

# Kapitel 4

## Statistische Modellierung in zwei Dimensionen

Die in diesem Kapitel vorgestellten statistischen Modellierungsverfahren sind aufgrund ihrer zweidimensionalen Struktur ideal geeignet, um Bilder zu modellieren. Das theoretisch geeignetste Modell, welches Gegenstand des folgenden Unterkapitels ist, basiert auf den sog. Markov-Random-Fields (MRFs), die eine mehrdimensionale Erweiterung der Markov-Quellen darstellen.

### 4.1 Markov-Random-Fields

In Kapitel 2 wurden die Hidden-Markov-Modelle als eine Erweiterung der Markov-Quellen vorgestellt, die einen *einfachen* statistischen Prozeß darstellen. Der Übergang zwischen den Modellzuständen ist bei der Markov-Quelle erster Ordnung durch die Übergangsmatrix  $A$  gegeben, deren Elemente durch  $a_{ij} = P(q_t = S_j | q_{t-1} = S_i)$  darstellbar sind. Diese Definition des Zustandsübergangs charakterisiert die Markov-Quelle als einen eindimensionalen Prozeß, der zudem kausal ist. Es handelt sich mithin um ein Modell, das sehr gut geeignet ist, um zeitabhängige Prozesse zu beschreiben.

Kartesisch abgetastete Bilder sind von zweidimensionaler Art und ihre Elemente (Pixel) verfügen nicht über eine natürliche kausale Ordnung. Aus diesen Gründen besitzt das den Markov-Modellen entsprechende statistische Modell für zweidimensionale, nichtkausale Prozesse eine erheblich komplexere Struktur. Das in der wissenschaftlichen Literatur gebräuchlichste statistische Modell für Bilder ist das zweidimensionale Wahrscheinlichkeitsfeld mit lokalen Abhängigkeiten, das als *Markov-Random-Field* (MRF) bekannt ist (siehe z.B. [Gem85, Der89] und [Li95]). Die Theorie der MRFs wurde ursprünglich in der statistischen Physik für interagierende Partikel wie z.B. Moleküle oder atomare Magnete entwickelt. Mit Hilfe der Markov-Random-Fields konnten Vorgänge beim Ferromagnetismus (Ising-Modell), in idealen Gasen, sowie in zweiwertigen Metallegierungen modelliert und

erklärt werden. Später wurden die MRFs in der Bildverarbeitung, beispielsweise für die Restaurierung von verrauschten oder optisch verzerrten Bildern ([Gem85]), der Segmentierung von Bildregionen mit homogener Textur ([Sim88]) oder der Oberflächenrekonstruktion aus Abstandsdaten eingesetzt.

Im folgenden wird angenommen, daß das MRF bei der Modellierung von Bildern eingesetzt wird und die Beispiele werden entsprechend gewählt. Bei Markov-Random-Fields werden Aufgabenstellungen der Bildverarbeitung umformuliert und als die Aufgabe dargestellt, Bildpunkten *Label* zuzuordnen. Solche Label können beispielsweise Indizes sein, die Texturklassen oder farblich homogenen Bildregionen entsprechen. Formal wird die Label-Aufgabe mit Hilfe von *Sites* und einer Label-Menge definiert. Die Menge der  $N$  Sites sei gegeben durch

$$S = \{s_1, s_2, \dots, s_N\} \quad (4.1)$$

Sites repräsentieren Punkte oder Regionen im Bildraum, wie z.B. Pixel oder Bildmerkmale. Für den Fall, daß die Sites Pixel eines  $m \times m$  dimensionalen Bildes repräsentieren, entspricht die folgende Indexierung eher den üblichen Notationen in der Bildverarbeitung:

$$S = \{s_{11}, s_{12}, \dots, s_{mm}\} \quad (4.2)$$

Es geht die in der Literatur zu MRFs gebräuchlichere Gleichung 4.1 aus Gleichung 4.2 durch Umindizierung hervor. Die Sites  $S$  interagieren miteinander durch das *Nachbarschaftssystem*. Ein Nachbarschaftssystem ist folgendermaßen gegeben:

$$\mathcal{N} = \{\mathcal{N}_i, s_i \in S\} \quad (4.3)$$

Die Sites  $\mathcal{N}_i$  sind die Nachbarn der Site  $s_i \in S$ . Von besonderem Interesse für die Bildverarbeitung sind homogene Nachbarschaftssysteme auf rechteckigem Raster der Form

$$\mathcal{N}_{ij} = \{s'_{ij} \in S, s_{kl} \in S, 0 < (k-i)^2 + (l-j)^2 \leq r\} \quad (4.4)$$

Es ist dabei zu beachten, daß an den Randpunkten der Bilder noch entsprechende Korrekturen vorzunehmen sind. Für  $r = 1$  liegt eine sog. Nachbarschaftsbeziehung erster Ordnung vor, bei der die Nachbarn der Site  $s_{ij}$  durch die Menge

$$\mathcal{N}_{ij} = \{s_{i,j-1}, s_{i,j+1}, s_{i-1,j}, s_{i+1,j}\} \quad (4.5)$$

gegeben ist. Den Sites können diskrete Label aus der Menge

$$\mathcal{L} = \{l_1, \dots, l_M\} \quad (4.6)$$

zugeordnet werden.

