

Universität Hamburg  
**Fachbereich Informatik**  
Vogt-Kölln-Str. 30  
D-22527 Hamburg

Fachbereichsmitteilung 349

**Detektion und Klassifizierung von  
Ligaturen in mittelalterlichen  
Musikhandschriften**

Arved Solth

FBI-HH-M-349

# Detektion und Klassifizierung von Ligaturen in mittelalterlichen Musikhandschriften

Arved Solth  
Universität Hamburg  
Fachbereich Informatik

Januar 2015

## Abstract

Manuscript research provides many interesting opportunities for applying computer-aided image analysis methods to large datasets. Detecting and counting special visual patterns that are of particular interest in extensive manuscripts is one example where image analysis can provide pattern matching functionalities to support the work of manuscript researchers. Together with a team of musicologists of the “Centre for the Study of Manuscript Cultures” we have developed a system that detects individual notes in medieval music manuscripts and classifies them using a Support Vector Machine. In a subsequent step the system detects ligatures as aggregates of individual notes and classifies them corresponding to certain geometric criteria.

## Zusammenfassung

Computergestützte Verfahren der Bildverarbeitung bieten vielfältige Möglichkeiten, um Paläografen und Manuskriptologen bei ihrer Arbeit zu unterstützen. Ein Beispiel hierfür ist das Detektieren und statistische Auswerten von Vorkommen relevanter visueller Muster in großen Datenbeständen digitalisierter Manuskriptseiten. In Kooperation mit Musikwissenschaftlern des “Centre for the Study of Manuscript Cultures” wurde ein System entwickelt, das es erlaubt, Ligaturen in Musikmanuskripten als Aggregate von Noten zu finden und zu zählen. Mit Hilfe einer Support Vector Machine werden Notenformen gelernt und anschließend zu Ligaturen verschiedener Typen verbunden.

## 1 Einleitung

Das “Center for the Study of Manuscript Cultures” (CSMC) beschäftigt sich mit der Erforschung von Manuskripten aus vielen verschiedenen Ländern und Kulturkreisen. Hierzu zählt unter anderem auch die Analyse mittelalterlicher Musikhandschriften. Eine der Herausforderungen, die durch Bildverarbeitungsmethoden unterstützt werden können, ist die Frage nach dem Vorkommen und der Häufigkeit verschiedener für den Manuskriptforscher interessante Zeichen bzw. Schriftmuster in großen Datenmengen. Dadurch können beispielsweise Rückschlüsse auf die Entstehung der Manuskripte gewonnen werden. Im Falle der in



Abbildung 1: Bildausschnitt aus mittelalterlicher Musikhandschrift

dieser Arbeit beschriebenen Kooperation mit Musikologen des CSMC handelte es sich konkret um die Frage nach dem Vorkommen verschiedener Ligaturtypen in Manuskripten unterschiedlichen Ursprungs. Zu diesem Zweck wurde ein System entwickelt, das nach der Vorverarbeitung der digitalisierten Manuskriptseiten zunächst einzelne Noten detektiert und klassifiziert und diese anschließend unter Berücksichtigung spezieller geometrischer Kriterien zu Ligaturen zusammenfasst. Als Ergebnis wird eine Statistik erstellt, die es dem Musikwissenschaftler erlaubt zu untersuchen, ob - und, wenn ja welcher - Zusammenhang zwischen dem Ursprung der Manuskripte und der Häufigkeiten der in ihnen verwendeten Ligaturtypen besteht. In Abschnitt 2 wird eine Auswahl von Arbeiten zum Thema der computergestützten Verarbeitung von Musikmanuskripten beschrieben. Abschnitt 3 erläutert unseren mehrstufigen Ansatz zur Detektion und Klassifizierung von Ligaturen. Ergebnisse werden im abschließenden Abschnitt 4 beschrieben.

