

Feinevaluierung des Probabilistischen Arm Modells zur Analyse von Klavierspielbewegungen

Aristotelis Hadjakos

telis@tk.informatik.tu-darmstadt.de

1 Einleitung

Beim Klavierspiel koordiniert der Spieler zahlreiche Bewegungen der Finger, der Arme, des Oberkörpers, etc. Besseres Bewusstsein über die eingesetzten Spielbewegungen und bessere Kenntnis der Bewegungsmöglichkeiten können sich positiv auf Spieltechnik, Musikalität und Gedächtnisleistung auswirken. Mit Hilfe von Sensortechnologie ist es möglich, dem Benutzer Feedback über die beim Spiel genutzten Spielbewegungen zu geben und so die Bewegungswahrnehmung zu schärfen. Für den typischen Benutzer sind jedoch die Sensordaten nur schwer mit den ausgeführten Spielbewegungen in Verbindung zu setzen. Bevor die Sensordaten dem Benutzer sonifiziert oder angezeigt werden, ist es daher vorteilhaft, eine automatische Analyse vorzuschalten, die aus dem Sensordatenstrom eine für den Benutzer leichter verständliche Bewegungsbeschreibung ermittelt.

Eine grundlegende Aufgabe bei der automatischen Analyse ist es, zu bestimmen welche Bewegungen des Armes beim Anschlag einer Taste involviert waren. Diese scheinbar einfache Frage wird durch das Auftreten von *sekundären Bewegungen*, die sich aus dem mechanischen Wechselspiel von Spielapparat und Klaviermechanik ergeben, erschwert. Das *Probabilistische Arm Modell (PAM)* [HAM09] ermöglicht es, die bei einem Anschlag auftretende sekundäre Bewegung abzuschätzen. Dazu wird die sekundäre Bewegung als normalverteilt modelliert. Mittels der Maximum-Likelihood-Methode wird aus einer umfassenden Datensammlung von Anschlägen eine Funktion f gelernt, die den Mittelwert und die Standardabweichung der in einem Gelenk auftretenden sekundären Bewegung abschätzt. Die Analyse bezieht sich hierbei auf die Bewegungen während eines Zeitintervalls fester Dauer d , das mit dem Erklängen des Tones endet. Mit Hilfe von inertialen Sensoren werden die Drehraten in allen sieben hauptsächlichen Freiheitsgraden des Arms bestimmt. Der Mittelwert μ_i und die Standardabweichung σ_i der sekundären Bewegung im i -ten Gelenk wird nun als Funktion der Bewegungen ($F_j, j \neq i$) in den anderen Gelenken und der Tastenanschlagsgeschwindigkeit (F_v) geschätzt: $(\mu_i, \sigma_i) = f(F_1, F_2, \dots, F_{i-1}, F_{i+1}, \dots, F_7)$. Eine Bewegung im i -ten Gelenk wird erkannt wenn die gemessene Bewegung den Mittelwert um einen Faktor c mal der Standardabweichung übersteigt, d.h. wenn $|F_i - \mu_i| > c \cdot \sigma_i$.

In diesem Aufsatz werden für den praktischen Einsatz relevante Fragen bezüglich PAM behandelt, die zuvor noch nicht systematisch betrachtet wurden, und zwar:

1. Welcher Wert soll für den Faktor c gewählt werden?
2. Wie wirkt sich die Veränderung der Dauer des Analyseintervalls bei der Erkennungsgenauigkeit aus? Kürzere Analyseintervalle haben den Vorteil, dass Überlappungen mehrerer Bewegungen während des Analyseintervalls minimiert werden. Allerdings ist dieser Vorteil gegenüber der bei isolierten Bewegungen bestimmten Erkennungsgenauigkeit abzuwägen. Von dieser ist nämlich zu erwarten, dass sie bei kürzeren Analyseintervallen sinkt.
3. Wieviel Daten sind nötig, um die Funktion f zu lernen?

2 Experimente

Um die im vorigen Abschnitt aufgestellten Fragen zu beantworten, wurden die Dauer des Analyseintervalls d , der Faktor c und die Menge der Daten systematisch verändert. Die Ergebnisse werden im Folgenden am Beispiel der Erkennung von Anschlägen mit Bewegung aus dem Handgelenk vorgestellt.

