

*“Which Line Is it Anyway”:  
Clustering for Staff Line Removal in  
Optical Music Recognition*

Sam Vinitzky

# Optical Music Recognition: An Overview

Allemanda



A musical score for a piece titled "Allemanda". The score is written in treble clef with a key signature of one sharp (F#) and a common time signature (C). It consists of eight staves of music. The notation includes various rhythmic values, including eighth and sixteenth notes, and rests. There are several trills marked with "tr" and some triplets marked with "3". The score is divided into two main sections: the first section ends at measure 11, and the second section begins at measure 12. The first section concludes with a double bar line and repeat signs. The second section continues with similar rhythmic patterns and includes a first ending bracket labeled "1." and a second ending bracket labeled "2.".

# Optical Music Recognition: An Overview

Allemanda



A musical score for a piece titled "Allemanda". The score is written in treble clef with a key signature of one sharp (F#) and a common time signature (C). It consists of eight staves of music. The notation includes various rhythmic values such as quarter notes, eighth notes, and sixteenth notes, along with rests and dynamic markings like *tr* (trill). There are also first and second endings indicated by bracketed numbers 1 and 2.



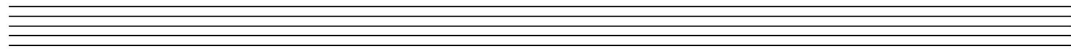
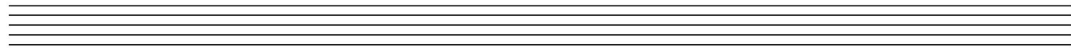
```
112 <!--=====-->
113 <part id="P1">
114 <measure number="1" width="324">
115 <print page-number="1">
116 <system-layout>
117 <system-margins>
118 <left-margin>150</left-margin>
119 <right-margin>0</right-margin>
120 </system-margins>
121 <top-system-distance>300</top-system-distance>
122 </system-layout>
123 <measure-numbering>system</measure-numbering>
124 </print>
125 <attributes>
126 <divisions>24</divisions>
127 <key>
128 <fifths>-3</fifths>
129 <mode>major</mode>
130 </key>
131 <time>
132 <beats>3</beats>
133 <beat-type>4</beat-type>
134 </time>
135 <clef>
136 <sign>G</sign>
137 <line>2</line>
138 </clef>
139 </attributes>
140 <note default-x="136">
141 <rest/>
142 <duration>24</duration>
143 <voice>1</voice>
144 <type>quarter</type>
145 </note>
146 <note default-x="196">
147 <pitch>
148 <step>B</step>
149 <alter>-1</alter>
150 <octave>4</octave>
151 </pitch>
152 <duration>24</duration>
153 <voice>1</voice>
154 <type>quarter</type>
155 <stem default-y="-55.5">down</stem>
156 </note>
157 <note default-x="256">
158 <pitch>
```

# Sheet Music Primer

## Allemanda

The image displays a page of sheet music for a piece titled "Allemanda". The music is written in G major (one sharp) and common time. It consists of four staves of music. The first staff begins with a treble clef and a key signature of one sharp (F#). The melody starts with a quarter note G, followed by a half note A, and then a quarter note B. A trill (tr) is placed over the first eighth note of the second measure. The second staff continues the melody with a triplet of eighth notes (G, A, B) in the first measure, followed by a quarter note C and a half note D. The third staff features a trill (tr) over the first eighth note of the first measure. The fourth staff contains several triplet markings (3) over groups of eighth notes. The piece concludes with a final cadence in G major.

# *Sheet Music Primer: Staff Lines*



# Sheet Music Primer: Symbols

Allemanda

The image displays a musical score for a piece titled "Allemanda". The score is written on four staves, each beginning with a treble clef and a key signature of two sharps (F# and C#). The time signature is common time (C). The music is characterized by intricate rhythmic patterns and various musical ornaments. The first staff includes two trills, indicated by the symbol "tr" above the notes. The second staff features a triplet of eighth notes, marked with a "3" above the notes. The third staff contains another trill, also marked with "tr". The fourth staff is filled with multiple triplet markings, each indicated by a "3" above the notes. The score is presented in a clear, professional layout, suitable for educational purposes in a music primer.

# Back to Optical Music Recognition...

Allemanda



A musical score for a piece titled "Allemanda". The score is written in treble clef with a key signature of one sharp (F#) and a 3/4 time signature. It consists of eight staves of music. The notation includes various rhythmic values such as eighth and sixteenth notes, as well as rests. There are several trills marked with "tr" and some triplets indicated by a "3" over a group of notes. The piece concludes with a double bar line and repeat signs.



```
112 <!--=====-->
113 <part id="P1">
114 <measure number="1" width="324">
115 <print page-number="1">
116 <system-layout>
117 <system-margins>
118 <left-margin>150</left-margin>
119 <right-margin>0</right-margin>
120 </system-margins>
121 <top-system-distance>300</top-system-distance>
122 </system-layout>
123 <measure-numbering>system</measure-numbering>
124 </print>
125 <attributes>
126 <divisions>24</divisions>
127 <key>
128 <fifths>-3</fifths>
129 <mode>major</mode>
130 </key>
131 <time>
132 <beats>3</beats>
133 <beat-type>4</beat-type>
134 </time>
135 <clef>
136 <sign>G</sign>
137 <line>2</line>
138 </clef>
139 </attributes>
140 <note default-x="136">
141 <rest/>
142 <duration>24</duration>
143 <voice>1</voice>
144 <type>quarter</type>
145 </note>
146 <note default-x="196">
147 <pitch>
148 <step>B</step>
149 <alter>-1</alter>
150 <octave>4</octave>
151 </pitch>
152 <duration>24</duration>
153 <voice>1</voice>
154 <type>quarter</type>
155 <stem default-y="-55.5">down</stem>
156 </note>
157 <note default-x="256">
158 <pitch>
```





