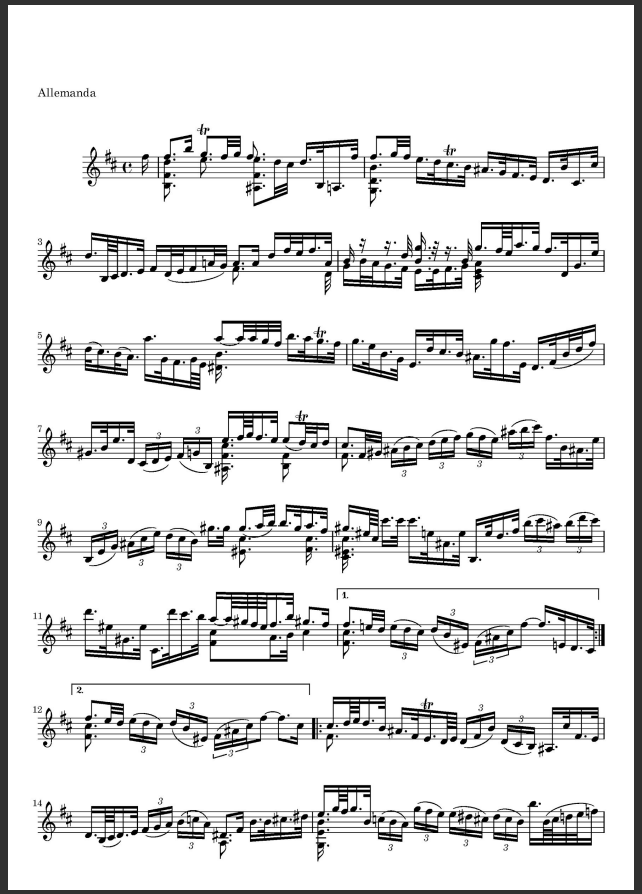
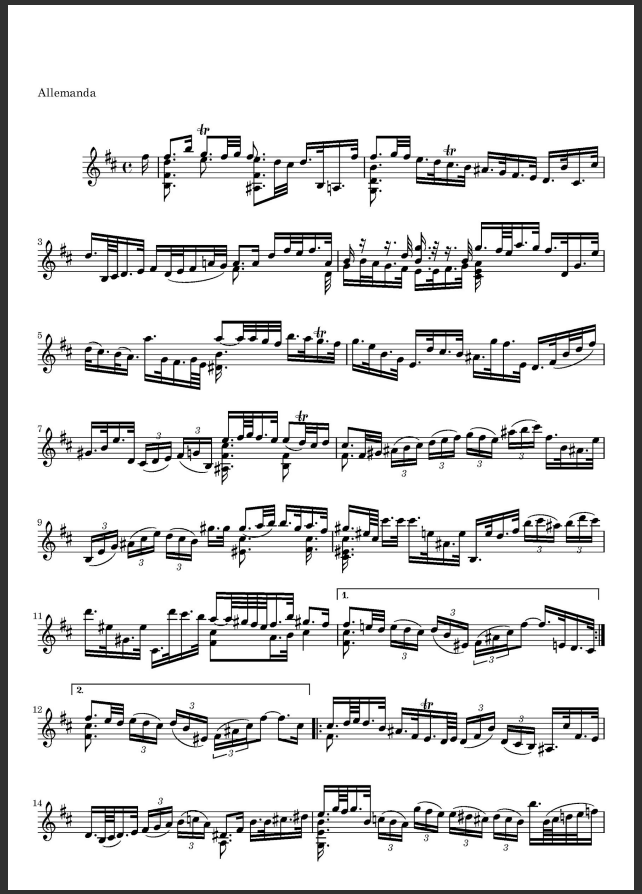
"Which Line Is it Anyway": Clustering for Staff Line Removal in Optical Music Recognition