In Abb. 1 ist der Einfluss des Faktors c auf die Erkennung eines Anschlags aus dem Handgelenk dargestellt. Dazu wurde die Gesamtdatenmenge im Verhältnis 2:1 in Trainings- und Testdaten aufgeteilt. Abb. 1 wurde basierend auf den Testdaten erstellt. Man erkennt, dass sich bei hohen Werten von c die fälschliche Erkennung von Anschlägen, die ohne Beteiligung des Handgelenks (Anschläge aus dem Finger, aus dem Ellenbogengelenk, ...) ausgeführt wurden, sinkt. Allerdings vermindert sich bei hohen Werten von c auch die Erkennung von tatsächlich aus dem Handgelenk ausgeführten Anschlägen. Mit Hilfe einer Kostenfunktion kann der optimale Wert für den Faktor c bestimmt werden. Mit einer Kostenfunktion, die einen Fehler 1. Art fünfmal schwerer gewichtet als einen Fehler 2. Art, wurde der optimale Faktor basierend auf den Trainingsdaten ermittelt. Er beträgt im betrachteten Beispiel $c = 2,35$.

Um den Einfluss des Analyseintervalls d zu bestimmen, wurde die Funktion f mit verschiedenen Analyseintervallen im Bereich von 0,01 bis 0,2 Sekunden trainiert. Daraufhin wurden die Kosten wie oben ausgeführt bestimmt. Eine Gegenüberstellung von Analyseintervall und Kosten für Anschläge aus dem Handgelenk befindet sich in Abb. 2. Die zugehörigen Faktoren c , die auf Grundlage der Trainingsdaten bestimmt wurden, sind auf der Sekundärachse aufgetragen. Ab einem Analyseintervall der Länge 0,1 s verringern sich die Kosten nur marginal. Für die Erkennung von Anschlägen aus Handgelenk, Ellenbogengelenk und Schulter wurde daher dieses Intervall gewählt. Für Anschläge mit Rotation des Unterarms wurde ein besonders kurzes Analyseintervall von 0,08 Sekunden gewählt, um zu verhindern, dass gegenläufige Rotationen, die beim Spielen von Klavierliteratur auftreten können, sich während des Analyseintervalls gegenseitig überdecken.

Um den Einfluss der Menge der Trainingsdaten auf die Güte der gelernten Funktion f zu bestimmen, wurde der Anteil der Trainingsdaten an der Gesamtdatenmenge von ca. 18.000 Anschlägen zwischen 1 und 99% variiert. Basierend auf den Trainingsdaten wurde die Funktion f trainiert und der optimale Faktor c ermittelt. Die Kostenfunktion wurde