# Optical Music Recognition: Pipeline

Allemanda



A musical score for a piece titled "Allemanda". The score is written in treble clef with a key signature of one sharp (F#) and a common time signature (C). It consists of eight staves of music. The notation includes various rhythmic values, including eighth and sixteenth notes, and rests. There are several trills marked with "tr" and some triplets marked with "3". The score is divided into two main sections: the first section ends at measure 11, and the second section begins at measure 12. The first section has a first ending bracket labeled "1." and a second ending bracket labeled "2.". The second section continues from measure 12 to the end of the piece. The music is characterized by a steady, rhythmic pattern with occasional melodic flourishes.

# Optical Music Recognition: Pipeline

## 1. Find and remove staff lines

Allemanda

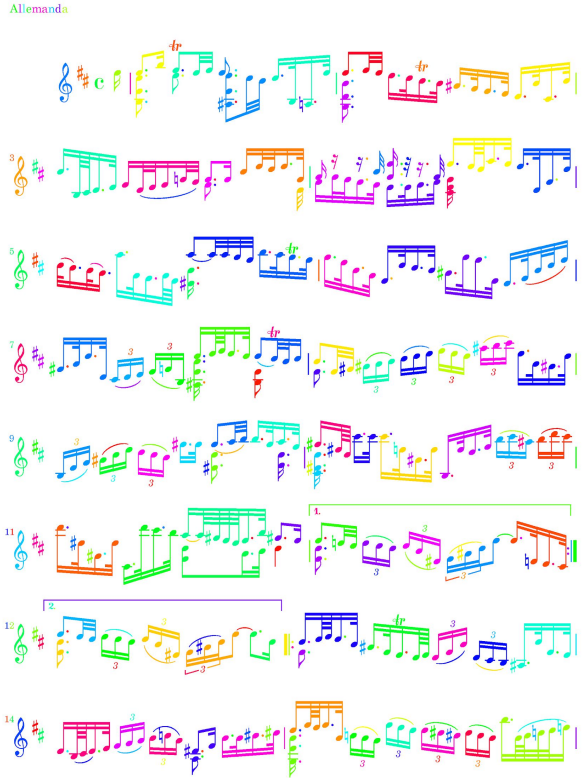


A musical score for a piece titled "Allemanda". The score is written in treble clef with a key signature of one sharp (F#) and a common time signature (C). It consists of eight staves of music. The first staff begins with a treble clef, a key signature of one sharp, and a common time signature. The music is written in a single system. The score includes various musical notations such as eighth notes, sixteenth notes, and triplets. There are also dynamic markings like *tr* (trill) and *tr* (trill) and *tr* (trill). The score is divided into two sections by a double bar line. The first section ends at measure 11, and the second section begins at measure 12. The score concludes with a double bar line and repeat dots.

# Optical Music Recognition: Pipeline

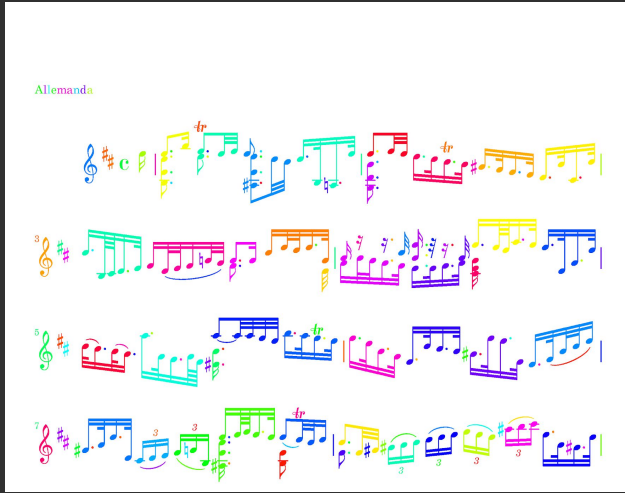
1. Find and remove staff lines
2. Find symbols

Allemanda



The image shows a page of musical notation for a piece titled "Allemanda". The notation is presented on eight staves, with the staff lines removed. The notes and rests are colored in various colors (red, green, blue, yellow, purple) to represent different symbols. The piece is in 3/4 time and features a key signature of one sharp (F#). The notation includes various rhythmic values, including eighth and sixteenth notes, and rests. There are also some special symbols like triplets and slurs. The piece is divided into two sections, labeled 1 and 2, with a repeat sign at the end of the first section.