## 2 Vergleichbare Arbeiten

Verschiedene Veröffentlichungen aus den vergangenen Jahren befassen sich mit unterschiedlichen Problemen im Zusammenhang mit der computergestützten Analyse von mittelalterlichen Musikhandschriften.

In [3] wird ein System zur automatischen Erkennung von Quadratnoten in historischen Musikhandschriften beschrieben. Das System verwendet zur Symbolklassifizierung wie unser Ansatz eine Support Vector Machine und erreicht eine hohe Präzision auf den analysierten Datensätzen. Im Gegensatz zu unserem Ansatz zielt das System allerdings nur auf die Detektion von einfachen Symbolen wie Noten ab, das Finden von komplexeren Strukturen wie aus einzelnen Noten zusammengesetzten Ligaturen wird nicht berücksichtigt.

Im Zusammenhang mit Musikhandschriften wurden auch andere Fragestellungen bearbeitet, z.B. das Problem der Schreiberidentifikation bzw. -Verifikation. [1] beschreibt einen zeichenabhängigen Ansatz, der die Verfahren der “Blurred Shape Model Descriptors” und des “Dynamic Time Warping” kombiniert, um gesuchte Zeichen in einer Handschrift zu identifizieren und anschließend zu analysieren.

Manuskript	Seiten	Auflösung Einzelseiten (Pixel)
M1	405	2000 x 3035
M2	534	2000 x 3045
M3	885	625 x 910

Tabelle 1: Musikhandschriften

### 3 Ligaturdetektion in Musikhandschriften

#### 3.1 Vorverarbeitung

Das Ziel der vorliegenden Arbeit ist es, in drei mittelalterlichen Musikhandschriften (M1, M2, M3), die von Musikwissenschaftlern am CSMC in Hamburg untersucht werden, verschiedenartige Ligaturen zu detektieren und zu zählen (siehe Tabelle 1).

Das von uns entwickelte System arbeitet in mehreren Schritten. Die digitalisierten Manuskriptseiten werden zunächst durch eine Farbraumreduktion in Grauwertbilder umgewandelt (siehe Abbildung 2). Die in dieser Arbeit untersuchten Manuskripte enthalten Noten in schwarzer Tinte und Notenlinien, die in roter Tinte erzeugt wurden. Aus diesem Grund ist es in unserem Fall möglich, Noten und Text vom Rest der digitalisierten Manuskriptseiten zu trennen, ohne eine spezielle Notenlinienentfernung anwenden zu müssen. Dazu wird bei der Umwandlung der Farbbilder in Graubilder der Rotkanal verwendet, da in diesem Schrift bzw. Noten die einzigen Elemente sind, die niedrige Werte enthalten und sich dadurch in diesem Farbkanal deutlich von Notenlinien und Papierhintergrund abheben.

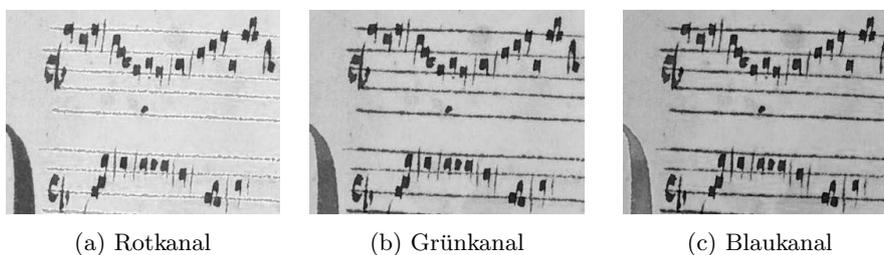
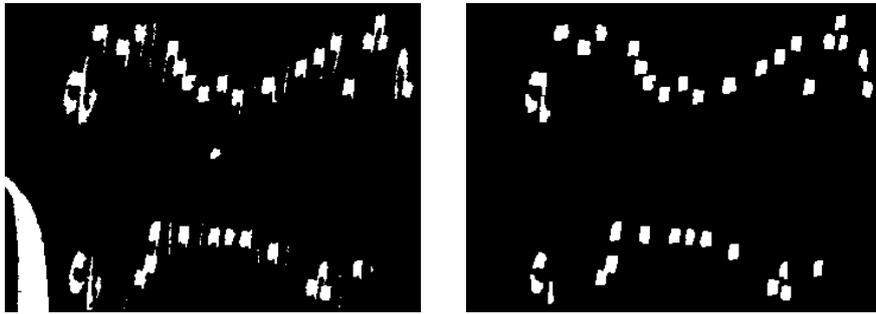