Sam Vinitsky

Optical Music Recognition: An Overview

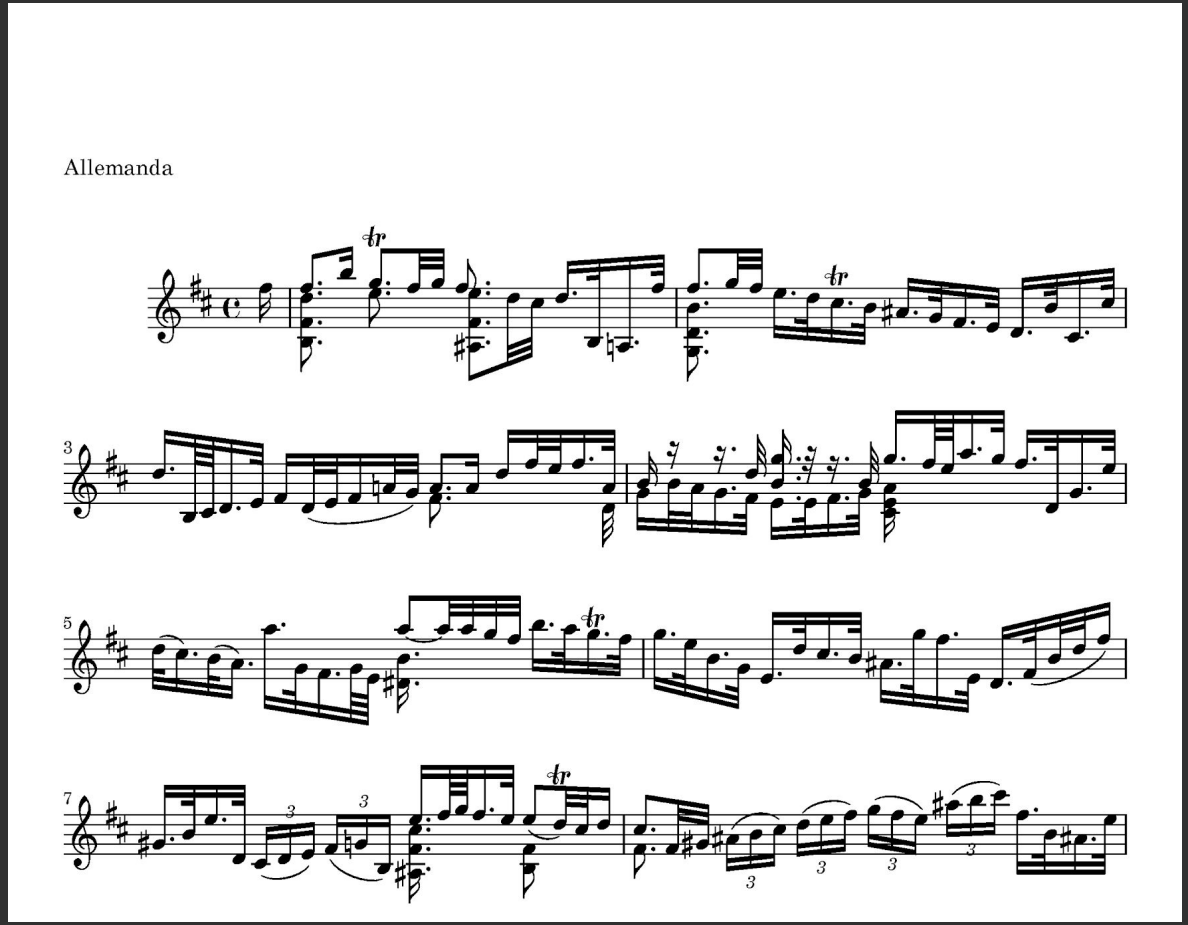


Optical Music Recognition: An Overview

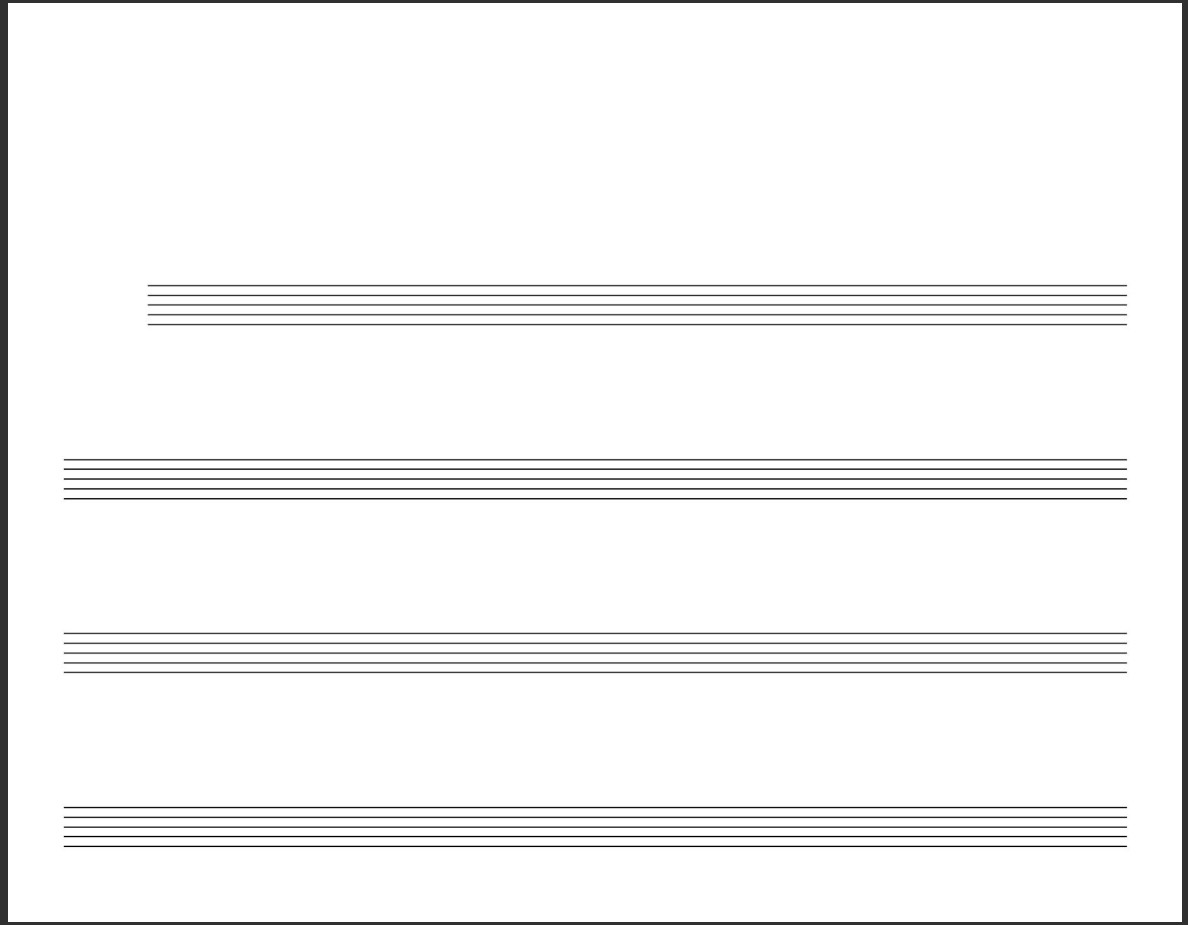


```
<!--=========================================================-->
<part id="P1">
  <measure number="1" width="324">
    <print page-number="1">
      <system-layout>
        <system-margins>
          <left-margin>150</left-margin>
          <right-margin>0</right-margin>
        </system-margins>
        <top-system-distance>300</top-system-distance>
      </system-layout>
      <measure-numbering>system</measure-numbering>
    </print>
    <attributes>
      <divisions>24</divisions>
      <key>
        <fifths>-3</fifths>
        <mode>major</mode>
      </key>
      <time>
        <beats>3</beats>
        <beat-type>4</beat-type>
      </time>
      <clef>
        <sign>G</sign>
        <line>2</line>
      </clef>
    </attributes>
    <note default-x="136">
      <rest/>
      <duration>24</duration>
      <voice>1</voice>
      <type>quarter</type>
    </note>
    <note default-x="196">
      <pitch>
        <step>B</step>
        <alter>-1</alter>
        <octave>4</octave>
      </pitch>
      <duration>24</duration>
      <voice>1</voice>
      <type>quarter</type>
      <stem default-y="-55.5">down</stem>
    </note>
    <note default-x="256">
      <pitch>
```