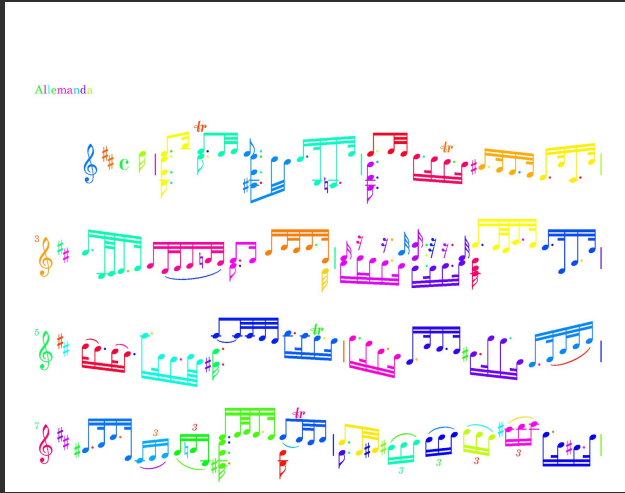
# Optical Music Recognition: Pipeline



1. Find and remove staff lines
2. Find symbols
3. Determine type of each symbol

1 = note  
2 = treble clef  
3 = sharp  
4 = sharp  
5 = three  
6 = note  
...

# Optical Music Recognition: Pipeline

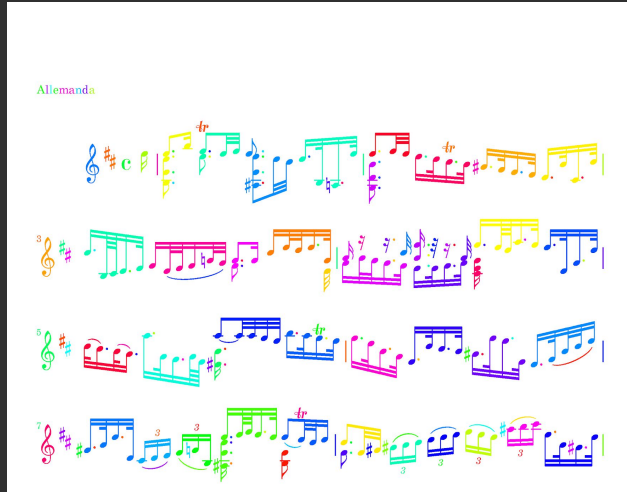


1. Find and remove staff lines
2. Find symbols
3. Determine type of each symbol
4. Determine how symbols relate to each other

1 = note (G4)  
2 = treble clef (first clef)  
3 = sharp (key sig with 4)  
4 = sharp (key sig with 3)  
5 = three (triplet for 6 8 11)  
6 = note (E5, triplet 5)

...

# Optical Music Recognition: Pipeline



1. Find and remove staff lines
2. Find symbols
3. Determine type of each symbol
4. Determine how symbols relate to each other
5. Convert into computer readable format (musicXML, MIDI, etc...)

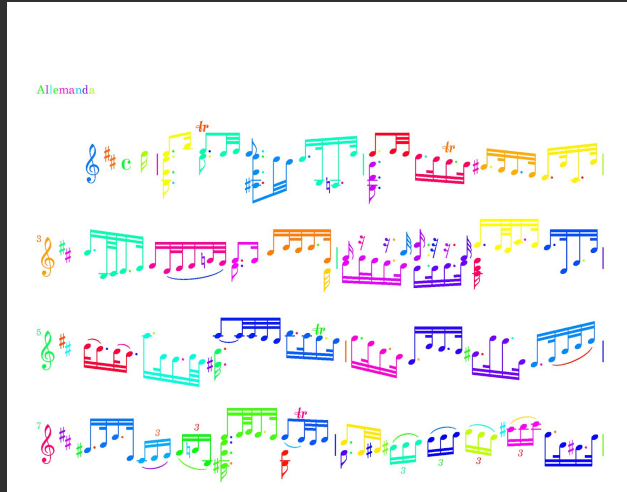
1 = note (G4)  
2 = treble clef (first clef)  
3 = sharp (key sig with 4)  
4 = sharp (key sig with 3)  
5 = three (triplet for 6 8 11)  
6 = note (E5, triplet 5)

...



```
112 <!------->
113 <part id="P1">
114   <measure number="1" width="324">
115     <print page-number="1">
116       <system-layout>
117         <system-margins>
118           <left-margin>150</left-margin>
119           <right-margin>0</right-margin>
120         </system-margins>
121         <top-system-distance>300</top-system-distance>
122       </system-layout>
123       <measure-numbering>system</measure-numbering>
124     </print>
125     <attributes>
126       <divisions>24</divisions>
127       <key>
128         <fifths>-3</fifths>
129         <mode>major</mode>
130       </key>
```

# Optical Music Recognition: Pipeline



1. \*\*\* Find and remove staff lines \*\*\*
2. Find symbols
3. Determine type of each symbol
4. Determine how symbols relate to each other
5. Convert into computer readable format (musicXML, MIDI, etc...)

