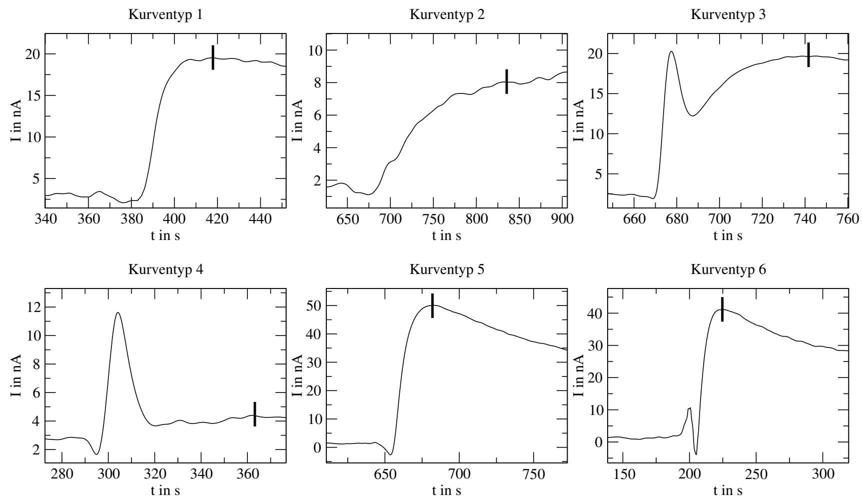

Abbildung 1. Unterschiedliche Kurvenformen mit eingezeichneten Meßpositionen für einen Sensortyp

genetische Algorithmen zur Kalibrierung elektrochemischer Sensoren benutzt. In [6] wird die Extrapolation von Meßkurvenabschnitten zur Beschleunigung der Messung verwendet. Zur Auswertung von Sensorarrays sind Neuronale Netze weit verbreitet [4,10,8].

Im folgenden wird zunächst auf die Modellierung der Meßkurven mittels HMM eingegangen. Anschließend werden die Vorverarbeitung und Merkmalsextraktion, die Trainings- sowie die Arbeitsphase des Systems erläutert. Schließlich werden die bisher erreichten experimentellen Resultate vorgestellt.

2 Systembeschreibung

2.1 Modellierung

Eine Biosensor-Meßkurve läßt sich in mehrere Abschnitte einteilen, die sich als eine Abfolge von Zuständen darstellen lassen (siehe Abbildung 2). Es liegt daher nahe, für die Modellierung (strikte) Links-Rechts-HMM zu verwenden.

Jede Kurve beginnt mit einer typischen Einstellphase. Diese Phase wird mit einer vom Kurventyp abhängenden Anzahl von n Zuständen $E_1, \ldots, E_n$ modelliert. Der Einstellphase schließt sich ein Bereich an, der die Meßphase einleitet. Diesem Bereich wird ein Zustand V zugeordnet. Es folgt der eigentliche Meßabschnitt, der allein für die Meßwertbildung relevant ist. Der Meßwert wird im Meßabschnitt durch Mittelung oder Maximumbildung ermittelt. Diese Zone (in der Regel ein bestimmter Peak oder ein Plateau) kann sich je nach Kurventyp in ihrer Lage und Form unterscheiden. Der Meßabschnitt wird mit dem Zustand M modelliert. Dem Meßabschnitt folgt ein Bereich, in dem der Strom in der