Unter Verwendung der bisher definierten Größen kann nun das MRF formal definiert werden. Seien  $F = \{F_1, \dots, F_N\}$  Zufallsvariablen über einer Menge von Sites  $S = \{s_1, \dots, s_N\}$

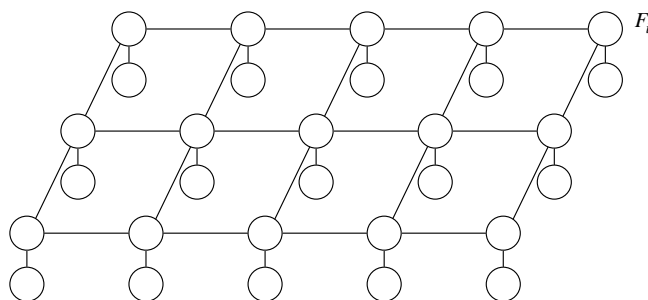


Abbildung 4.1: Darstellung eines Markov-Random-Fields mit Nachbarschaftsbeziehung erster Ordnung als ungerichteten Graph (aus [Luc95])

unter Verwendung eines Nachbarschaftssystems  $\mathcal{N}$ , so ist der statistische Prozeß  $F$  ein Markov-Random-Field, falls

$$P(F = f) > 0 \quad \text{und} \\ P(F_i = f_i | F_k = f_k, k \neq i) = P(F_i = f_i | F_k = f_k, s_k \in \mathcal{N}_i) \quad (4.7)$$

In der obigen Gleichung ist mit  $f_i$  eine Realisierung der Zufallsvariablen  $F_i$  bezeichnet.  $f_i$  nimmt Werte aus der Menge der Label  $\mathcal{L}$  an. Gleichung 4.7 stellt eine mehrdimensionale und nicht notwendigerweise kausale Verallgemeinerung der in Gleichung 2.1 für Markov-Quellen vorgestellten eindimensionalen Markovbedingung dar. Das einer Site zugeordnete Label hängt ausschließlich von den Labels der benachbarten Sites ab. Diese Abhängigkeit kann für den Fall einer Nachbarschaftsbeziehung erster Ordnung (Gleichung 4.5) in Form eines graphischen Modells visualisiert werden (siehe auch Abb. 2.4). Abb. 4.1 zeigt die zweidimensionale Anordnung der Sites und die den Sites zugeordneten Labels als ungerichteten Graphen. Wie schon in Unterkapitel 2.2.6 bzw. in Abb. 2.4 deutet die Abwesenheit von Verbindungen in Abb. 4.1 auf statistische Unabhängigkeit hin. An dieser Stelle sei angemerkt, daß sowohl für Markov-Quellen bzw. Hidden-Markov-Modelle als auch für MRFs graphische Modelle existieren (vgl. Abb. 2.4 und Abb. 4.1), eine Darstellung als endlicher stochastischer Automat ist jedoch nur für die erstgenannten, eindimensionalen Modelle möglich (siehe Abb. 2.1). Der Grund hierfür liegt in der fehlenden Kausalität der Markov-Random-Fields. Die Abb. 4.1 und die Markovbedingung in Gleichung 4.7 beschreiben die *lokale* Charakteristik der MRFs. Neben der Möglichkeit, ein MRF über die lokalen Eigenschaften und mithin über die bedingten Wahrscheinlichkeiten  $P(F_i = f_i | F_k = f_k, s_k \in \mathcal{N}_i)$  zu definieren, existiert auch die Möglichkeit die globalen Eigenschaften zur Definition zu verwenden. Das Hammersly-Clifford Theorem besagt, daß jedes Wahrscheinlichkeitsfeld  $F$  auf  $S$  unter Verwendung des Nachbarschaftssystems  $\mathcal{N}$  genau dann ein MRF ist, falls  $P(F = f)$  in Form einer Gibbs'schen Verteilung darstellbar ist (Beweis siehe [Li95]). Die Gibbs'sche

Verteilung ist gegeben durch:

$$P(F = f) = \frac{1}{Z} e^{-\frac{U(f)}{T}} \quad \text{mit} \quad (4.8)$$

$$Z = \sum_{\text{alle } f} e^{-\frac{U(f)}{T}}$$

In Gleichung 4.8 ist  $Z$  die sog. *Partition Function*,  $T$  eine Konstante, die auch Temperatur genannt wird und  $U(f)$  die Energiefunktion. Die Energie kann berechnet werden durch

$$U(f) = \sum_{c \in C} V_c(f) \quad (4.9)$$

$U(f)$  ist die Summe über alle sog. *Clique-Potentiale*  $V_c(f)$ . Cliques sind Untermengen der Sites ( $C \subseteq S$ ) und sie sind durch die Eigenschaft charakterisiert, daß jedes Site-Paar in einem Clique ein Nachbarpaar ist. Dies soll am Beispiel der Nachbarschaftsbeziehung erster Ordnung (siehe Gleichung 4.5) verdeutlicht werden. Die Cliques für die Site  $s_{ij}$  und dem letztgenannten Nachbarschaftssystem bilden die folgende Menge von Sites:

$$C = \{\{s_{i,j}\}, \{s_{i,j}, s_{i,j+1}\}, \{s_{i,j}, s_{i,j-1}\}, \{s_{i,j}, s_{i+1,j}\}, \{s_{i,j}, s_{i-1,j}\}\} \quad (4.10)$$

Die Clique-Potentiale sind ein Bestandteil des festgelegten Modells und charakterisieren die lokalen Interaktionen der Sites. In [Li95] ist aus Gleichung 4.8 eine Gleichung für die Wahrscheinlichkeit  $P(F_i = f_i | F_k = f_k, s_k \in \mathcal{N}_i)$ , also die Wahrscheinlichkeit, daß die Site  $s_i$  den Zustand  $f_i$  einnimmt, unter der Annahme, daß die Nachbarsites festgelegt sind, abgeleitet worden. Diese Gleichung ist:

$$P(F_i = f_i | F_k = f_k, s_k \in \mathcal{N}_i) = \frac{e^{-\sum_{c \in C} V_c(f)}}{\sum_{f_k; s_k \in \mathcal{N}_i; k \neq i} e^{-\sum_{c \in C} V_c(f')}} \quad (4.11)$$

Diese bedingte Wahrscheinlichkeit entspricht einer *nichtkausalen* Verallgemeinerung der Übergangswahrscheinlichkeit  $a_{ij}$  bei der Markov-Quelle bzw. dem Markov-Modell (vgl. die Gleichungen 2.2 und 2.3).