1 = note (G4)  
2 = treble clef (first clef)  
3 = sharp (key sig with 4)  
4 = sharp (key sig with 3)  
5 = three (triplet for 6 8 11)  
6 = note (E5, triplet 5)

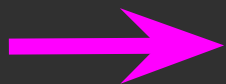
...



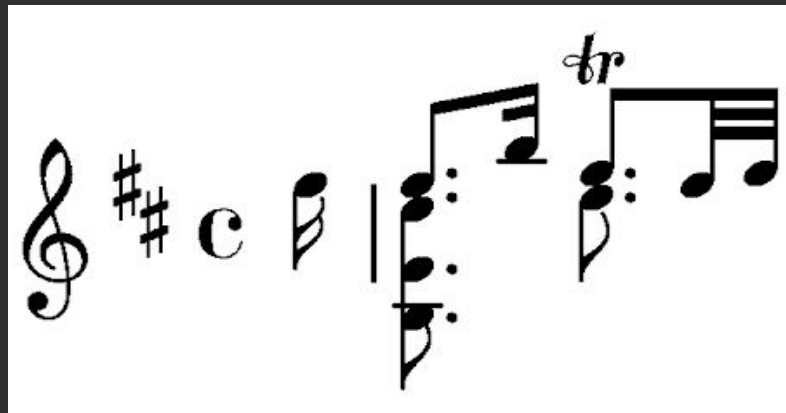
```
112 <!--=====-->
113 <part id="P1">
114   <measure number="1" width="324">
115     <print page-number="1">
116       <system-layout>
117         <system-margins>
118           <left-margin>150</left-margin>
119           <right-margin>0</right-margin>
120         </system-margins>
121         <top-system-distance>300</top-system-distance>
122       </system-layout>
123       <measure-numbering>system</measure-numbering>
124     </print>
125     <attributes>
126       <divisions>24</divisions>
127       <key>
128         <fifths>-3</fifths>
129         <mode>major</mode>
130       </key>
```

# Staff Line Removal: Overview

Input: original score image

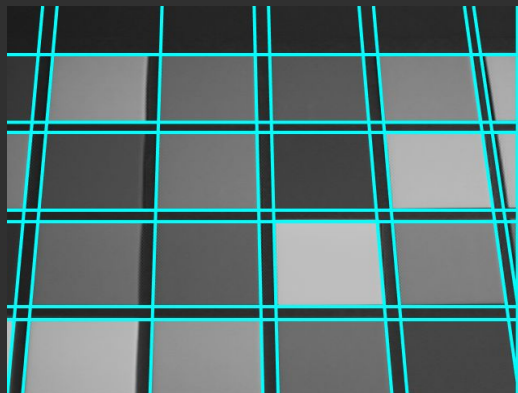
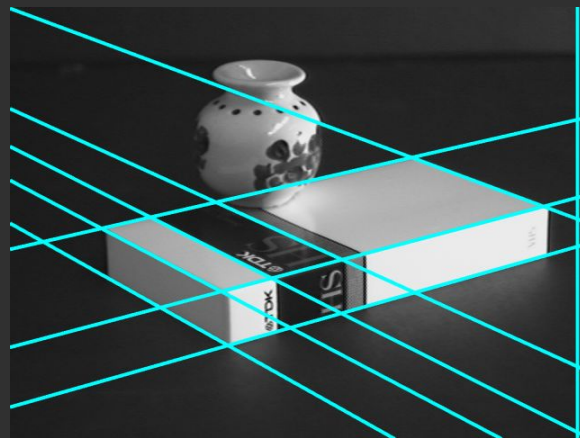
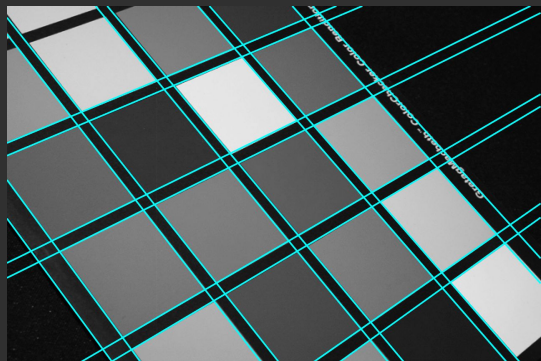


Output: staffless image





# Reminder: Hough Transform...



# Staff "Line" Removal...

The image displays a handwritten musical score consisting of five staves. The notation includes various musical symbols such as notes, rests, beams, and dynamic markings. The score is annotated with several performance instructions in French:

- Staff 1:** *modéré* (moderate), *f* (forte), *p* (piano), *sf* (sforzando).
- Staff 2:** *lent* (slow), *modéré*, *lent*, *Rall.* (ritardando), *res modéré* (return to moderate), *Presser* (press forward).
- Staff 3:** *Modéré*, *Presser toujours* (press forward always), *res a poco a poco* (return gradually), *mp cresc.* (mezzo-piano crescendo).
- Staff 4:** *movt. de valse* (waltz movement), *pp* (pianissimo), *cresc.* (crescendo).

The score also features various time signatures (e.g., 5/4, 3/4, 6/8, 4/4) and dynamic markings such as *mf* (mezzo-forte), *p* (piano), *pp* (pianissimo), and *sf* (sforzando). There are also markings for *tr* (trill) and *3* (triplets).

