

Kapitel 7

Zusammenfassung

Ziel dieser Arbeit war die Nutzung der integrierten Segmentierungs- und Klassifikationseigenschaften der Hidden-Markov-Modelle für die Erkennung von Mustern in Bildern und Bildsequenzen. Diese besondere Eigenschaft ist durch den Viterbi-Algorithmus gegeben, der eine Merkmal-Zustandszuordnung ausgibt, die als Segmentierung interpretiert werden kann. Ferner liefert der Viterbi-Algorithmus einen Schätzwert dafür, daß ein gegebenes Muster von einem HMM produziert wurde. Unter Verwendung dieses Schätzwertes erfolgt die Musterklassifikation. Obwohl Hidden-Markov-Modelle schon seit den 80er Jahren erfolgreich bei der Erkennung von zeitlich veränderlichen Mustern, wie beispielsweise Sprache oder Online-Handschrift, eingesetzt werden und in diesen Anwendungsszenarien somit schon lange Gebrauch gemacht wird von der Fähigkeit in einem Schritt segmentieren und klassifizieren zu können, so ist die Nutzung dieser Eigenschaften auf dem Gebiet der Bild- und Bildsequenzerkennung als neu anzusehen. Dabei ist die Verwendung der populären Hidden-Markov-Modelle bei der Erkennung von Bildern, also zweidimensionaler Muster, keinesfalls trivial, da eine Erweiterung der eindimensionalen Struktur dieser Modelle erforderlich ist. Gleiches gilt für Bildsequenzen, die dreidimensionale Muster darstellen.

In dieser Arbeit konnte zunächst ohne Verwendung der höherdimensionalen Hidden-Markov-Modelle die erfolgreiche Anwendung der integrierten Segmentierungs- und Klassifikationseigenschaften auf dem Gebiet der automatischen Bildererkennung demonstriert werden. Dazu wurden neuartige eindimensionale HMM-Topologien vorgestellt, die zusammen mit einer polaren Abtastung eine translations-, skalierungs- und rotationsunabhängige Modellierung von Objektformen bzw. handskizzierten Piktogrammen ermöglichen. Die integrierten Segmentierungs- und Klassifikationseigenschaften der HMMs wurden dazu genutzt, die Orientierung der gedrehten Objekte herauszufinden und das Objekt zu erkennen. In den Experimenten wurden Erkennungsgenauigkeiten von bis zu 99.5% mit Piktogramm-Datenbasen erreicht, die aus 20 Klassen bestehen. Die Erkennungsergebnisse lagen über denen, die mit konventionellen Erkennungsmethoden, nämlich Momenten in Kombination mit künstlichen neuronalen Netzen erzielt wurden. Die vorgestellten Methoden konnten erfolgreich auf die Erkennung natürlicher Bilder erweitert werden. Es wurden Ergebnisse präsen-

tiert, die mit einem Bilddatenbanksystem, das intuitiv über Skizzen des Benutzers abgefragt werden kann und das die neuartigen eindimensionalen Modelltopologien verwendet, erzielt wurden.