Zusammenfassend kann festgestellt werden, daß die Markov-Random-Fields eine nicht-kausale, multidimensionale Erweiterung der Markov-Quellen sind. In Gleichung 4.11 ist eine den Übergangswahrscheinlichkeiten der Markov-Quellen entsprechende Größe definiert worden und somit ist ein dem *ersten* stochastischen Prozeß der Hidden-Markov-Modelle entsprechender Prozeß vorhanden. Auf ähnliche Weise, wie bei den HMMs kann nun auch bei den MRFs ein zweiter stochastischer Prozeß eingeführt werden. Dies geschieht auf folgende Weise: Die Label entsprechen nicht mehr direkt z.B. einem Grauwert oder allgemeiner formuliert, einem Merkmal in einem Bild, sondern stellen einen Index einer Wahrscheinlichkeitsverteilung bzw. -dichte dar. Die Ausgabe von Grauwerten erfolgt dann wie bei den Hidden-Markov-Modellen über diesen zweiten stochastischen Prozeß. Ein solches *Hidden*

Markov-Random-Field stellt ein sehr geeignetes Modell für Bilder dar, da der nichtkausale Aufbau der Bilder modelliert werden kann und durch die Ausgabe von Merkmalen über Wahrscheinlichkeitsverteilungen zusätzliche Variationen berücksichtigt werden. Für die praktische Verwendung in der Mustererkennung erweist sich jedoch das Fehlen von effizienten Algorithmen zur Parameterbestimmung bzw. zur Bestimmung der Produktionswahrscheinlichkeiten als Hindernis. Durch die fehlende Kausalität sind rekursive Verfahren, wie Baum-Welch- und Viterbi-Algorithmus nicht anwendbar ([Luc95]) und es müssen sehr rechenaufwendige Verfahren für die Bestimmung der Produktionswahrscheinlichkeit eingesetzt werden, wie z.B. *Stochastic Relaxation* ([Gem85]). Dieses Fehlen effizienter Algorithmen hat dazu geführt, daß die Modellierung mit MRFs im Bereich der zwei- und mehrdimensionalen Muster nicht dieselbe, dominierende Verbreitung gefunden hat, wie die Hidden-Markov-Modelle in der Modellierung eindimensionaler Muster und insbesondere von Sprachmustern. Da für die nichtkausalen MRFs keine effizienten Algorithmen existieren, wurde in der wissenschaftlichen Literatur eine Vielzahl von kausalen Variationen untersucht. Ein wichtiger Beitrag im Kontext dieser Arbeit ist die Betrachtung von zweidimensionalen Hidden-Markov-Modellen, die von Levin und Pieraccini in der Arbeit [Lev92] vorgestellt wurden.

## 4.2 Zweidimensionale Hidden-Markov-Modelle

Zweidimensionale Hidden-Markov-Modelle (2DHMMs), die auch planare HMMs genannt werden, sind kausale MRFs mit einem zusätzlichen zweiten stochastischen Prozeß, der durch die Ausgabe von Observations über Ausgabeverteilungen bzw. Ausgabedichten gegeben ist. Die Observations  $o_{xy}$ , beispielsweise Grauwerte eines Bildes, werden als zweidimensionale Matrix, die in den Bilddimensionen  $x$  mit  $(1 \leq x \leq X)$  und  $y$  mit  $(1 \leq y \leq Y)$  indiziert sind, betrachtet. Statistische Abhängigkeiten bestehen beim 2DHMM zu den beiden *Vorgängern* in beiden Dimensionen. Der Zustand  $q_{(x,y)}$  am Ort  $(x,y)$  hängt ausschließlich von den Zuständen  $q_{(x-1,y)}$  und  $q_{(x,y-1)}$  ab. Dies ist eine kausale Markovbedingung erster Ordnung für zwei Dimensionen und kann formal folgendermaßen angegeben werden:

$$P(q_{(x,y)} = S_{(i,j)} | q_{(\tilde{x},\tilde{y})} = S_{(\tilde{k},\tilde{l})}; 1 \leq \tilde{x} \leq (x-1), 1 \leq \tilde{k} \leq N_x, 1 \leq \tilde{y} \leq (y-1), 1 \leq \tilde{l} \leq N_y) = \quad (4.12)$$

$$P(q_{(x,y)} = S_{(i,j)} | q_{(x-1,y)} = S_{(k,l)}, q_{(x,y-1)} = S_{(m,n)})$$

In der obigen Gleichung ist mit  $S_{(i,j)}$  der Zustand mit Index  $(i,j)$  auf einem kartesischen Raster von Zuständen bezeichnet, während  $q_{(x,y)}$  die Zufallsvariable für die Einnahme eines Zustandes aus der Menge der möglichen Zustände  $(S_{(1,1)}, S_{(1,2)}, \dots, S_{(2,1)}, \dots, S_{(N_x, N_y)})$  am Ort  $(x,y)$  darstellt (vgl. Gleichung 2.1 für den eindimensionalen Fall). Wie schon bei den eindimensionalen HMMs, gehört zu jedem Modellzustand  $S_{(i,j)}$  eine z.B. diskrete Ausgabeverteilung  $b_{(i,j)}(k)$  über einem festgelegten Alphabet. Aus Gleichung 4.12 lassen sich Übergangs-

wahrscheinlichkeiten zwischen Modellzuständen der folgenden Form ableiten (aus [Lev92]; Siehe auch [Mer00]):

$$A_{ij,kl,mn} = P(q_{(x,y)} = S_{(m,n)} | q_{(x-1,y)} = S_{(i,j)}, q_{(x,y-1)} = S_{(k,l)}) \quad (4.13)$$