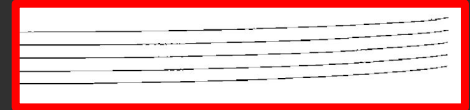
# Staff "Line" Removal...

The image shows a handwritten musical score with five staves. The notation includes various musical symbols such as clefs, time signatures, dynamics, and performance instructions. The staves are as follows:

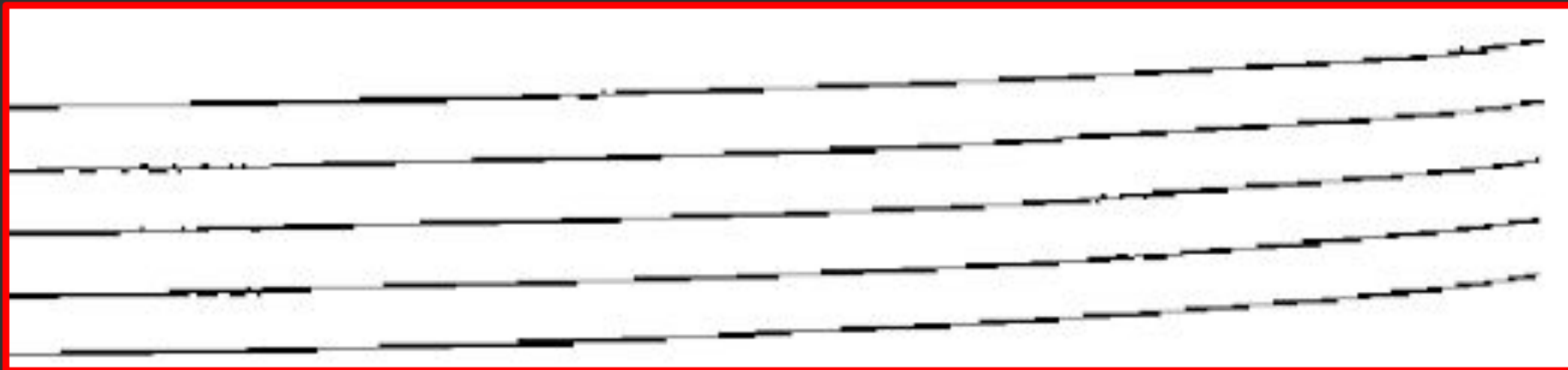
- Staff 1:** Treble clef, 5/4 time signature. Dynamics: *f*, *f*, *f*, *f*, *f*. Tempo: *Moderé*. Includes a fermata and a triplet.
- Staff 2:** Bass clef, 3/4 time signature. Dynamics: *mf*, *p*, *mf*. Tempo: *lent*, *Moderé*, *lent*. Includes a fermata, a triplet, and a section marked *Rall.* with a triplet.
- Staff 3:** Bass clef, 4/4 time signature. Dynamics: *p*, *mf*. Tempo: *Moderé*. Includes a fermata and a section marked *Presser*.
- Staff 4:** Treble clef, 4/4 time signature. Dynamics: *p*, *mf*. Tempo: *Moderé*. Includes a fermata, a section marked *Presser toujours*, and a section marked *res moderé* with a triplet.
- Staff 5:** Bass clef, 4/4 time signature. Dynamics: *mf*, *p*, *pp*. Tempo: *Moderé*. Includes a fermata, a section marked *mouv. de valse*, and a section marked *resc.* with a triplet.

The bottom staff (Staff 5) is highlighted with a red rectangular box.

# *Staff "Line" Removal...*



# *Staff "Line" Removal...*



# *Staff Line Removal: Algorithms*

## **Handcrafted:**

- Run-length [Carter & Bacon '92]
- Shortest Path [Cardoso et al. '08]
- Etc... [see Dalitz et al. '08]

# *Staff Line Removal: Algorithms*

## **Handcrafted:**

- Run-length [Carter & Bacon '92]
- Shortest Path [Cardoso et al. '08]
- Etc... [see Dalitz et al. '08]

## **Supervised Machine Learning:** (State of the Art)

- Classification of pixels [Calvo-Zaragoza et al. '16]
- Deep learning on image
  - Convolutional Neural Networks [Calvo-Zaragoza et al. '17]
  - Generative Adversarial Networks [Bhunja et al. '18]

# *Staff Line Removal: Algorithms*

## **Handcrafted:**

- Run-length [Carter & Bacon '92]
- Shortest Path [Cardoso et al. '08]
- Etc... [see Dalitz et al. '08]

## **Supervised Machine Learning:** (State of the Art)

- Classification of pixels [Calvo-Zaragoza et al. '16]
- Deep learning on image
  - Convolutional Neural Networks [Calvo-Zaragoza et al. '17]
  - Generative Adversarial Networks [Bhunja et al. '18]

## **Unsupervised Machine Learning:**

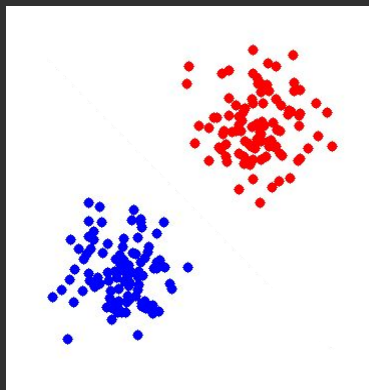
- Clustering of pixels [Vinitzky '18]