Die Erkenntnisse, die durch die Experimente mit den eindimensionalen Hidden-Markov-Modellen erzielt wurden, konnten genutzt werden, um die kombinierten Segmentierungs- und Klassifikationseigenschaften auch im zweidimensionalen Fall nutzen zu können. Die vorliegende Arbeit präsentierte einen Ansatz, der es ermöglicht, zweidimensionale Muster in komplexen Umgebungen aufzufinden und zu klassifizieren. Dabei wurden pseudo zweidimensionale Hidden-Markov-Modelle in Kombination mit an den Bildkontext angepaßten Umgebungszuständen verwendet. Pseudo zweidimensionale HMMs stellen eine hierarchische Erweiterung der eindimensionalen Modelle dar und sind geeignet, um zweidimensionale Muster zu modellieren. Es existieren effiziente Algorithmen für das Training und die Klassifikation mit diesen Modellen und somit bietet sich die Anwendung dieser Modelle an. Die Anpassung der Parameter der Umgebungszustände kann auf verschiedene Weisen erfolgen, je nachdem, ob Vorwissen über die zu analysierende Szene vorliegt oder nicht. Für den letztgenannten Fall wurde ein Verfahren entwickelt, bei dem die Parameter der Umgebungszustände auf allen Merkmalen des zu analysierenden Bildes bestimmt wurden. Nach der Ausführung des Viterbi-Algorithmus liegt eine Zuordnung der Merkmale zu den Umgebungszuständen und den Zuständen des gesuchten Musters vor, die als Segmentierung des Bildes in Muster und Umgebung interpretiert werden kann. Zusätzlich liefert der Viterbi-Algorithmus einen Schätzwert für die Produktionswahrscheinlichkeit, der zur Klassifikation genutzt werden kann. Das neuartige Verfahren wurde zunächst auf die Erkennung von handskizzierten Piktogrammen in komplexen Szenen angewendet. Dabei wurden Erkennungsgenauigkeiten von 90% auf einer Piktogrammdatenbasis erreicht, die aus 20 Klassen bestand. Weitere Experimente wurden beschrieben, die die Eignung des Ansatzes für das Auffinden von benutzerdefinierten Formen in technischen Zeichnungen belegen. Somit ist es z.B. möglich, eine durch eine Skizze spezifizierte Schraube in komplexen technischen Zeichnungen aufzufinden. Schließlich wurde der P2DHMM-Ansatz für das Personen-Tracking in Bildfolgen eingesetzt. Dabei konnte gezeigt werden, daß Muster bzw. Personen auch in Grauwert- und Farbbildern mit dem vorgestellten Ansatz gefunden werden können. Es zeigte sich ebenfalls, daß der Ansatz gut kombinierbar ist mit einem Kalman-Filter, das die Dynamik der Bewegung einer Person modelliert. Obwohl diese zweistufige Methode, die P2DHMMs für das Auffinden der Person und das Kalman-Filter für die Bewegungsmodellierung verwendet, gute Ergebnisse zeigte, so ist eine dreidimensionale Modellierung sehr viel geeigneter für die Erkennung von Bildfolgen.

Eine solche dreidimensionale Modellierung wurde im Rahmen dieser Arbeit entwickelt. Die neuartigen sog. pseudo dreidimensionalen Hidden-Markov-Modelle ermöglichen es, Merkmale, die auf Einzelbildern berechnet werden, gemeinsam mit Merkmalen zu modellieren, die aus der temporalen Abfolge der Bilder bestimmt werden. Somit können dynamische

und statische Bewegungsmuster gemeinsam mit einem Modell erkannt werden. Die Evaluierung des P3DHMM-Ansatzes erfolgte anhand einer selbsterstellten Gestendatenbank, die aus 12 Gesten besteht, die der Steuerung von Baukränen dienen. Es wurden experimentelle Ergebnisse mit diesen Modellen erzielt, die denen, die mit einem alternativen Ansatz, der eindimensionale HMMs in Kombination mit geometrischen Momenten verwendet, überlegen sind. Als Ausblick wurde auf die integrierte Segmentierung und Klassifikation von Bildsequenzen mit P3DHMMs und Umgebungsmodell hingewiesen, die z.B. eine positionsunabhängige Gestenerkennung ermöglicht, oder für die Abfrage von Filmdatenbanken mit Beispielgesten genutzt werden kann.

Die Methode der gemeinsamen Segmentierung und Klassifikation, die bei der Verwendung von Hidden-Markov-Modellen zur Verfügung steht, konnte erfolgreich für die Erkennung von Mustern in Bildern und Bildsequenzen genutzt werden. Es wurde eine Vielzahl von Experimenten mit neuartigen Modellierungsmethoden vorgestellt, die die Eignung dieser Methoden für die Mensch-Maschine-Kommunikation und verschiedenen multimedialen Anwendungen belegen. Die Arbeit zeigt somit das große Anwendungspotential der Hidden-Markov-Modelle im Bereich der Bild- und Bildsequenzerkennung.