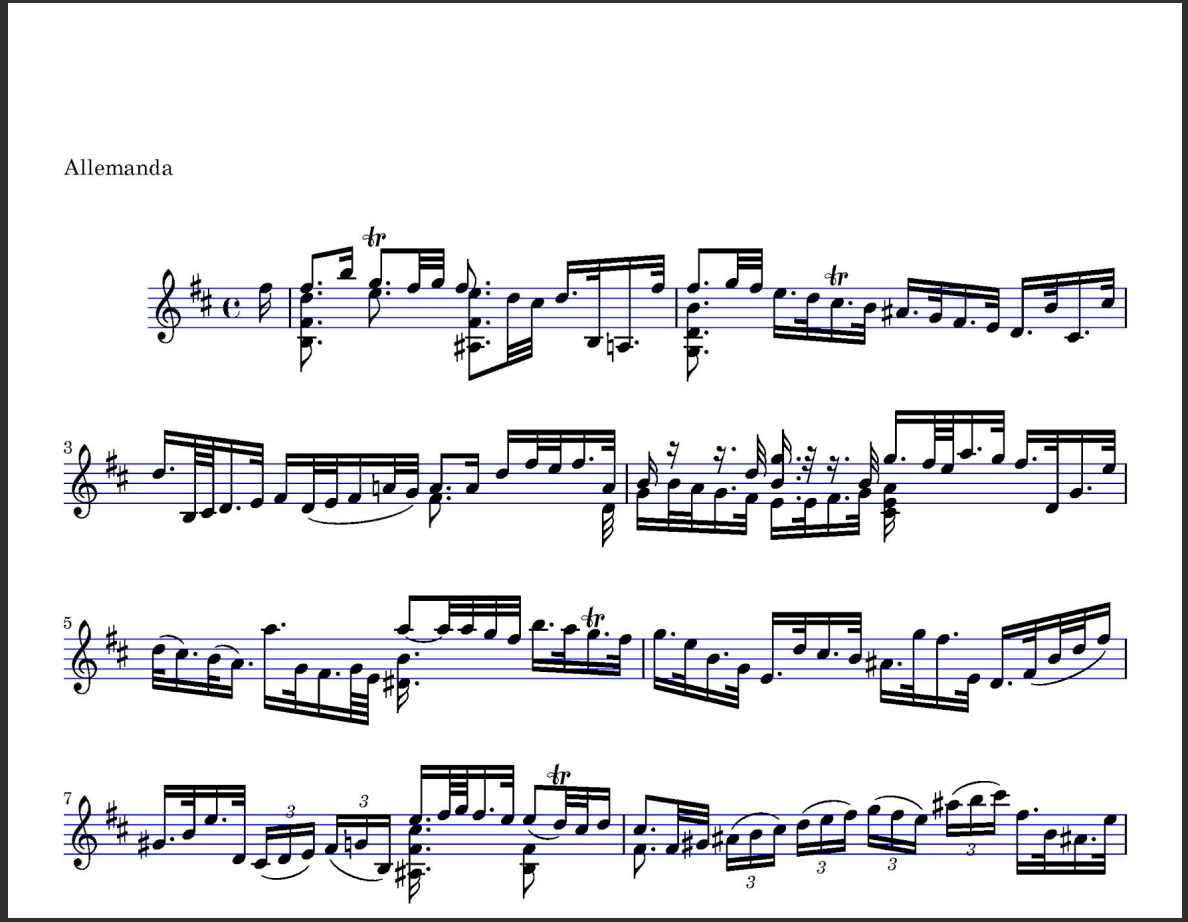
Sheet Music Primer



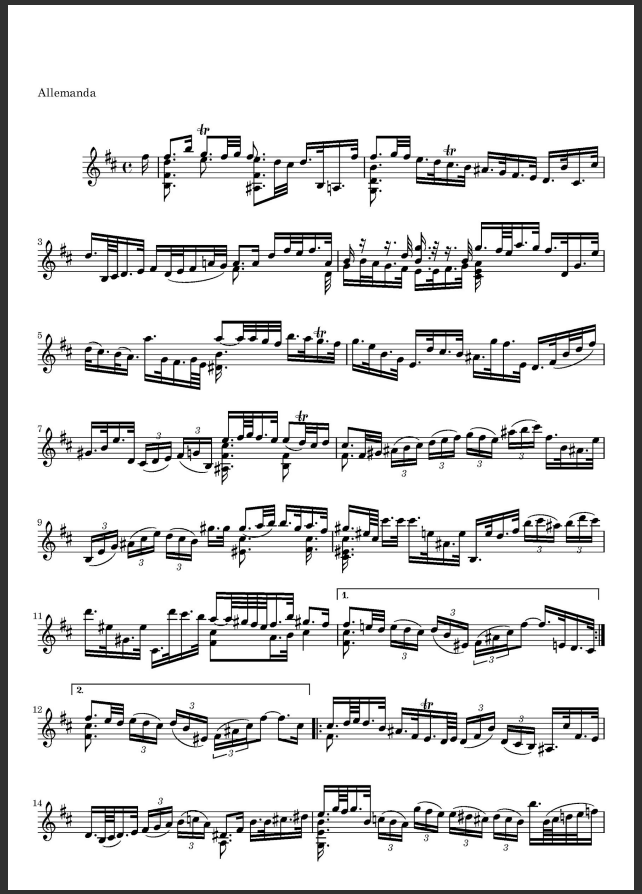
Sheet Music Primer: Staff Lines



Sheet Music Primer: Symbols

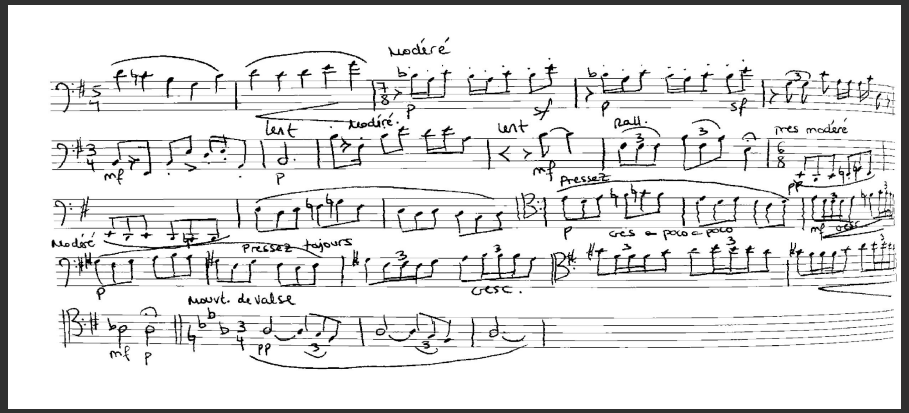


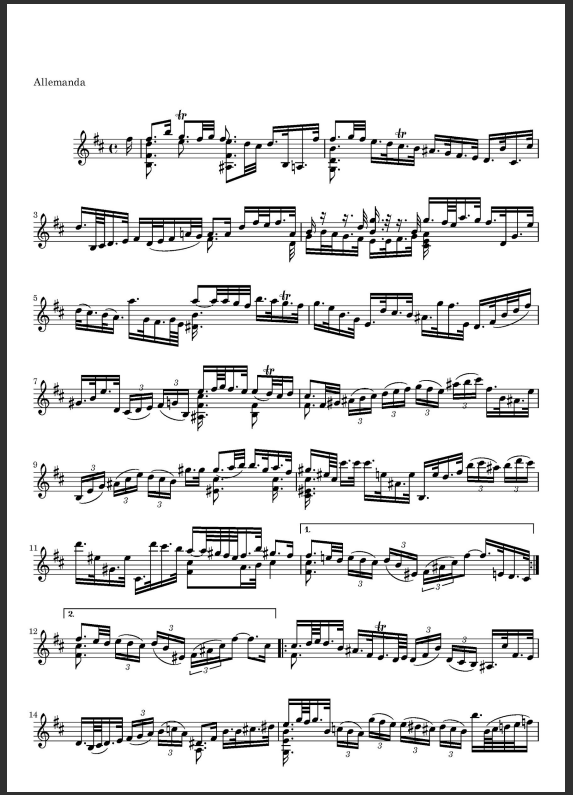
Back to Optical Music Recognition...



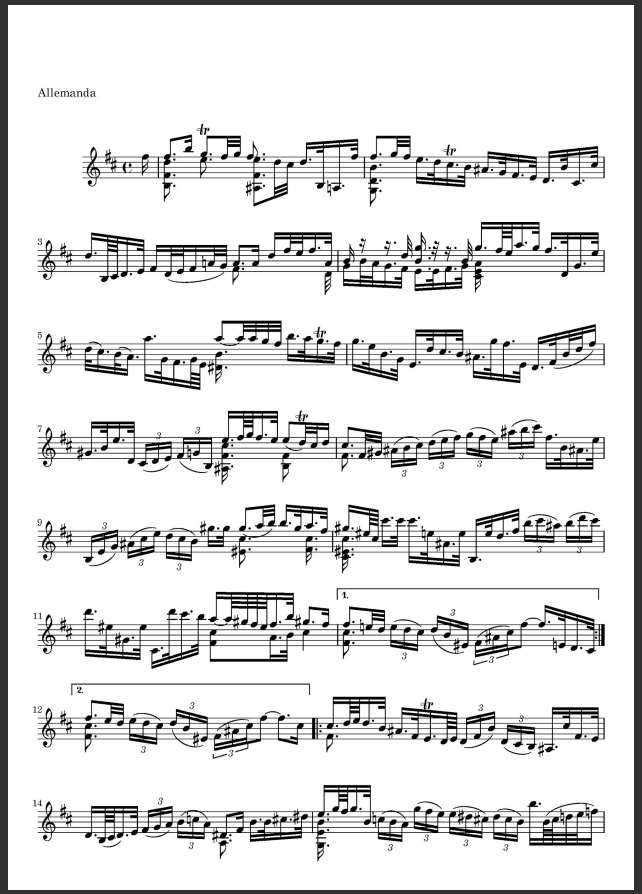
```
112   <!--=========================================================-->
113   <part id="P1">
114     <measure number="1" width="324">
115       <print page-number="1">
116         <system-layout>
117           <system-margins>
118             <left-margin>150</left-margin>
119             <right-margin>0</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>
```