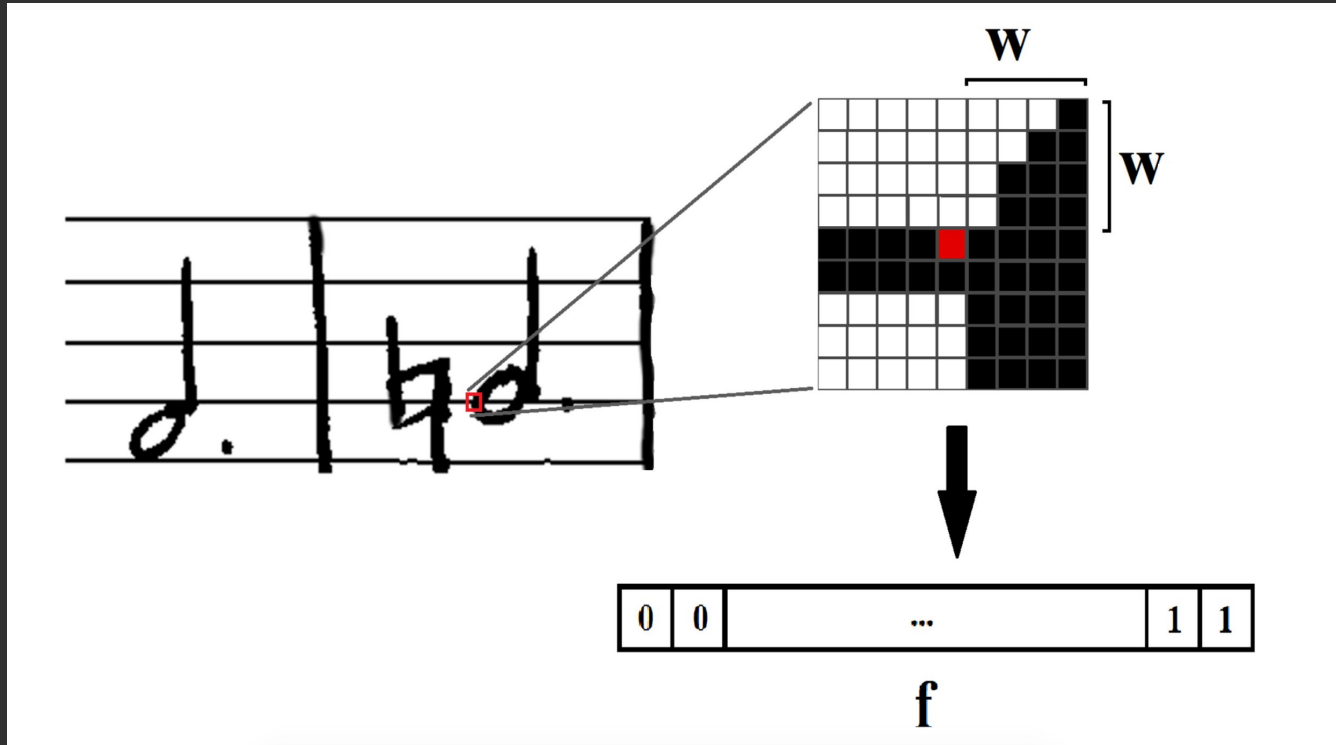
# *Staff Line Removal via Pixel Clustering*

## **Algorithm:** (Vinitzky '18)

1. Convert each black pixel into a feature vector



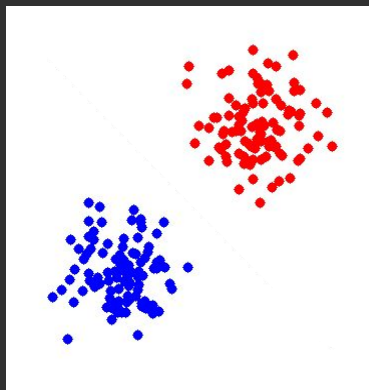
# Staff Line Removal via Pixel Clustering



# *Staff Line Removal via Pixel Clustering*

## **Algorithm:** (Vinitzky '18)

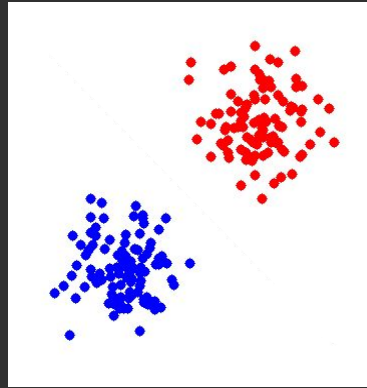
1. Convert each black pixel into a feature vector