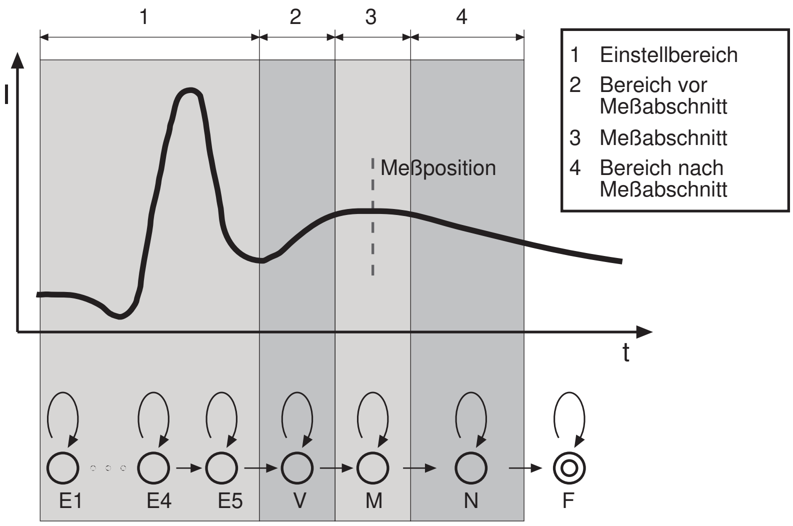

Abbildung 2. Zuordnung von Abschnitten einer Meßkurve zu den Zuständen eines HMM. Die Anzahl der Zustände E_i variiert von Kurventyp zu Kurventyp.

Regel konstant bleibt oder gleichmäßig ansteigt bzw. abfällt. Ihm wird der Zustand N zugeordnet. Wird dieser Bereich genügend lange durchlaufen, so ist das ein sicheres Zeichen zum Beenden der Messung. Das Erreichen des Endes der Messung wird durch einen ausgezeichneten Finalzustand F modelliert.

2.2 Vorverarbeitung und Merkmalsextraktion

Zur Rauschunterdrückung werden die Kurven in einem ersten Vorverarbeitungsschritt durch eine Mittelwertfilterung geglättet. Als Merkmale zur Generierung von Beobachtungssymbolen für die HMM werden die 1. und 2. Ableitung der Kurve verwendet. Hierzu werden die Meßwerte in einem Fenster der Größe 30 Punkte mit einer quadratischen Funktion approximiert, aus der dann die Ableitungen analytisch bestimmt werden. Die Ableitungen werden zu einem 2-dimensionalen Merkmalsvektor zusammengefasst. Zur Verarbeitung in einem diskreten HMM ist eine Vektorquantisierung (VQ) notwendig. Hierzu kommt eine Self-Organizing Map (SOM) nach Kohonen [7] mit 16 Vektoren zum Einsatz.[1] Die in jedem Zeitschritt t erhaltenen Symbole o_t bilden die Beobachtungssequenz $O = o_1, o_2, \ldots, o_T$.

2.3 Trainingsphase

Für jede Klasse von Kurven wird ein HMM mit Hilfe des Baum-Welch-Algorithmus trainiert. Da dieser die Modellparameter nur lokal maximieren kann, ist

[1] Es wurde das SOM_PAK 3.1 von T. Kohonen et al. verwendet.

die Wahl von geeigneten Initialparametern, speziell für die Emissionswahrscheinlichkeiten von großer Bedeutung [9]. Außerdem sollen hier, wie in Abschnitt 2.1 beschrieben, bestimmte Kurvenabschnitte durch spezielle Zustände des HMM repräsentiert werden. Daher werden in allen Trainingskurven die zur Meßwertbestimmung relevanten Abschnitte mit Hilfe einer dafür entwickelten Software auf Basis des Expertenwissens manuell markiert: In jeder Kurve werden zunächst die Zeitpunkte für den Start der Einstellphase (und damit der Messung), das Ende der Einstellphase, der Meßzeitpunkt und das Ende der Messung markiert. Der Bereich von ± 4s um den markierten Meßzeitpunkt wird als Meßabschnitt M festgelegt. Dadurch erhält man automatisch die Zeitpunkte für die Abschnitte vor V und nach N dem Meßabschnitt. In allen Bereichen werden die mittleren Symbolhäufigkeiten für alle Trainingskurven geschätzt und als initiale Emissionswahrscheinlichkeiten für die zugeordneten Zustände im Training verwendet. Ein wesentliches Problem neben der Detektion des Kurventyps sowie des Meßzeitpunktes ist die Erkennung des Endes der Messung. Dazu wurde ein spezieller Finalzustand F eingeführt. Es erwies sich als schwierig, die Übergangswahrscheinlichkeiten vom Zustand N in den Zustand F mit dem Baum-Welch-Algorithmus zu trainieren, da in der Regel die Kurvencharakteristik in dem Zeitraum, der den Zuständen N und F zugeordnet ist, im wesentlichen identisch ist. Die Übergangswahrscheinlichkeiten lassen sich jedoch auch aus der mittleren Verweildauer in diesem Zustand bestimmen [9]. Daher wird das HMM zunächst nur bis zum Zustand N trainiert und anschießend um den Finalzustand F erweitert, welcher die gleichen Emissionswahrscheinlichkeiten wie der Zustand N erhält. Anhand aller Trainingskurven wird die mittlere Verweildauer im Zustand N mit Hilfe der Markierungen geschätzt und daraus die Übergangswahrscheinlichkeiten vom Zustand N in den Zustand F berechnet.