Applications of OMR: Handwritten Music

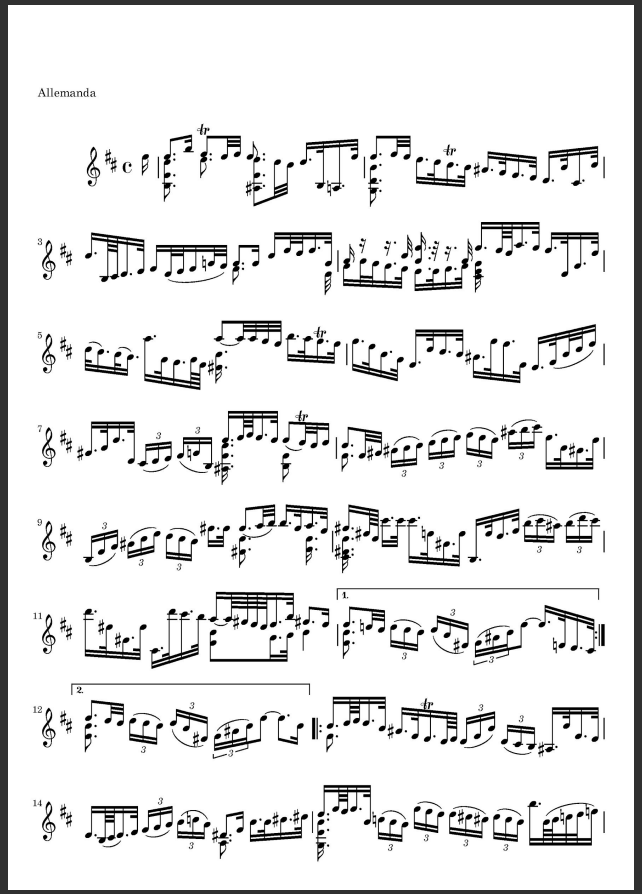




Optical Music Recognition: Pipeline

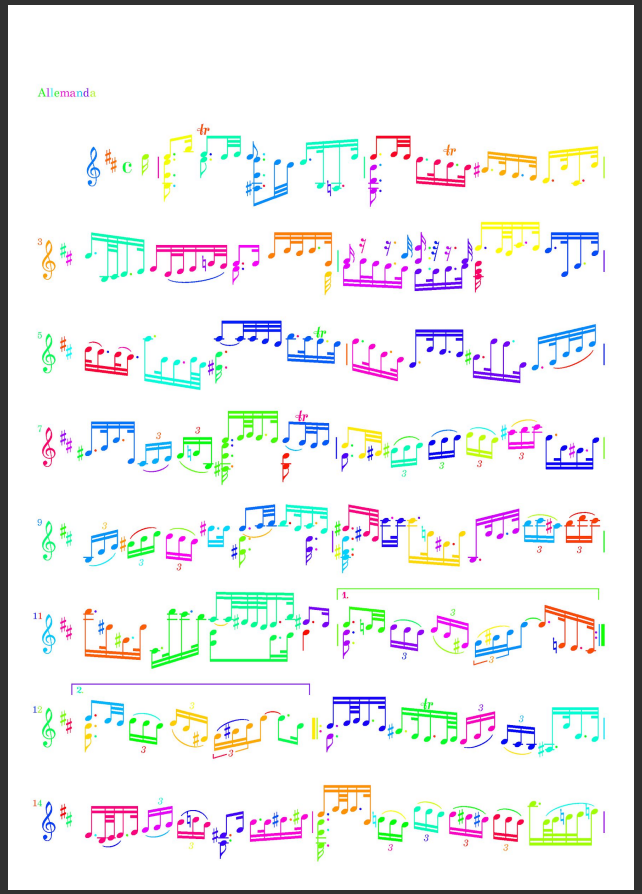


Optical Music Recognition: Pipeline



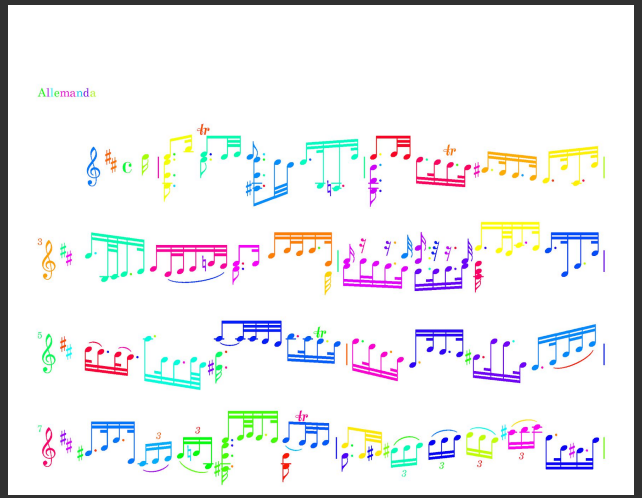
1. Find and remove staff lines

Optical Music Recognition: Pipeline



1. Find and remove staff lines
2. Find symbols

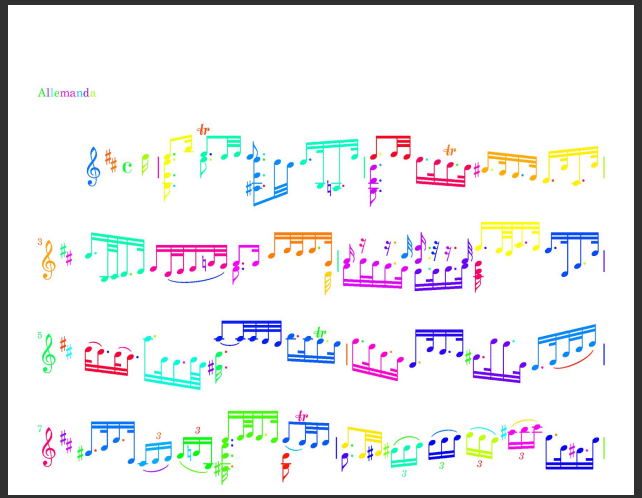
Optical Music Recognition: Pipeline



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

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

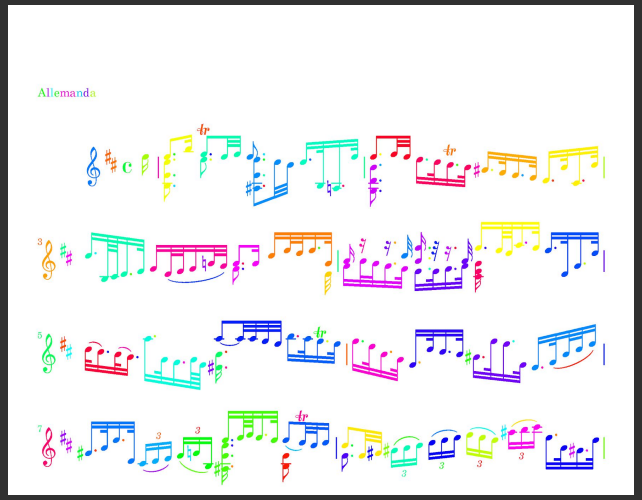
Optical Music Recognition: Pipeline



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

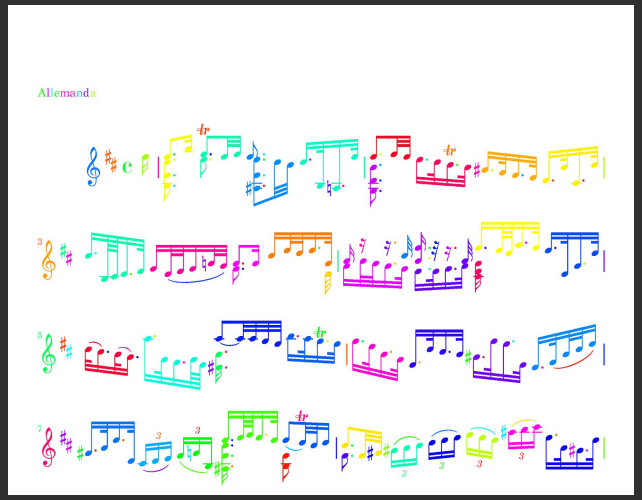


1. Find and remove staff lines
2. Find symbols
3. Determine type of each symbol
4. Determine how symbols relate to each other
5. Convert into computer readable format (musicXML, MIDI, etc...)

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)
...