# Staff Line Removal via Pixel Clustering

## Algorithm: (Vinitisky '18)

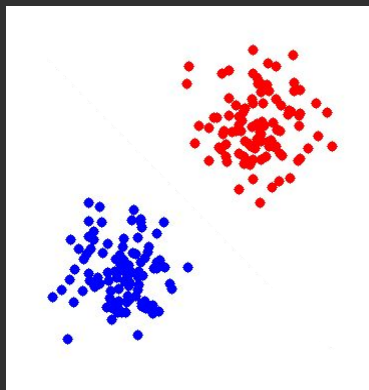
1. Convert each black pixel into a feature vector
2. “Cluster” the feature vectors into two groups



# Staff Line Removal via Pixel Clustering

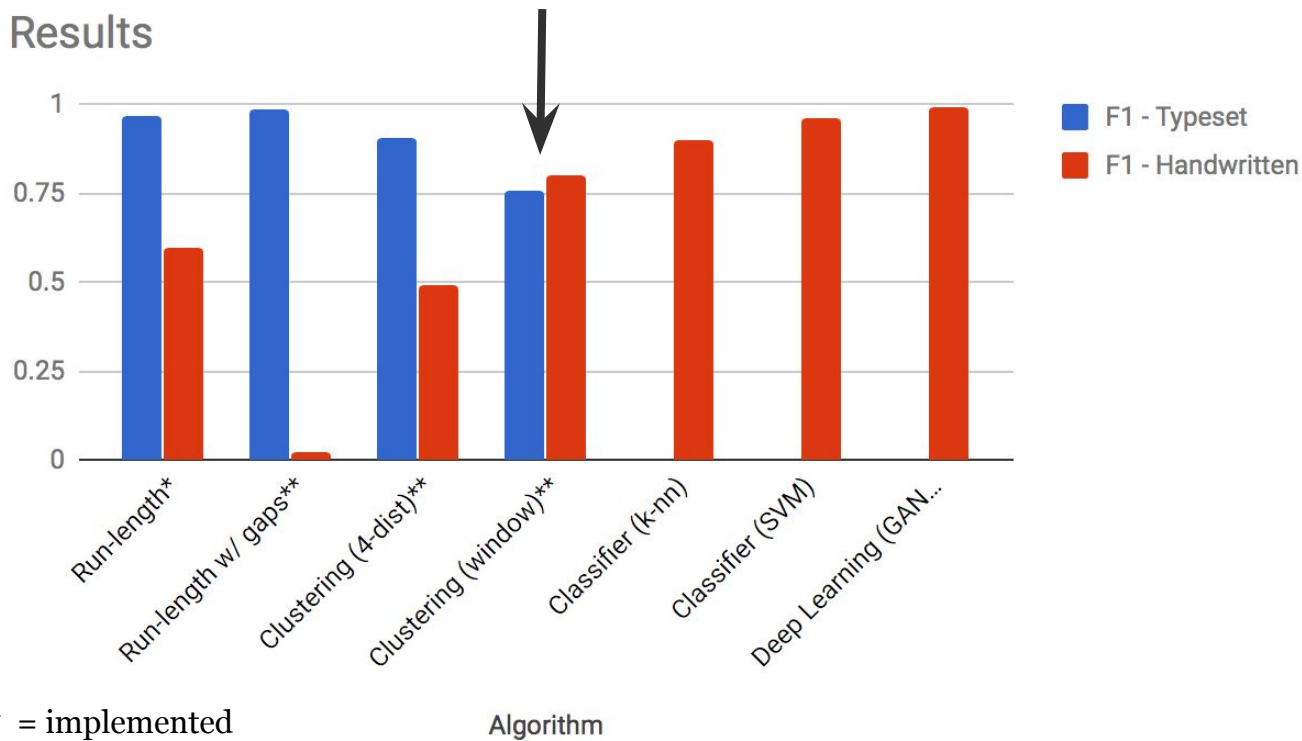
## Algorithm: (Vinitzky '18)

1. Convert each black pixel into a feature vector
2. “Cluster” the feature vectors into two groups
3. Figure out which cluster is “staff” and which is “non-staff”