Zusätzlich sind zwei Ränder des Bildes ( $x = 1$  bzw.  $y = 1$ ) als Sonderfälle zu betrachten. Hier können die folgenden Übergangswahrscheinlichkeiten verwendet werden, die den eindimensionalen Übergangswahrscheinlichkeiten über den diskreten Werten  $x$  und  $y$  entsprechen (vgl. Gleichung 2.2):

$$a_{ij,kl}^V = P(q_{(x=1,y)} = S_{(k,l)} | q_{(x=1,y-1)} = S_{(i,j)}) \quad \text{für } (x = 1, y) \quad (4.14)$$

$$a_{ij,kl}^H = P(q_{(x,y=1)} = S_{(k,l)} | q_{(x-1,y=1)} = S_{(i,j)}) \quad \text{für } (x, y = 1) \quad (4.15)$$

Zusätzlich zu den Übergangswahrscheinlichkeiten und den Ausgabewahrscheinlichkeiten fehlt für die vollständige Definition eines Modells noch die Angabe der Wahrscheinlichkeiten für die Einnahme des Anfangszustands, die gegeben sind durch:

$$\pi_{ij} = P(q_{(1,1)} = S_{(i,j)}) \quad (4.16)$$

Nach einer vollständigen Definition eines Modells  $\lambda$  kann die Produktionswahrscheinlichkeit für eine gegebene Symbolmatrix  $O_{XY}$  berechnet werden. Das Finden einer effizienten Methode zur Berechnung der Produktionswahrscheinlichkeiten war in Kapitel 2 als eines der zu lösenden Aufgaben definiert worden, um die Markov-Modelle zur Musterklassifikation verwenden zu können. Der direkte Weg, um die gesuchte Wahrscheinlichkeit bei gegebener Matrix  $O_{XY}$  zu berechnen, führt zunächst über die Wahrscheinlichkeit  $P(Q|\lambda)$ , der Wahrscheinlichkeit, daß eine Zustandsfolge  $Q$  durchlaufen wurde. Diese ist in [Lev92] angegeben als:

$$P(Q|\lambda) = \pi_{q_{(1,1)}} \prod_{x=2}^X a_{q_{(x-1,1)}, q_{(x,1)}}^H \prod_{y=2}^Y a_{q_{(1,y-1)}, q_{(1,y)}}^V \prod_{y=2}^Y \prod_{x=2}^X A_{q_{(x-1,y)}, q_{(x,y-1)}, q_{(x,y)}} \quad (4.17)$$

Die Verbundwahrscheinlichkeit für das gemeinsame Eintreten der Zustandsfolge  $Q$  und der Observation  $O_{XY}$  ist

$$P(O_{XY}, Q|\lambda) = P(O_{XY}|Q, \lambda) \cdot P(Q|\lambda) = \prod_{x=1}^X \prod_{y=1}^Y b_{q_{x,y}}(o_{xy}) \pi_{q_{(1,1)}} \prod_{x=2}^X a_{q_{(x-1,1)}, q_{(x,1)}}^H \prod_{y=2}^Y a_{q_{(1,y-1)}, q_{(1,y)}}^V \prod_{y=2}^Y \prod_{x=2}^X A_{q_{(x-1,y)}, q_{(x,y-1)}, q_{(x,y)}} \quad (4.18)$$

Schließlich ergibt sich die gesuchte Produktionswahrscheinlichkeit als die Summation dieser Verbundwahrscheinlichkeit über alle möglichen Zustandsfolgen zu:

$$P(O|\lambda) = \sum_{\text{alle } Q} P(O, Q|\lambda) \quad (4.19)$$

Es ist ebenso wie bei den eindimensionalen Hidden-Markov-Modellen möglich, diese Produktionswahrscheinlichkeit unter Verwendung der wahrscheinlichsten Zustandsfolge  $Q^*$  anzunähern. Die Bestimmung beider Wahrscheinlichkeiten erfordert jedoch  $(N_X N_Y)^{XY}$  Berechnungen (siehe [Lev92]). Die Anzahl an erforderlichen Berechnungen wächst also mit der Bildgröße exponentiell an. Diese Komplexität läßt sich nicht auf die gleiche Weise, wie im eindimensionalen Fall reduzieren, da sowohl der Baum-Welch-Algorithmus, als auch der Viterbi-Algorithmus nur bei den einfachen statistischen Abhängigkeiten der eindimensionalen Modelle anwendbar sind ([Lev92, Li99, Mer00]). Die Komplexität der Algorithmen kann erheblich reduziert werden, wenn die nichtlinearen Verzerrungen der Zeilen und Spalten eines Bildes als unabhängig voneinander angesehen werden. Diese Annahme führt zu den im nächsten Unterkapitel vorgestellten sog. pseudo zweidimensionalen Hidden-Markov-Modellen.

### 4.3 Pseudo zweidimensionale Hidden-Markov-Modelle

Pseudo zweidimensionale Hidden-Markov-Modelle (P2DHMMs) wurden von Agazzi und Kuo in ([Aga93b]) vorgestellt. Durch den Verzicht auf eine Modellierung, die Musterverzerrungen in beiden Dimensionen gemeinsam betrachtet, also ein Verzicht auf Zustandsübergänge wie die in Gleichung 4.13 für zweidimensionale Hidden-Markov-Modelle vorgestellten Modellgrößen  $A_{ij,kl,mn}$ , konnten für das Training und die Erkennung effiziente Algorithmen gefunden werden. Diese effizienten Algorithmen basieren auf dem in [Kuo94] vorgestellten sog. zweifachverschachtelten Viterbi-Algorithmus, der eine Variante des in Unterkapitel 2.2.2 für eindimensionale HMMs beschriebenen Algorithmus darstellt. Durch das Vorhandensein effizienter Trainings- und Erkennungsalgorithmen wurden die P2DHMMs, im Gegensatz zu den bisher in diesem Kapitel vorgestellten statistischen Modellen (MRF bzw. 2DHMM) vielfach für die Erkennung von zweidimensionalen Mustern eingesetzt.