2.4 Arbeitsphase

Bei der Analyse einer Abtastfolge $I(t)$ muß zunächst festgestellt werden, welches der trainierten HMM die aktuelle Meßkurve repräsentiert. Ist dieses Modell bestimmt, muß das Ende der Messung detektiert und gegebenenfalls der Meßabschnitt ermittelt werden. Hierzu wird folgendes Vorgehen vorgeschlagen (siehe auch Abbildung 3):

1. $t = 0$
2. Wiederhole die Schritte (a) bis (f) solange, bis der letzte Zustand der Zustandsfolge $q_1, \ldots, q_t$ gleich dem Endzustand F ist, oder die Maximalzeit $t_{\max}$ überschritten wurde:
   (a) $t = t + 1$
   (b) Messung eines neuen Abtastwertes $I(t)$.
   (c) Mit Merkmalsextraktion und VQ wird ein neues Symbol o_t erzeugt.
   (d) Bestimme mit dem Forward-Algorithmus für $o_1, \ldots, o_t$ die Wahrscheinlichkeit $P(o_1, \ldots, o_t | HMM_i)$ für jedes Modell i.
   (e) Wähle das Modell mit der maximalen Wahrscheinlichkeit. Dieses Modell wird als korrekte Interpretation des Kurvenverlaufs angenommen.

(f) Berechne mit dem Viterbi-Algorithmus die zur Beobachtungssequenz wahrscheinlichste Zustandsfolge $q_1, \ldots, q_t$ für das ausgewählte Modell. Jedem Zustand q_j der Zustandsfolge wird für die spätere Auswertung der entsprechende Stromwert I_j zugeordnet.

3. Wurde der Endzustand F erreicht, dann erfolgt die eigentliche Meßwertbildung: Anhand der ermittelten Zustandsfolge $q_1, \ldots, q_t$ werden alle Stromwerte $I_k, \ldots, I_{k+n}$ der Kurve, denen der Meßzustand M zugeordnet wurde, zur Meßwertbestimmung herangezogen. Das kann eine Mittelung über diese Werte oder eine Maximumbildung sein.
   Wurde die vorgegebenen Maximalzeit überschritten und der Endzustand F nicht erreicht, muß überprüft werden, ob wenigstens der Meßzustand M erreicht wurde. Ist dies der Fall, kann trotzdem ein Meßwert gebildet werden, es wird jedoch eine Warnung ausgegeben. Wurde der Meßzustand M nicht erreicht, dann wird die Messung verworfen. Die Gründe dafür können sein: ein neuer oder unbekannter Kurventyp, zu starke Störungen (Rauschen) oder ein defekter bzw. verbrauchter Sensor.

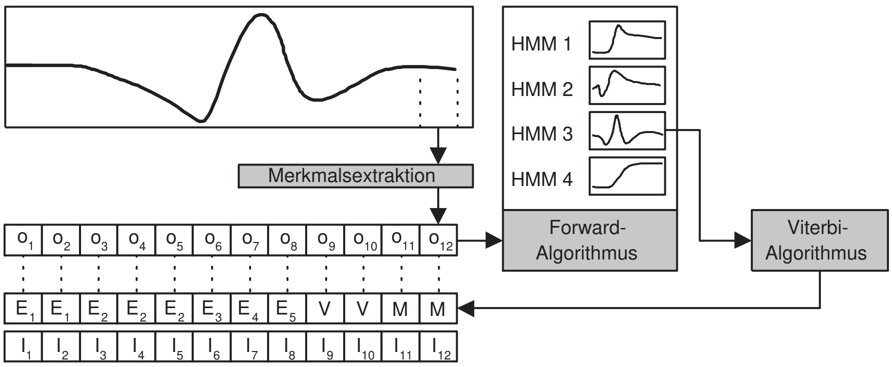

Abbildung 3. Prinzipieller Ablauf in einem Abtastschritt (vereinfacht)