# Staff Line Removal: Results

## Results



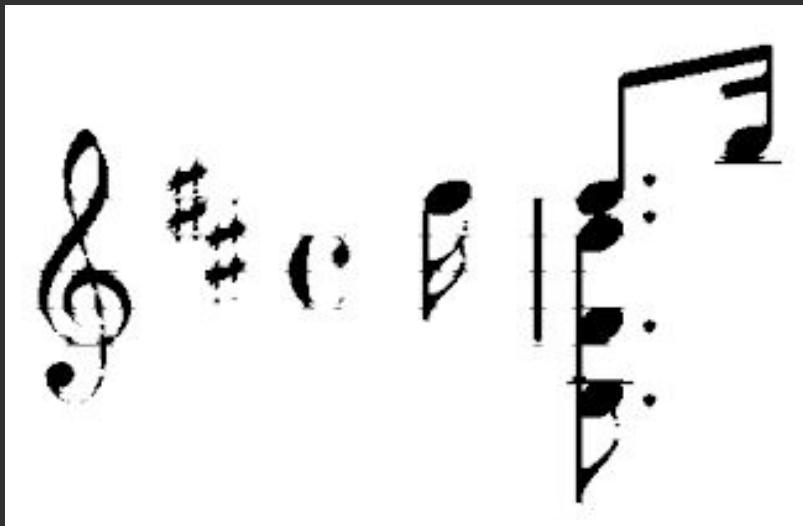
\* = implemented

\*\* = original algorithm

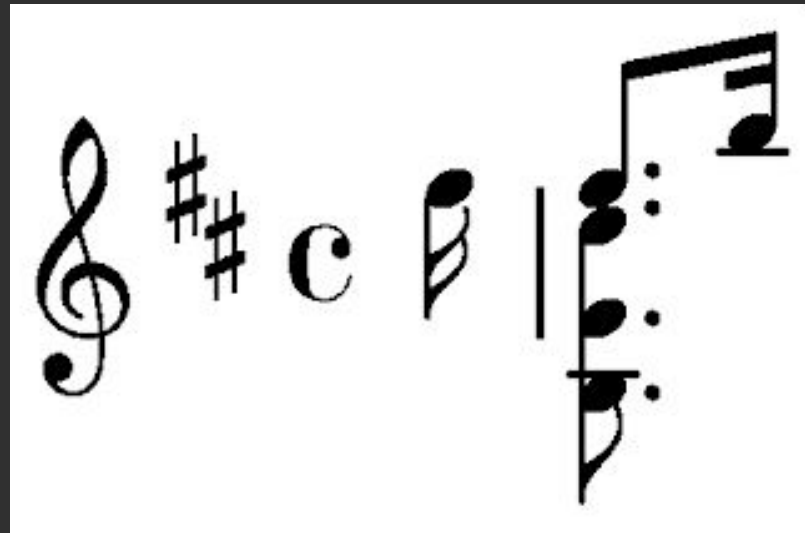
Algorithm

# *Staff Line Removal: Results*

Clustering output

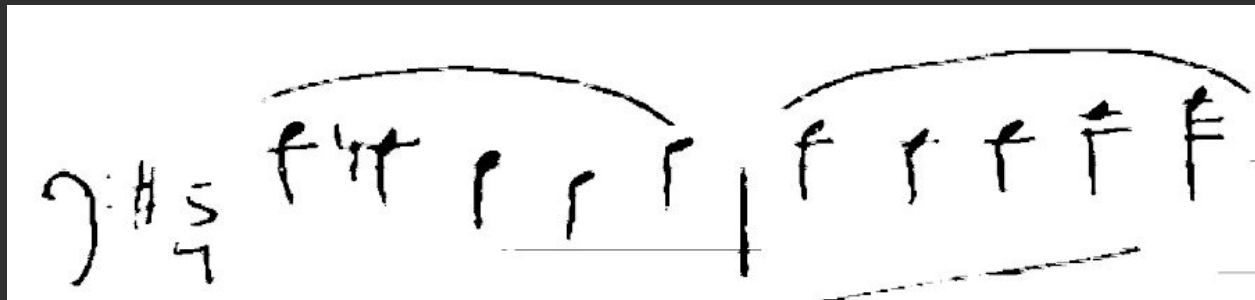


Ground truth staff-less



# Staff Line Removal: Results

Clustering output



Ground truth staff-less





## *Future Work*

### **Clustering for Staff Line Removal:**

- Other clustering methods

## *Future Work*

### **Clustering for Staff Line Removal:**

- Other clustering methods
- Better feature vectors

## *Future Work*

### **Clustering for Staff Line Removal:**

- Other clustering methods
- Better feature vectors
- Picking the “staff” cluster

## *Future Work*

### **Clustering for Staff Line Removal:**

- Other clustering methods
- Better feature vectors
- Picking the “staff” cluster
- White pixels can be staff (due to noise)

## *Future Work*

### **Clustering for Staff Line Removal:**

- Other clustering methods
- Better feature vectors
- Picking the “staff” cluster
- White pixels can be staff (due to noise)

### **Optical Music Recognition:**

- The rest of the pipeline...

# References

- [1] Jorge Calvo-Zaragoza, Luisa Mico, and Jose Oncina. Music staff removal with supervised pixel classification. *International Journal on Document Analysis and Recognition (IJDAR)*, 19(3):211–219, Sep 2016.
- [2] Jorge Calvo-Zaragoza, Jose J. Valero-Mas, and Antonio Pertusa. End-to-end optical music recognition using neural networks. In *ISMIR*, 2017.
- [3] Jaime Cardoso, Artur Capela, Ana Rebelo, and Carlos Guedes. A connected path approach for staff detection on a music score, 2008.
- [4] I. Fujinaga, M. Droettboom, C. Dalitz, and B. Pranzas. A comparative study of staff removal algorithms. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 30:753–766, 2008.
- [5] A. Konwer, A. K. Bhunia, A. Bhowmick, P. Banerjee, P. Pratim Roy, and U. Pal. Staff line Removal using Generative Adversarial Networks. *ArXiv e-prints*, 2018.
- [6] J. Novotny and J. Pokorny. Introduction to optical music recognition: Overview and practical challenges. 2015.
- [7] N.P. Carter, R.A. Bacon: *Automatic Recognition of Printed Music*. In H.S. Baird, H. Bunke, K. Yamamoto (editors): “Structured Document Image Analysis”, pp. 454-65, Springer, 1992.