Anwendungsbeispiele für pseudo zweidimensionale Hidden-Markov-Modelle sind die Gesichtserkennung [Sam94b, Eic00], inhaltsbasierte Bilddatenbankabfragen [Lin97], die Erkennung von gedruckter Schrift (OCR) [Aga93a] und von Handschrift [Bip97, Bip00]. In der vorliegenden Arbeit werden ebenfalls die P2DHMMs verwendet und um die Fähigkeit erweitert, in komplexe Umgebungen eingebundene zweidimensionale Muster, auf integrierte Weise, zu Segmentieren und zu Klassifizieren. Dieses integrierte Auffinden und Erkennen wird durch die Modellierung mit zusätzlichen Umgebungszuständen ermöglicht und ist in dem sich anschließenden Kapitel 5 ausführlich beschrieben. Eine weitere Bedeutung kommt den P2DHMMs in Kapitel 6 zu, da sie hier als Baustein einer im Rahmen dieser Arbeit entwickelten pseudo dreidimensionalen Hidden-Markov-Modell-Struktur verwendet werden. Zunächst ist es jedoch erforderlich, die P2DHMMs formal zu definieren.

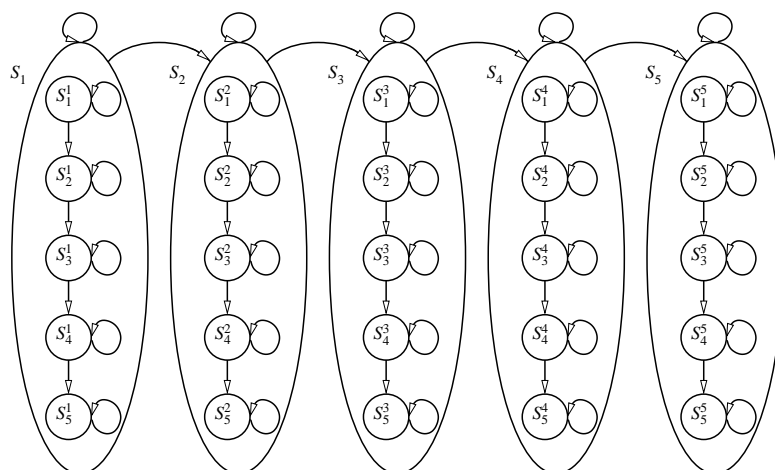


Abbildung 4.2: Pseudo zweidimensionales Hidden-Markov-Modell

### 4.3.1 Modelldefinition der pseudo zweidimensionalen Hidden-Markov-Modelle

Pseudo zweidimensionale HMMs modellieren, im Gegensatz zu den MRFs und den 2DHMMs, die Abhängigkeiten von räumlich benachbarten Merkmalen durch einen zweistufigen, hierarchischen Prozeß. Der übergeordnete statistische Prozeß modelliert die Abhängigkeiten von beispielsweise Bildspalten mit einem Markov-Modell erster Ordnung. Die Bildspalten selbst werden mit weiteren Markov-Modellen modelliert, die in das übergeordnete Modell eingebunden sind. Die Abb. 4.2 zeigt die Darstellung eines P2DHMMs, die das Modell als zweistufigen statistischen Automaten interpretiert. Dargestellt sind fünf, mit  $S_1, \dots, S_5$  bezeichnete, übergeordnete Zustände, die üblicherweise in der englischsprachigen Literatur mit *Superstates* bezeichnet werden ([Aga93b, Aga93a, Kuo94]). In dieser Arbeit wird für die übergeordneten Zustände die in [Bip00] eingeführte Bezeichnung *Metazustand* verwendet. In Abb. 4.2 sind ebenfalls die den Metazuständen zugeordneten eindimensionalen HMMs mit jeweils fünf Zuständen ( $S_1^n, \dots, S_5^n; 1 \leq n \leq 5$ ) dargestellt. Liegt nun ein zweidimensionales Muster  $O_{XY}$  in Form einer  $X \times Y$  Matrix vor, so kann dieses von dem P2DHMM auf folgende Weise modelliert werden: Jede Bildspalte ( $o_{xy}, 1 \leq y \leq Y, x = \text{const.}$ ) wird einem Metazustand zugeordnet, was eine nichtlineare Musterverzerrung in horizontaler Richtung ermöglicht. Zusätzlich werden die Bildspalten selbst von eindimensionalen Hidden-Markov-Modellen modelliert, und dies ermöglicht eine nichtlineare Musterverzerrung in vertikaler Richtung. Auf die gleiche Weise können auch Bildzeilen den Metazuständen zugeordnet werden. Die Unterscheidung in entweder zeilen- oder spaltenbasiert Modellierung verdeutlicht den hierarchischen Aufbau der P2DHMMs.

Formal wird ein pseudo zweidimensionales Hidden-Markov-Modell ( $\Lambda$ ) durch die folgenden Parameter beschrieben. In der höheren hierarchischen Modellebene ist die Anzahl  $N$  der Metazustände ( $S_1, \dots, S_N$ ) zu spezifizieren. Diesen Zuständen sind Übergangswahr-



scheinlichkeiten  $a_{ij}$  zugeordnet, die definiert sind durch:

$$a_{ij} = P(q_x = S_j | q_{x-1} = S_i) \quad (4.20)$$

In Gleichung 4.20 ist mit  $q_x$  die Zufallsvariable für die Einnahme eines Metazustands für die Bildspalte  $x$  bezeichnet. Es wird an dieser Stelle deutlich, daß bei der Modellfestlegung auch die *Ausrichtung* des P2DHMMs angegeben werden muß. Die Wahrscheinlichkeit für die Anfangszustände der höheren Hierarchieebene ist gegeben durch:

$$\pi_j = P(q_1 = S_j) \quad (4.21)$$

Die Modellgrößen der Gleichungen 4.20 und 4.21 können zu Matrizen bzw. Vektoren zusammengefaßt werden. Jedem Metazustand  $S_j$  ist ein eindimensionales HMM zugeordnet. Diese können wie in Kapitel 2 dargestellt, definiert werden. Jeder Modellparameter erhält einen zusätzlichen Index, der die Zugehörigkeit zum entsprechenden Metazustand angibt. Es ergeben sich mithin die folgenden Parameter für die Modelle  $\lambda_1, \dots, \lambda_N$ :

- $N^j$  ist die Anzahl der Zustände  $(S_1^j, \dots, S_{N^j}^j)$  des dem  $j$ -ten Metazustand  $S_j$  zugeordneten Modells.
- Die Übergangswahrscheinlichkeiten  $a_{kl}^j$  sind folgendermaßen definiert:

$$a_{kl}^j = P(q_{x,y} = S_l^j | q_{x,y-1} = S_k^j) \quad (4.22)$$

Dabei ist mit  $q_{x,y}$  die Zufallsvariable für die Einnahme eines Zustands für die Bildspalte  $x$  und die Bildzeile  $y$  bezeichnet.