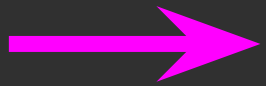
```
112   <!--=========================================================-->
113   <part id="P1">
114     <measure number="1" width="324">
115       <print page-number="1">
116         <system-layout>
117           <system-margins>
118             <left-margin>150</left-margin>
119             <right-margin>0</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...)

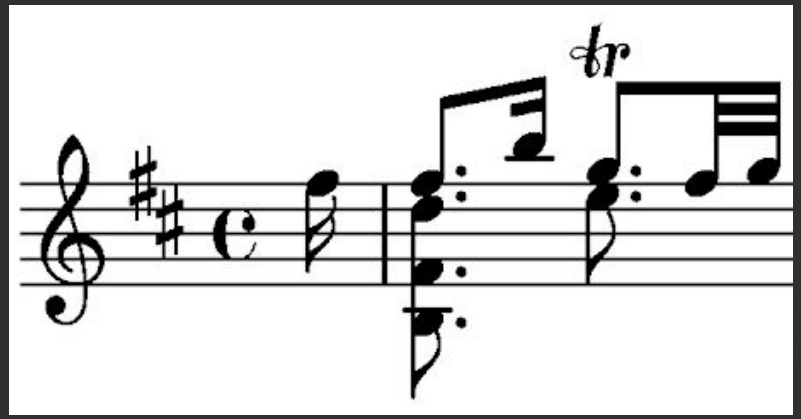
1 = note (G4)
2 = treble clef (first clef)
3 = sharp (key sig with 4)
4 = sharp (key sig with 3)
5 = three (triplet for 6 8 11)
6 = note (E5, triplet 5)
...



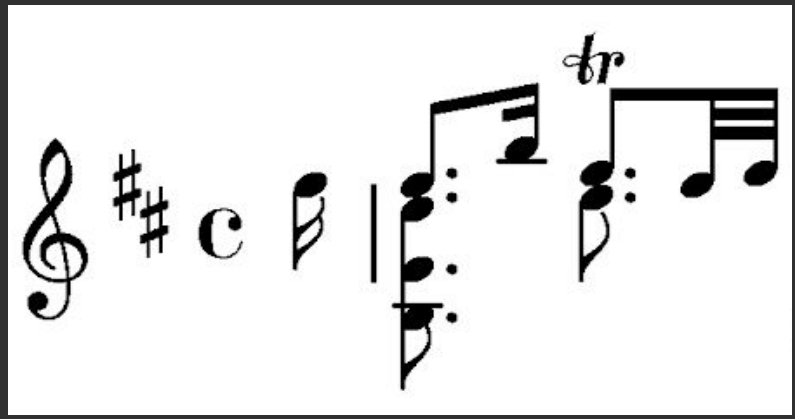
```
112   <!--=========================================================-->
113   <part id="P1">
114     <measure number="1" width="324">
115       <print page-number="1">
116         <system-layout>
117           <system-margins>
118             <left-margin>150</left-margin>
119             <right-margin>0</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