Abbildung 2: Farbkanäle (Helligkeit = Amplitude des Farbkanals)

Anschließend werden die Grauwertbilder mit Hilfe eines einfachen Schwellenwertes binarisiert. Bei Bildern mit starken Hintergrundstörungen, wie zum Beispiel ungleichmäßiger Ausleuchtung der Manuskriptseiten, kann die Verwendung eines adaptiven Schwellenwertverfahrens notwendig sein. Das auf diese Weise erzeugte S/W-Bild wird anschließend von allen Regionen gesäubert, deren Flächeninhalt entweder zu groß oder zu klein ist, um einen einzelnen Notenkopf zu repräsentieren. Hierdurch werden Störungen wie Schmutz oder Staub sowie große Regionen, die beispielsweise zu Initialen gehören können, entfernt. Abbildung 3 zeigt ein binarisiertes Bild vor und nach der Anwendung dieses Flächenfilters.



(a) Vor Anwendung des Flächenfilters (b) Nach Anwendung des Flächenfilters

Abbildung 3: Binarisierter Ausschnitt einer Musikhandschrift

### 3.2 Notendetektion

Die verbliebenen Flächen werden mit Hilfe eines von einer Support Vector Maschine (SVM) erzeugten Klassifikators in eine der drei Klassen “Quadratnotation”, “Rhombennotation” und “Blob” klassifiziert, wobei “Blob” eine Fläche mit passendem Flächeninhalt aber unpassender Form bezeichnet. Die Klassen “Quadratnotation” sowie “Rhombennotation” beschreiben Notenköpfe mit quadratischer bzw. Rauten-Form. Abbildung 4 zeigt einige Beispiele dieser Notenklassen.

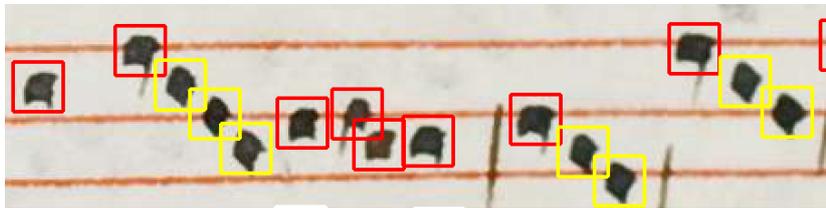


Abbildung 4: Quadratnotation (rot) und Rhombennotation (gelb)

Der Klassifikator wird erzeugt, indem die SVM mit Beispielen für die einzelnen Klassen trainiert wird. Zu diesem Zweck wurden Trainings- und Testmengen aus den einzelnen Manuskripten erstellt, die jeweils durchschnittlich 200 Objekte aus jeder Klasse und jedem Manuskript enthalten (siehe Tabelle 2). Die Stichproben, mit denen die SVM trainiert und getestet wird, bestehen aus um die Mittelpunkte der einzelnen Noten zentrierten Bildausschnitten einheitlicher Größe. Die Sequenz aller schwarzen und weißen Pixel eines Bildausschnittes bilden einen Merkmalsvektor, mit dem der Bildausschnitt in der SVM repräsentiert wird. Um die unterschiedlichen Auflösungen zu kompensieren, mit denen die verschiedenen Manuskripte digitalisiert wurden, werden alle Stichproben in den Test- und Trainingsmengen vor der Verwendung mit der SVM auf eine einheitliche Größe skaliert.