- Die Wahrscheinlichkeiten für die Anfangszustände sind gegeben durch:

$$\pi_i^j = P(q_{x,1} = S_i^j) \quad (4.23)$$

- Die Ausgabeverteilungen können wie im eindimensionalen Fall diskret oder kontinuierlich sein. Hier wird ein diskretes Hidden-Markov-Modell angenommen und somit ergibt sich die diskrete Ausgabeverteilung  $b_i^j(k)$  über einem für das gesamte P2DHMM festgelegten Alphabet zu:

$$b_i^j(k) = P(v_k | q_{x,y} = S_i^j) \quad (4.24)$$

Die in Kapitel 2.2.4 dargestellten Gaußschen Mischverteilungen können alternativ verwendet werden um kontinuierliche P2DHMMs zu erhalten.

Es sind bei den vorgestellten Modellgrößen die üblichen statistischen Randbedingungen einzuhalten (vgl. Gleichung 2.4). Somit ist der Aufbau des pseudo zweidimensionalen Hidden-Markov-Modells mit seinen Parametern vollständig definiert. Im eindimensionalen Fall (Kapitel 2) konnte gezeigt werden, daß durch den Viterbi-Algorithmus die Möglichkeiten gegeben sind, HMMs für die Klassifikation einzusetzen und zudem auch die HMMs an Trainingsdaten anzupassen. Letzteres wird durch den in Unterkapitel 2.2.3 vorgestellten Viterbi-Trainingsalgorithmus ermöglicht, der aufbauend auf einer Viterbi-Segmentierung die sich

daraus ergebenden Häufigkeiten, z.B. der Verwendung einer bestimmten Übergangswahrscheinlichkeit  $a_{ij}$  (vgl. Gl. 2.38), bestimmt und nachfolgend die Parameter neu schätzt. Für den Fall, daß ein verallgemeinerter Viterbi-Algorithmus existiert, ist es möglich, die P2DHMMs zu trainieren und Bildmuster mit den Modellen zu klassifizieren. Solch ein verallgemeinerter Viterbi-Algorithmus existiert und wird als zweifachverschachtelter Viterbi-Algorithmus bezeichnet, der im folgenden vorgestellt wird.

### 4.3.2 Zweifachverschachtelter Viterbi Algorithmus

Der zweifachverschachtelte Viterbi-Algorithmus wurde in [Kuo94] von Agazzi und Kuo vorgestellt. Es wird wie im eindimensionalen Fall die wahrscheinlichste Zustandszuordnung  $Q_{XY}^*$  bestimmt und basierend auf dieser Kenntnis ein Schätzwert für die Produktionswahrscheinlichkeit  $P(O_{XY}, Q_{XY}^* | \lambda) = P^*(O_{XY} | \lambda)$  ermittelt. Da es sich bei der Observation  $O_{XY}$  um eine  $X \times Y$  Matrix handelt, besteht auch die Größe  $Q^*$  im zweidimensionalen Fall aus Elementen, die in Form einer  $X \times Y$ -Matrix angeordnet sind.

Bedingt durch die Struktur der P2DHMMs, die ein hierarchisches Modell darstellen, wird die Ermittlung der wahrscheinlichsten Zustandssequenz durch zwei ineinander verschachtelte Berechnungen durchgeführt. Zunächst werden die Schätzwerte der Produktionswahrscheinlichkeiten der Bildspalten mit den den Metazuständen zugeordneten HMMs ( $\lambda_i$  für  $1 \leq i \leq N$ ) bestimmt. Diese Wahrscheinlichkeiten  $P_i(O_x) = P(O_x | \lambda_i)$  werden durch den schon bekannten eindimensionalen Viterbi-Algorithmus berechnet. Im nachfolgenden Schritt werden diese Wahrscheinlichkeiten als Ausgabewahrscheinlichkeiten der Zustände des übergeordneten Markov-Modells ( $S_i$  für  $1 \leq i \leq N$ ) verwendet. Auf der Ebene des übergeordneten Markov-Modells werden ebenfalls dem eindimensionalen Fall entsprechende Viterbi-Berechnungen durchgeführt. Der Algorithmus wird im folgenden formal dargestellt, wobei mit den Berechnungen der Zustandszuordnungen auf Bildspaltenebene begonnen wird:

- Initialisierung

$$\begin{aligned} \vartheta_{x1}^j(i) &= \pi_i^j b_i^j(O_{x1}) \\ \psi_{x1}^j(i) &= 0 \end{aligned} \quad (4.25)$$

- Rekursionsschritt

$$\begin{aligned} \vartheta_{xy}^j(i) &= \max_{i-2 \leq k \leq i} (\vartheta_{x,y-1}^j(k) \cdot a_{ki}^j) b_i^j(O_{xy}) \\ \psi_{xy}^j(i) &= \arg \max_{i-2 \leq k \leq i} (\vartheta_{x,y-1}^j(k) \cdot a_{ki}^j) \end{aligned} \quad (4.26)$$