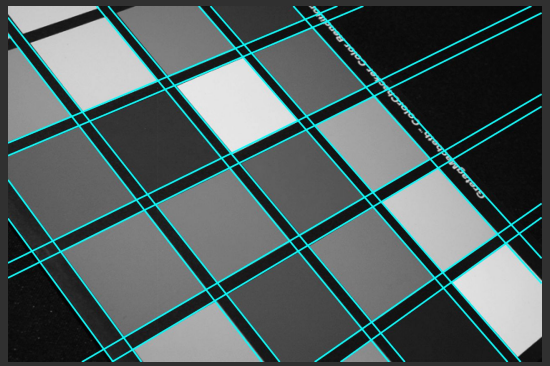
Input: original score image

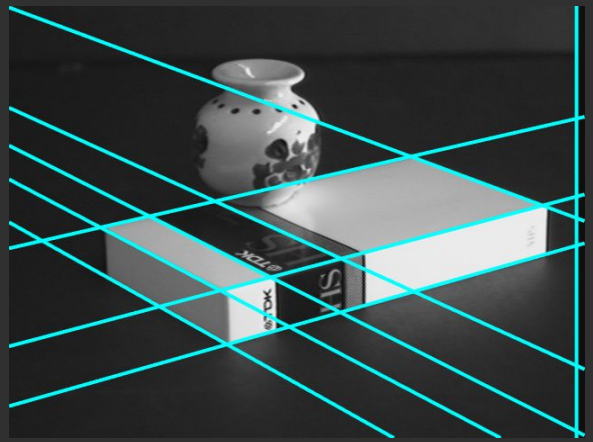


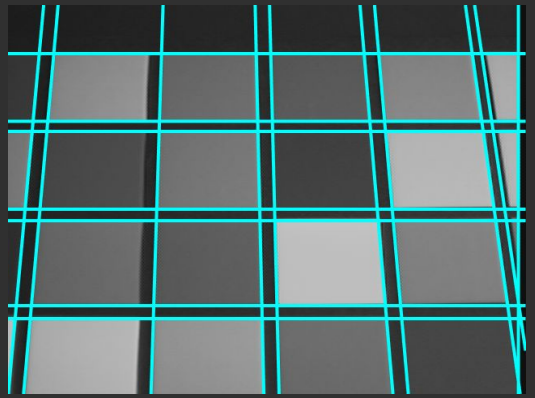
Output: staffless image



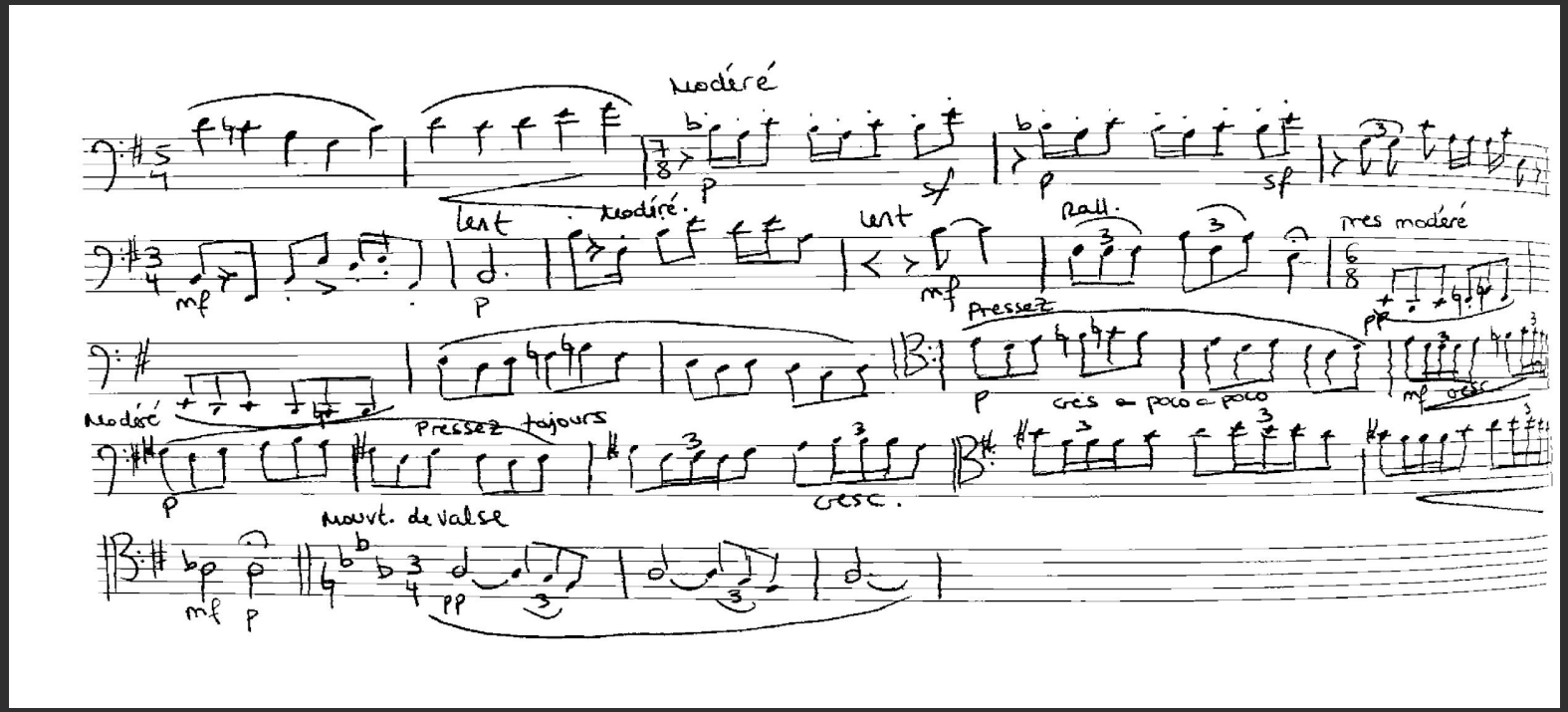
Reminder: Hough Transform...



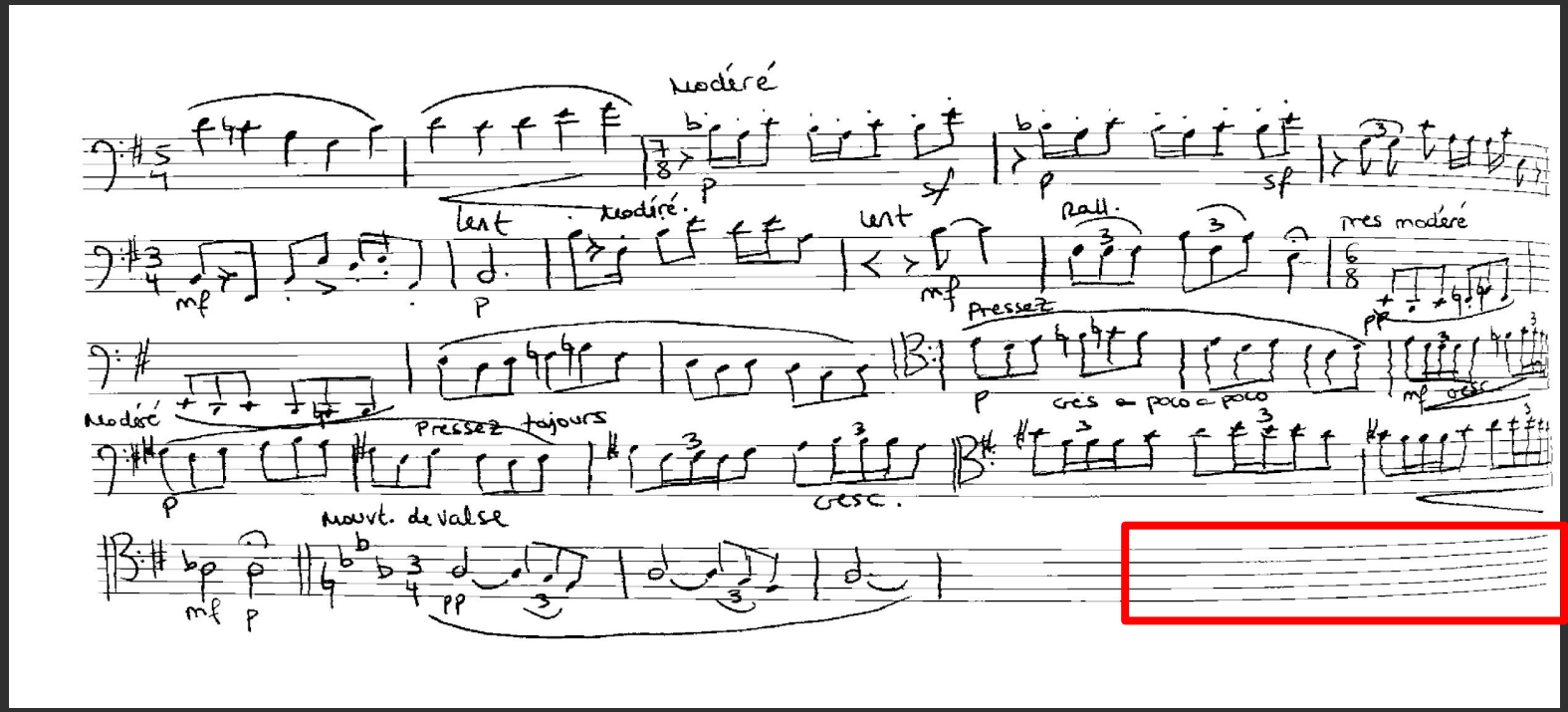




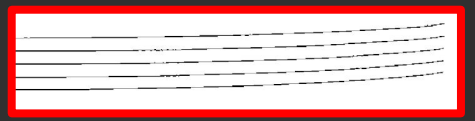
Staff "Line" Removal...