Das beste Klassifizierungsergebnis erreicht die SVM, wenn die Noten für die Test- und Trainingsmengen eines SVM-Durchlaufs aus dem gleichen Manuskript stammen. Wird die SVM mit Noten verschiedener Manuskripte trainiert, verschlechtert sich das Klassifizierungsergebnis von Testsets einzelner Manuskripte.

Set	Rhomben	Quadrate	Blobs
M1-Training	333	390	218
M1-Test	319	260	242
M2-Training	128	181	139
M2-Test	207	252	188
M3-Training	222	218	178
M3-Test	287	445	244

Tabelle 2: Trainings- und Testmengen

Tabelle 3 verdeutlicht diese Ergebnisse.

Manuskript	Korrekt klassifiziert	Korrekt klassifiziert (%)
M1	719 / 821	87,57 %
M2	624 / 647	96,44 %
M3	623 / 967	64,42 %

Tabelle 3: Ergebnisse der Notenklassifikation

### 3.3 Ligaturdetektion

Alle im vorherigen Schritt detektierten Noten werden als Bestandteil einer potentiellen Ligatur in Betracht gezogen. Anschließend werden die nächsten Nachbarn dieser Noten betrachtet, die innerhalb eines zugelassen Maximalabstandes liegen. Mehrere sequentiell hintereinander liegende Noten werden zu einer Ligatur zusammengefasst, wenn sie folgende Kriterien erfüllen:

1. Der Abstand zwischen jeweils zwei aufeinander folgenden Noten  $N_i$  und  $N_{i+1}$  liegt unter einem Schwellenwert  $d_T$ :  $d(N_i, N_{i+1}) < d_T$
2. Die Orientierung der Verbindung zwischen den beiden Noten  $N_i$  und  $N_{i+1}$  liegt zwischen zwei Schwellenwerten  $\alpha_L$  und  $\alpha_H$ :  $\alpha_L < \alpha(N_i, N_{i+1}) < \alpha_H$

Durch die Definition einer Ligatur als Gruppe mehrerer benachbarter Einzelnoten, zwischen denen bestimmte Relationen bestehen, ist das System flexibel anpassbar, um beliebige Konfigurationen aus Einzelnoten erkennen zu können. Neben der Anzahl der Notenköpfe pro Ligatur können so die Abstände und Winkel zwischen den einzelnen Noten variabel eingestellt und auf diese Weise auch an Eigenschaften anderer Handschriften angepasst werden.

Die verschiedenen Ligaturtypen werden auf der Basis der in ihnen enthaltenen Noten sowie Relationen zwischen den Noten klassifiziert (siehe Tabelle 4). Ligaturtyp 1 besteht aus einer Note in Quadratnotation gefolgt von mehreren Noten in Rhombennotation. Ligaturtyp 2 enthält nur Noten in Quadratnotation. Als weiteres Merkmal sind alle Noten in Ligaturtyp 2 über Verbindungslinien miteinander verbunden. Da diese in der Praxis durch beschädigtes Ausgangsmaterial oder suboptimale Digitalisierung jedoch häufig nicht sichtbar sind, kann lediglich die Anwesenheit von Notenhälsen als Indiz für Ligaturtyp 2 gewertet werden. Die Abwesenheit von Notenhälsen kann jedoch nicht als Hinweis auf Ligaturtyp 1 gewertet werden. Abbildung 5 zeigt die unterschiedlichen Ligaturtypen.



(a) Ligaturtyp 1



(b) Ligaturtyp 2

Abbildung 5: Ligaturtypen

Da durch die oben genannten Qualitätsschwankungen eine Verbundenheit zwischen Noten und Notenhälsen in einem in Regionen aufgeteilten Bild nicht garantiert ist, wird die ungefähre Position von Strichen als schwaches Kriterium akzeptiert.