- Terminierung

$$\begin{aligned} P_j(O_x) &= \max_{1 \leq i \leq N^j} (\vartheta_{xY}^j(i)) \\ N_j(O_x) &= \arg \max_{1 \leq i \leq N^j} (\vartheta_{xY}^j(i)) \end{aligned} \quad (4.27)$$

Diese Darstellung des Viterbi-Algorithmus unterscheidet sich von der Darstellung in Kapitel 2 in der Verwendung des diskretisierten Ortes  $y$  anstelle des Zeitschritts  $t$  (vgl. Gleichungen 2.24 bis 2.26). Ferner sind die Größen mit dem hochgestellten Index  $j$  versehen, der den Bezug zum Markov-Modell  $\lambda_j$  und somit zum  $j$ -ten Metazustand anzeigt, versehen. Die Erläuterungen zu den Größen  $\vartheta$  und  $\psi$  sind in Kapitel 2 gegeben. Der zweite Durchlauf des Viterbi Algorithmus kann wie folgt dargestellt werden:

- Initialisierung

$$\begin{aligned} D_1(j) &= \pi_j P_j(1) \\ \gamma_1(j) &= 0 \end{aligned} \quad (4.28)$$

- Rekursionsschritt

$$\begin{aligned} D_x(j) &= \max_{j-2 \leq k \leq j} (D_{x-1}(k) \cdot a_{kj}) P_j(O_x) \\ \gamma_x(j) &= \arg \max_{j-2 \leq k \leq j} (D_{x-1}(k) \cdot a_{kj}) \end{aligned} \quad (4.29)$$

- Terminierung

$$\begin{aligned} P^* &= \max_{1 \leq j \leq N} (D_X(j)) \\ q_x &= \arg \max_{1 \leq j \leq N} (D_X(j)) \end{aligned} \quad (4.30)$$

In den unterschiedlichen hierarchischen Stufen entsprechen die Größen  $D$  und  $\vartheta$  bzw.  $\gamma$  und  $\psi$  einander. Nach einem vollständigen Durchlauf des zweifachverschachtelten Viterbi Algorithmus ist der Schätzwert für die Produktionswahrscheinlichkeit durch die Gleichung 4.30 gegeben. Somit ist es möglich, zweidimensionale Muster mit den P2DHMMs zu klassifizieren. Soll auch die optimale Zustandssequenz ermittelt werden, die verwendet werden kann um ein Viterbi-Training durchzuführen, so kann diese Sequenz mit folgender Rechenvorschrift ermittelt werden:

$$\begin{aligned} q_x &= \gamma_{x+1}(q_{x+1}) \\ q_{xy} &= \psi_{x,y+1}(q_{x,y+1}) \end{aligned} \quad (4.31)$$

Neben der Möglichkeit die Parameter des P2DHMMs mit einem Viterbi-Training zu bestimmen, existiert auch ein zweifachverschachtelter Baum-Welch-Trainingsalgorithmus. Dieser ist in [MM99] beschrieben. Anstelle der zweifachverschachtelten Trainings- bzw. Klassifikationsalgorithmen können jedoch auch die Algorithmen für eindimensionale Markov-Modelle verwendet werden, falls die im folgenden Abschnitt beschriebenen Modellierungsverfahren eingesetzt werden.

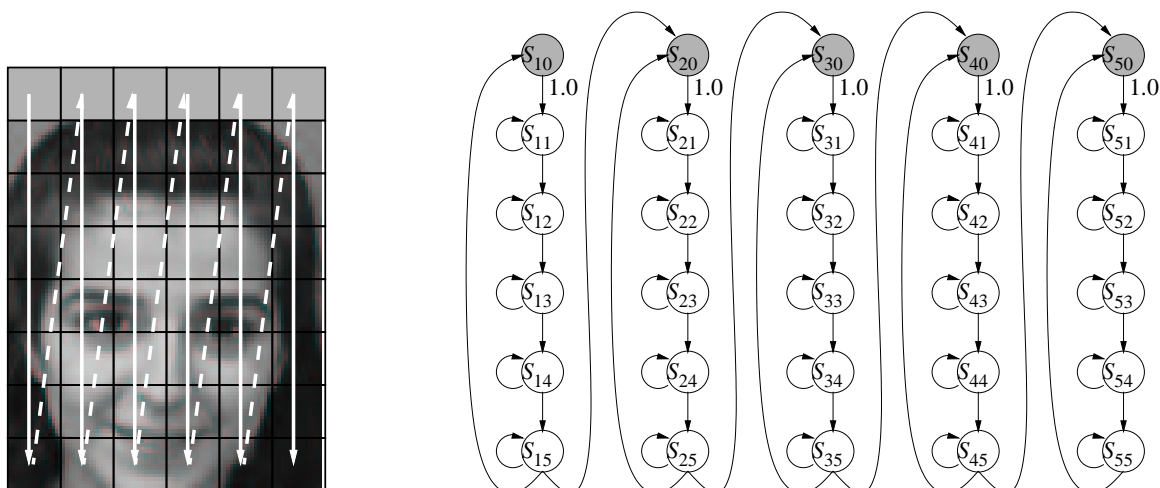


Abbildung 4.3: Eine den P2DHMMs gleichwertige Modellierung, die eindimensionale Hidden-Markov-Modelle verwendet.

### 4.3.3 Umformung in gleichwertige eindimensionale Hidden-Markov-Modelle

Samaria beschreibt in [Sam94b] eindimensionale Hidden-Markov-Modelle, die eine im Vergleich zu den P2DHMMs gleichwertige Modellierung ermöglichen. Dies wird durch das Einfügen von sog. *Markierungszuständen* in die eindimensionalen HMMs und durch das Hinzufügen von korrespondierenden *Markierungsmerkmalen* erreicht. Diese Modellierung ist in Abb. 4.3 schematisch dargestellt. Auf der linken Seite in Abb. 4.3 ist das zu modellierende zweidimensionale Muster gezeigt, bei dem es sich um ein Bild aus einer Gesichtsdatenbank handelt. Um eine Modellierung mit eindimensionalen HMMs zu ermöglichen, müssen zunächst die Merkmale des Bildes in eine Merkmalsequenz überführt werden. Dies geschieht durch die Abtastung des Bildes auf folgende Weise: Merkmale werden, wie in Abb. 4.3 durch weiße Pfeile visualisiert, mit einem Abtastfenster von oben nach unten und links nach rechts entnommen. Der Anfang einer jeden Bildspalte wird durch ein sog. Markierungsmerkmal angezeigt. Diese hinzugefügten Merkmale sind in der Abb. 4.3 grau dargestellt.