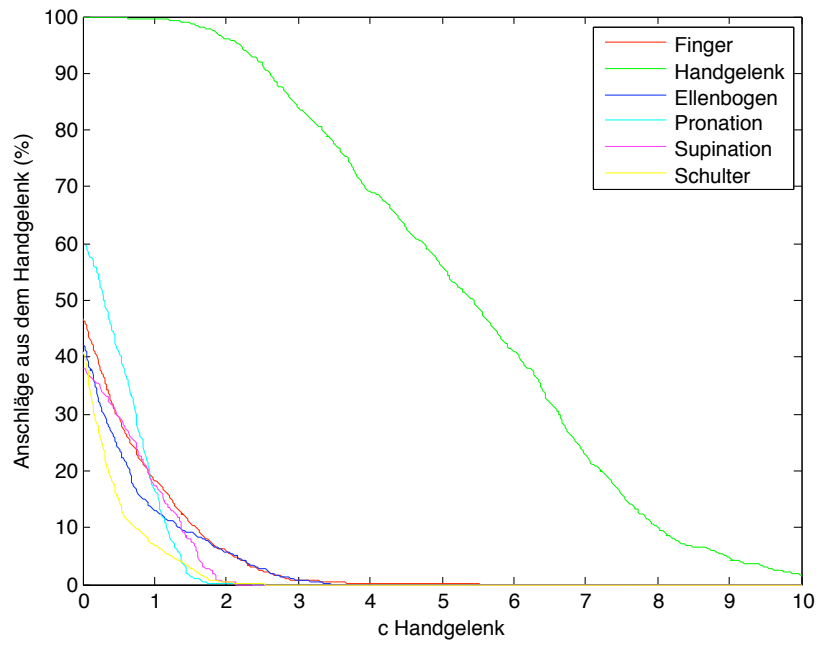


Abbildung 1: Erkennung von Anschlägen aus dem Handgelenk

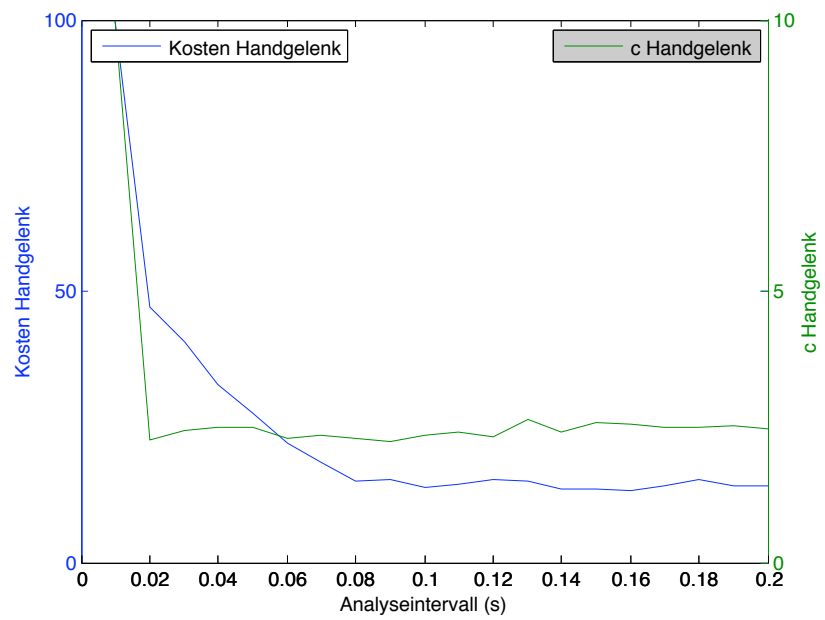


Abbildung 2: Analyseintervall und Kosten (Primärachse) bzw. Faktor c (Sekundärachse).

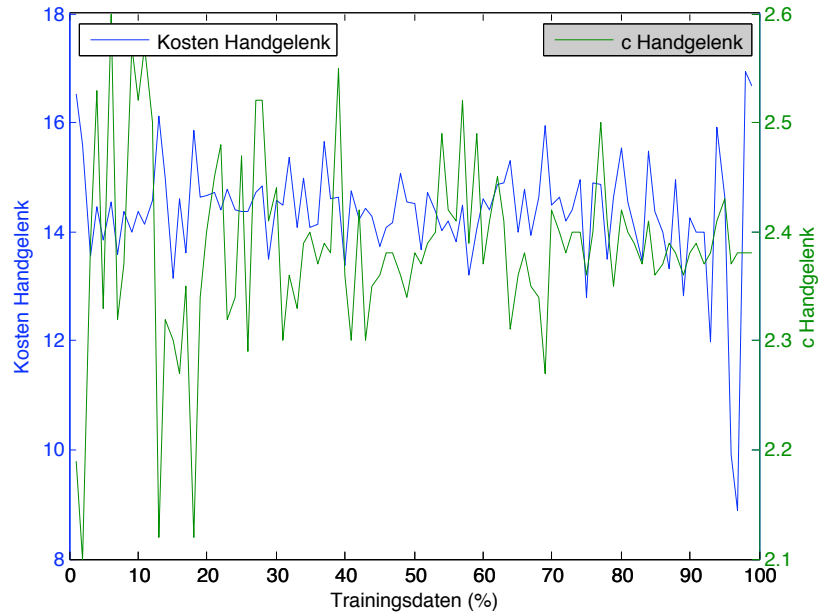


Abbildung 3: Trainingsdaten und Kosten (Primärachse) bzw. Faktor c (Sekundärachse).

dann anhand der Testdaten ausgewertet. In Abb. 3 befindet sich eine Gegenüberstellung von Menge der Trainingsdaten und Kosten. Die zugehörigen Faktoren c sind auf der Sekundärachse aufgetragen. Selbst bei der Nutzung von nur 1% der Gesamtdatenmenge ist kein deutlich negativer Effekt auf die ermittelten Kosten feststellbar. Es reicht also aus, PAM mit ca. 180 Anschlägen zu trainieren.

Literatur

[HAM09] Aristotelis Hadjakos, Erwin Aitenbichler und Max Mühlhäuser. Probabilistic Model of Pianists' Arm Touch Movements. In *NIME 2009 Proceedings*, 2009.