Die am Fachbereich Informatik der Universität Hamburg entwickelte Bildverarbeitungsbibliothek VIGRA liefert zahlreiche Verfahren zum Durchführen der einzelnen Arbeitsschritte (siehe [2]). Für die Auswertung der drei getesteten Manuskripte wurde ein Prototyp inklusiver graphischer Benutzeroberfläche entwickelt, der es erlaubt, den Trainings- und Testsets der einzelnen Manuskripte manuell neue Beispiele hinzuzufügen und die Kriterien der Ligaturklassifikation anhand von Parametern zu konfigurieren.

## 4 Ergebnisse

Nach Auswertung der drei Musikhandschriften M1, M2 und M3 mit teilweise unterschiedlichem Ursprung konnten auf diese Weise signifikante Unterschiede in den Verteilungen der beiden gesuchten Ligaturklassen festgestellt werden. Tabelle 5 listet die Anzahl der in den Manuskripten gefundenen Ligaturen beider Typen auf. Das Verhältnis zwischen Ligaturen vom Typ 1 und Ligaturen vom Typ 2 beträgt in Manuskript M3 ungefähr 3:1, in den Manuskripten M1 und M2 liegt es bei 2:1. Dieser erhebliche Unterschied in der Verteilung der verschiedenen Ligaturtypen in unterschiedlichen Manuskripten dient den Musikwissenschaftlern als Ausgangspunkt zur weiteren Analyse und muss im Kontext weiterer Informationen über die Manuskripte interpretiert werden.

Die Qualität der Ligaturdetektion hängt stark von den Ergebnissen der Notendetektion und -klassifizierung ab. Entsprechend variiert die Qualität der Ergebnisse in den einzelnen Manuskripten teilweise erheblich. Da Manuskript M3 mit einer deutlich geringeren Auflösung digitalisiert wurde als die Manus-

	Quadratnotation	Rautennotation
Noten verbunden	ja	nein
Notentyp 1. Note	quadratisch	quadratisch
Notentyp ab 2. Note	quadratisch	rautenförmig

Tabelle 4: Ligaturkriterien

Manuskript	M1	M2	M3
#Ligaturtyp 1	2073	1089	3891
#Ligaturtyp 2	1398	549	1364

Tabelle 5: Ergebnisse der Ligaturredetektion

kripte M1 und M2, ist die Notenklassifikation per SVM bei M3 um 9,42% (im Vergleich zu M1) bzw. 17,44% (im Vergleich zu M2) schlechter als bei den anderen beiden Manuskripten, worunter auch die Qualität der darauf aufbauenden Ligaturredetektion leidet. Daher ist eine hohe Qualität bei der Digitalisierung der Manuskripten in jedem Fall vorteilhaft und erstrebenswert.

Es konnte gezeigt werden, dass Aufgaben in der Manuskriptforschung mit etablierten Methoden der computergestützten Bildverarbeitung sinnvoll unterstützt werden können. Durch die Definition von Suchmustern als Zusammenfassung mehrerer Einzelnoten mit bestimmten geometrischen Eigenschaften kann das System außerdem flexibel für die Detektion weiterer für Musikwissenschaftler interessante Notenkonstellationen eingesetzt werden. Die erzielten Ergebnisse können als Grundlage für weiterführende statistische Analysen und Fragestellungen dienen.

Die Forschung für diese Arbeit wurde im Rahmen des von der Deutschen Forschungsgemeinschaft (DFG) geförderten Sonderforschungsbereiches 950: „Manuskriptkulturen in Asien, Afrika und Europa“ durchgeführt.

## Literatur

- [1] A. Fornes and J. Lladós. A symbol-dependent writer identification approach in old handwritten music scores. In *Frontiers in Handwriting Recognition (ICFHR), 2010 International Conference on*, pages 634–639, Nov 2010.
- [2] U. Köthe. *Generische Programmierung für die Bildverarbeitung*. PhD thesis, University of Hamburg, 2000.
- [3] C. Ramirez and J. Ohya. Automatic recognition of square notation symbols in western plainchant manuscripts. *Journal of New Music Research*, 43(4):390–399, 2014.