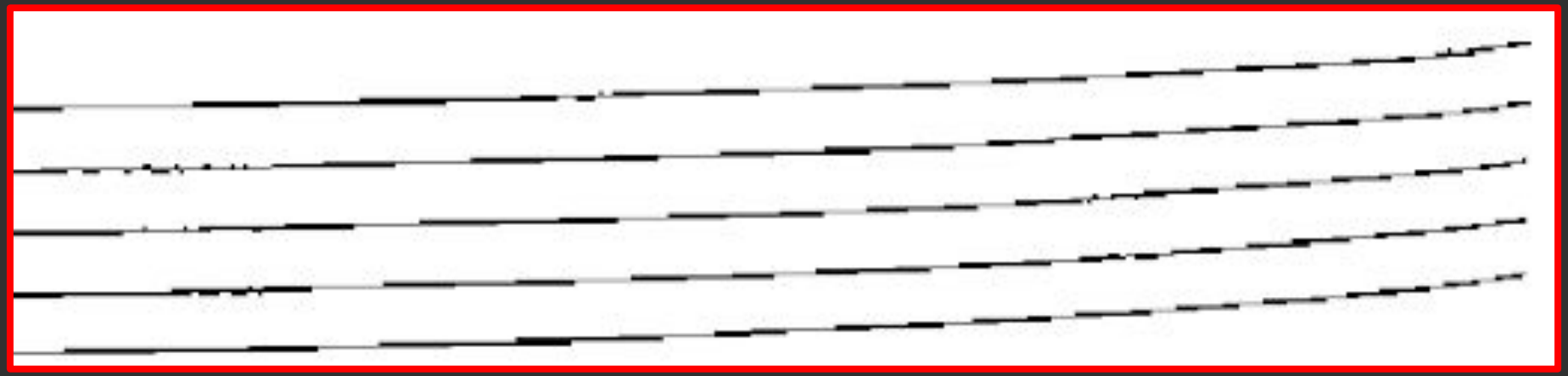
Staff "Line" Removal...



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-Zarogoza et al. ‘16]
- Deep learning on image
  - Convolutional Neural Networks [Calvo-Zarogoza et al. ‘17]
  - Generative Adversarial Networks [Bhunia 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-Zarogoza et al. ‘16]
- Deep learning on image
  - Convolutional Neural Networks [Calvo-Zarogoza et al. ‘17]
  - Generative Adversarial Networks [Bhunia et al. ‘18]

Unsupervised Machine Learning:

- Clustering of pixels [Vinitsky ‘18]

Staff Line Removal via Pixel Clustering

Algorithm: (Vinitsky ‘18)

1. Convert each black pixel into a feature vector

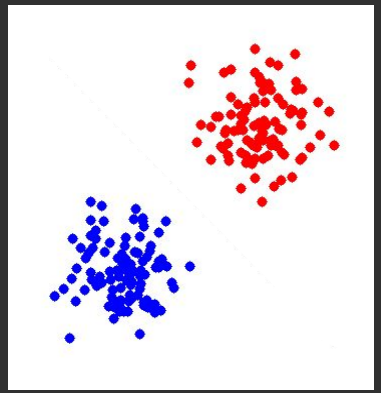


Image modified from http://sli.ics.uci.edu/Classes-CS178-Notes/LinearClassify

Staff Line Removal via Pixel Clustering

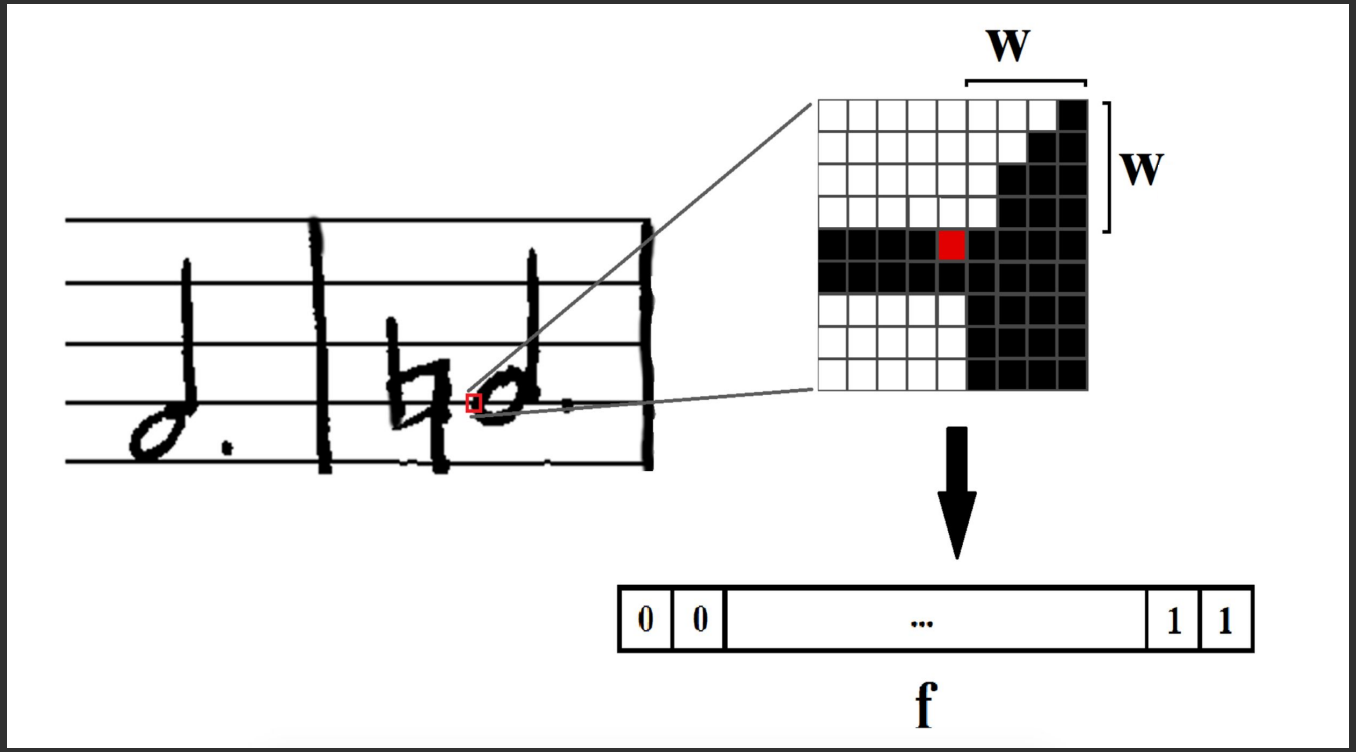


Image source: Calvo-Zaragova et al., ‘16

Staff Line Removal via Pixel Clustering

Algorithm: (Vinitsky ‘18)