3 Experimentelle Resultate

Zum Test des Verfahrens wurden mit 5 verschiedenen Glucose-Sensoren gleichen Typs 205 Meßkurven mit einem Abtastintervall von 250 ms aufgezeichnet. Um einen guten Querschnitt über die möglichen Kurvenformen zu erhalten, wurden bei der Kurvenaufzeichnung einerseits verschiedene Analytkonzentrationen (Kalibrierkonzentration im Verhältnis 1:1 und 1:5), andererseits verschiedene Meßverfahren (Auftropfen des Analyten oder Eintauchen des Sensors) verwendet. Entgegen ursprünglicher Annahmen ließen sich die Kurven nicht ausschließlich anhand der Analytkonzentration bzw. des Meßverfahrens in Klassen einteilen.

Daher wurden die Kurven einer visuellen Beurteilung unterzogen und so in 6 verschiedene Klassen unterteilt (siehe auch Abbildung 1). Als Unterscheidungskriterium dienten die Form der Einstellphase und die Meßposition mit den dazugehörenden charakteristischen Kurvenabschnitten. Außerdem wurde versucht, eine spezielle Klasse für verbrauchte Sensoren zu verwenden (Kurventyp 2). Für jede Kurve wurden die Markierungspunkte für die einzelnen Kurvenabschnitte manuell gesetzt (siehe Abschnitt 2.3). Das Training der einzelnen HMM erfolgte mittels Baum-Welch-Algorithmus (Batch-Training), wobei die leave-one-out-Methode angewendet wurde. Dabei muß das Training für jede Testkurve wiederholt werden, die Kurve selbst wird hierzu aus der Traingsmenge entfernt. Um die Genauigkeit zu ermitteln, wurde anstelle der Konzentrationen der manuell bestimmte Meßwert I_{D} mit dem Meßwert I_{D} aus dem HMM-Verfahren verglichen. Die Ergebnisse sind in Tabelle 1 zusammengefaßt.
Die Tests ergaben gute Erkennungsraten und für den Anwendungsfall ausreichende Genauigkeiten. Die vorhandenen Fehlklassifikationen sind auf relativ große Ähnlichkeiten zwischen einigen Kurventypen (insbesondere Typ 1 und 2 sowie Typ 5 und 6) zurückzuführen. Dabei ist zu bemerken, daß die Klassifikationsrate allein eigentlich nur bei der Erkennung von verbrauchten Sensoren relevant ist. Trotzdem ist zu sehen, daß bei richtiger Klassifikation die Meßgenauigkeit steigt (Ausnahme: Kurventyp 3). Beim Kurventyp 2 und 4 sind höhere Ungenauigkeiten zu verzeichnen. Diese sind auf Störsignale zurückzuführen, die aufgrund der bei diesen Kurventypen auftretenden kleinen Signalströmen die Genauigkeit relativ stark beeinflussen.

Tabelle 1. Verwechslungsmatrix und Genauigkeiten

Verwechslungsmatrix		Kurventyp (soll)					
		1	2	3	4	5	6
Erkennungsrate in %	1	84	9	0	0	0	0
	2	4	82	0	0	0	0
	3	5	9	92	7	0	0
	4	0	0	8	93	0	0
	5	4	0	0	0	96	14
	6	3	0	0	0	4	86
mittlerer rel. Fehler in %	richtig klassifiziert	2.8	5.1	4.5	9.6	0.9	0.3
	falsch klassifiziert	2.8	5.8	1.7	10.1	2.8	3.0
	gesamt	2.8	5.2	4.3	9.6	1.0	0.7
Standard-abweichung in %	richtig klassifiziert	4.0	7.4	4.9	7.5	1.4	0.4
	falsch klassifiziert	5.0	8.8	0	0	0	2.9
	gesamt	4.2	7.7	4.7	7.2	1.4	1.5