Die Topologie des eindimensionalen Modells ist auf der rechten Seite der Abb. 4.3 illustriert. Die grau-schattierten Zustände sind die sog. Markierungszustände, die bei der Emission von Markierungsmerkmalen hohe Wahrscheinlichkeiten ausgeben. Da dem Markierungsmerkmal unmittelbar kein weiteres folgt, sind die Zustandsübergänge von einem Markierungszustand auf sich selbst auf Null gesetzt (z.B.  $a_{10,10} = P(q_k = S_{10} | q_{k-1} = S_{10}) = 0$ ,  $a_{20,20} = 0, \dots$ ). Die Übergangswahrscheinlichkeiten, die zu den Markierungszuständen führen, also z.B.  $a_{15,10}$  und  $a_{15,20}$  modellieren die statistischen Abhängigkeiten aufeinanderfolgender Bildspalten. Sie entsprechen somit den Übergangswahrscheinlichkeiten des übergeordneten Markov-Modells bei den pseudo zweidimensionalen Modellen (vgl. Gleichung 4.20).

Bei der Verwendung dieser Modellierungsmethode muß darauf geachtet werden, daß die Markierungsmerkmale ausschließlich den Markierungszuständen zugeordnet werden. Dies kann bei der Verwendung von diskreten Ausgabeverteilungen durch das Reservieren eines Ausgabesymbols  $v_m$  geschehen. Die Wahrscheinlichkeitsverteilung der Markierungszustände wird für alle Ausgabesymbole  $v_i$  mit  $i \neq m$  zu Null gesetzt und für das Symbol  $v_m$  zu Eins. Für alle anderen Zustände ist die Wahrscheinlichkeit für die Ausgabe des Markierungssymbols auf Null zu setzen. Durch diese Maßnahmen ist gesichert, daß die vorgestellte eindimensionale Modelltopologie die Abfolge von Bildspalten richtig modelliert und somit mit den P2DHMMs vergleichbare Eigenschaften aufweist. Sollen kontinuierliche Ausgabeverteilungen verwendet werden, so erweist sich die Tatsache als problematisch, daß die verwendeten Gaußfunktionen (siehe Unterkapitel 2.2.4) theoretisch auch im beliebig großen Abstand zum Mittelwert noch von Null verschiedene Werte aufweisen. Somit wäre eine fälschliche Zuordnung von Markierungsmerkmalen und gewöhnlichen Zuständen sowie die Zuordnung von gewöhnlichen Merkmalen und Markierungszuständen möglich. Dies kann jedoch durch programmtechnische Maßnahmen verhindert werden, indem z.B. auf ein festgelegtes Markierungsmerkmal abgefragt wird und die Wahrscheinlichkeit Null ausgegeben wird. Dabei darf das gewählte Markierungsmerkmal nicht im Wertebereich der gewöhnlichen Merkmale vorhanden sein, was bei natürlichen Bildern stets möglich ist. Falls die Merkmale z.B. Grauwerte, die mit 8bit kodiert wurden, repräsentieren, so ist der Wertebereich für die Merkmale  $0, \dots, 255$  und eine geeignete Wahl für das Markierungsmerkmal wäre somit ein Wert außerhalb des 8bit-Wertebereichs.

Diese Modellierungsmethode erlaubt die direkte Verwendung der in Kapitel 2 vorgestellten Trainings- und Klassifikationsalgorithmen für eindimensionalen Hidden-Markov-Modelle. Aus diesem Grund kann somit Software, die für den Einsatz in der automatischen Spracherkennung entwickelt wurde, wie etwa das Hidden-Markov-Toolkit (HTK), unter den erwähnten Voraussetzungen verwendet werden. Daher wurde diese Modellierungstechnik, bei den in dem folgenden Kapitel 5 vorgestellten Experimenten, eingesetzt.

## 4.4 Kapitelzusammenfassung

In diesem Kapitel wurde in die Theorie der zweidimensionalen statistischen Modellierung eingeführt. Zunächst wurden die Markov-Random-Fields vorgestellt, die ein nichtkausales Modell darstellen, welches in der statistischen Physik entwickelt wurde. Aus diesen gehen durch die Einführung einer kausalen Abhängigkeit von benachbarten Bildelementen und eines zweiten statistischen Prozesses die zweidimensionalen Hidden-Markov-Modelle hervor. Dieser zweite statistische Prozeß ist wie im eindimensionalen Fall die Ausgabe von Merkmalen durch Ausgabeverteilungen bzw. -dichten. Sowohl die Markov-Random-Fields, als auch die zweidimensionalen HMMs sind für die Mustererkennung wenig geeignet, da keine effizienten Algorithmen existieren, um ein Training der Modelle anhand von Musterbeispielen

len zu ermöglichen. Durch weitere Vereinfachungen wurden die pseudo-zweidimensionalen Hidden-Markov-Modelle abgeleitet, für die ein erweiterter Viterbi-Algorithmus existiert. Somit existiert ein effizienter Algorithmus für diese Modelle, der eine Parameterbestimmung durch ein Training mit Beispielen ermöglicht. Zusätzlich kann durch diesen Viterbi-Algorithmus eine Segmentierung eines Musters in Kombination mit einer Klassifikation erfolgen. Dies wird im folgenden Kapitel verwendet, um zweidimensionale Muster in einer komplexen Szene aufzufinden und zu erkennen.