4 Zusammenfassung und Ausblick

In diesem Beitrag wurde ein auf diskreten HMM basierendes Verfahren zur Analyse von Biosensor-Meßkurven vorgestellt. Bei den ersten Tests wurden bereits gute Erkennungsraten und Genauigkeiten erreicht. Es zeigte sich, daß diskrete HMM zur Behandlung der hier vorliegenden Problemstellung gut geeignet sind. Zur Erhöhung von Genauigkeit und Erkennungsrate sind Verbesserungen sowohl bei der Datenvorverarbeitung, als auch bei der Modellierung geplant. Um die Datenvorverarbeitung robuster gegen die teilweise sehr starken Störsignale zu gestalten, wird eine verbesserte Filterung bzw. Approximation des Kurvenverlaufes angestrebt. Um die Anzahl von einander ähnlichen Modellen zu reduzieren, sollen ähnliche Kurventypen durch Mischung von Modellen zusammengefaßt werden. Unterschiede z.B. in der Einstellphase werden dann durch Verzweigungen in den HMM modelliert. Außerdem sollen neue bzw. verbesserte Modelle zur Detektion von defekten oder verbrauchten Sensoren genutzt werden.
Die für weitere Experimente notwendige Datenbasis an Meßkurven wird ständig durch weitere Sensorexemplare und Sensortypen vergrößert und aktualisiert. Letztendlich ist geplant, das Verfahren für die Anwendung in einem neuen Biosensor-Handmeßgerät auf ein Mikrocontrollersystem zu portieren.

Literatur

1. R.-D. Bippus. Pseudo zweidimensionale HMM zur Erkennung handgeschriebener Beiträge. In: E. Paulus, F. M. Wahl (Hrsg.). Mustererkennung 97, 19. DAGM-Symposium Braunschweig, Springer-Verlag, Berlin (1997) 245–253
2. S. Eickeler, S. Müller, G. Rigoll. Gesichtserkennung mit Hidden Markov Modellen. In: W. Förstner, J. M. Buhmann, A. Faber (Hrsg.). Mustererkennung 99, 21. DAGM-Symposium Bonn, Springer-Verlag, Berlin (1999)
3. W. Fichtner, H. Kaden, S. Herrmann. Kalibrierung elektrochemischer Sensoren unter Nutzung genetischer Algorithmen. In: Technisches Messen 61. Oldenbourg-Verlag (1994) 432–438
4. J. W. Gardner, E. L. Hines Pattern Analysis Techniques. In: E. Kress-Rogers (Hrsg.). Handbook of Biosensors and Electronic Noses. CRC Press (1997) 633–651
5. E. A. H. Hall. Biosensoren. Springer-Verlag, Berlin/Heidelberg (1995)
6. K. Hirose, c/o Kyoto Daliichi. Method of measuring sample by enzyme electrodes. European patent 0 623 681 A1. (1994)
7. T. Kohonen. Self-Organizing Maps. Springer-Verlag, Heidelberg (1995)
8. G. Niebling. Strategien der Signalverarbeitung in der Chemosensorik – Konventionelle Methoden und neuronale Netze. Herbert Utz Verlag Wissenschaft, München (1996)
9. L. R. Rabiner and B. H. Juang. Fundamentals of Speech Recognition. Prentice Hall (1993)
10. S. Vaihinger, W. Göpel. Multi-Component Analysis in Chemical Sensing. In: W.Göpel, T. A. Jones, M. Kleitz, J. Lundström, T.Seiyama (Hrsg.). Sensors – A Comprehensive Survey, vol. 2, VCH Weinheim (1991) 192–237
11. J. Weitzenberg. Methoden der Fuzzy-Logik zur Auswertung von Meßsignalen amperometrischer Biosensoren. Diplomarbeit, MLU Halle-Wittenberg (1998)