1. Convert each black pixel into a feature vector

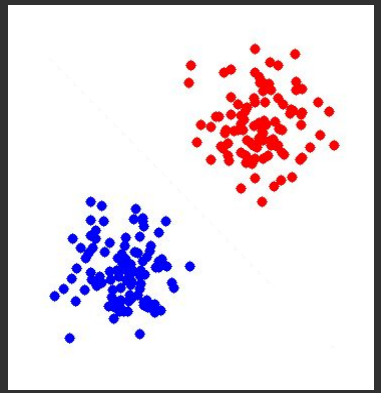


Image modified from http://sli.ics.uci.edu/Classes-CS178-Notes/LinearClassify

Staff Line Removal via Pixel Clustering

Algorithm: (Vinitsky ‘18)

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

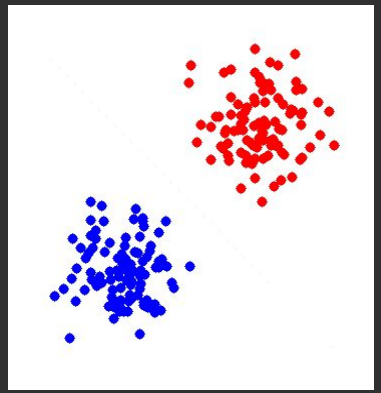


Image modified from http://sli.ics.uci.edu/Classes-CS178-Notes/LinearClassify

Staff Line Removal via Pixel Clustering

Algorithm: (Vinitsky ‘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”

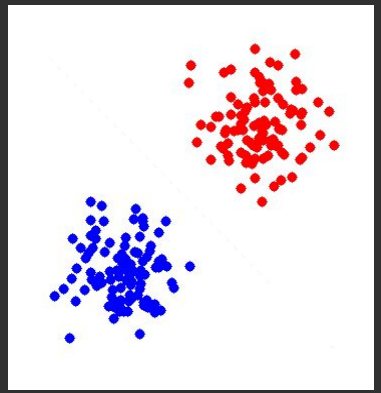
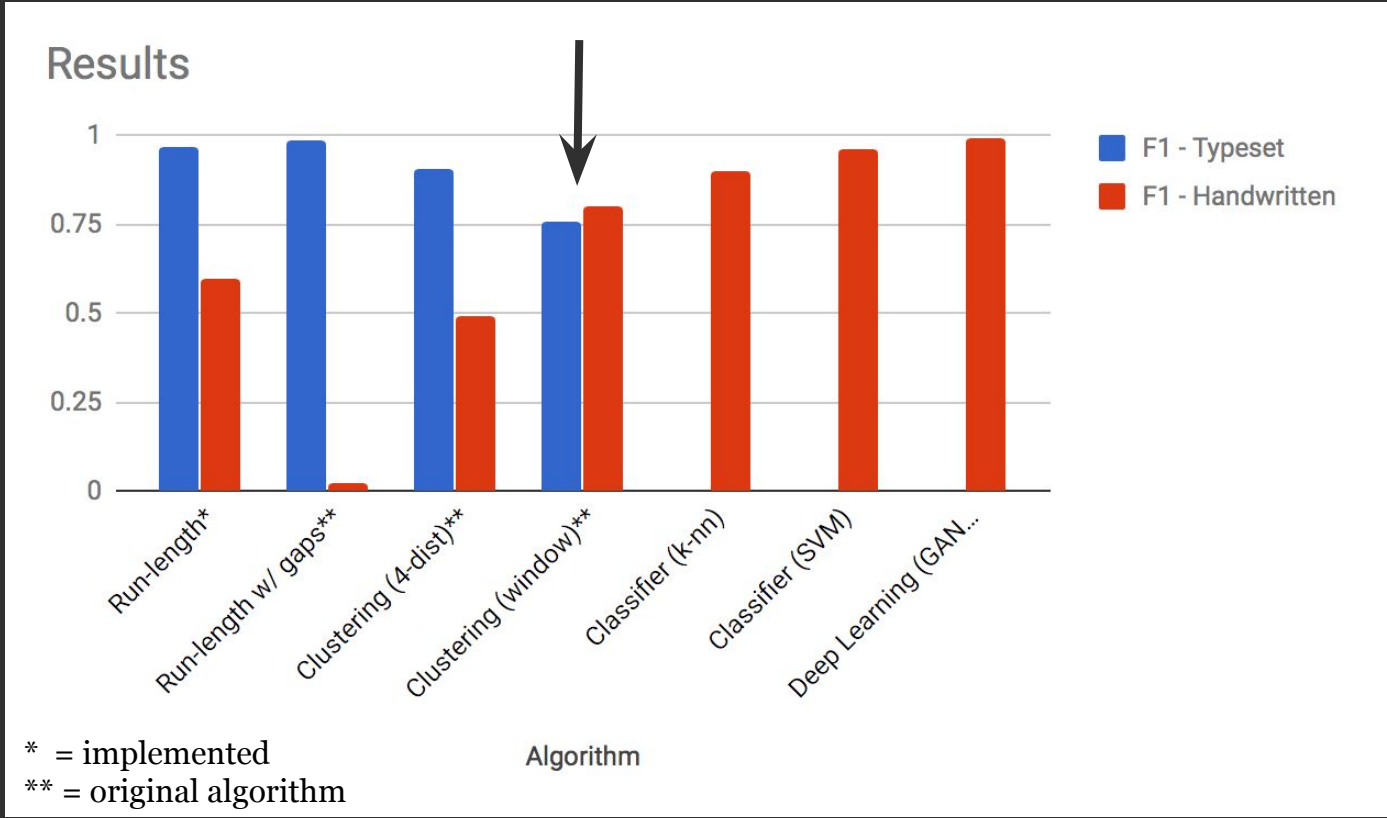


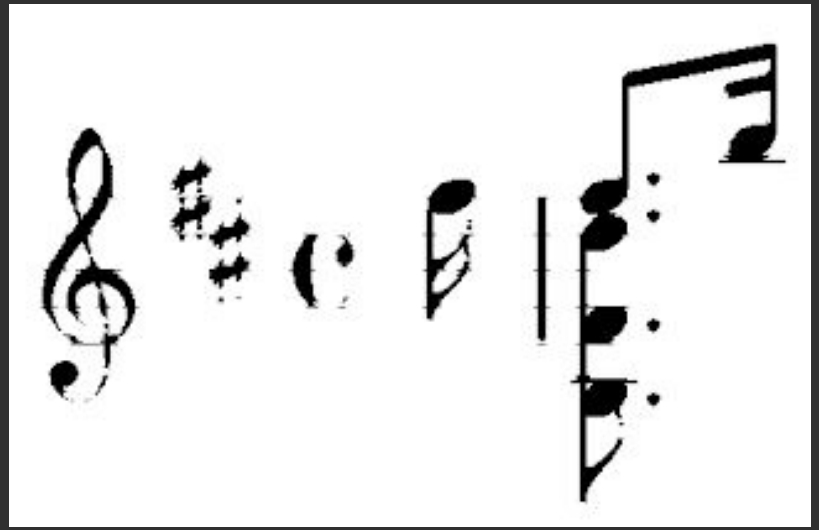
Image modified from http://sli.ics.uci.edu/Classes-CS178-Notes/LinearClassify

Staff Line Removal: Results

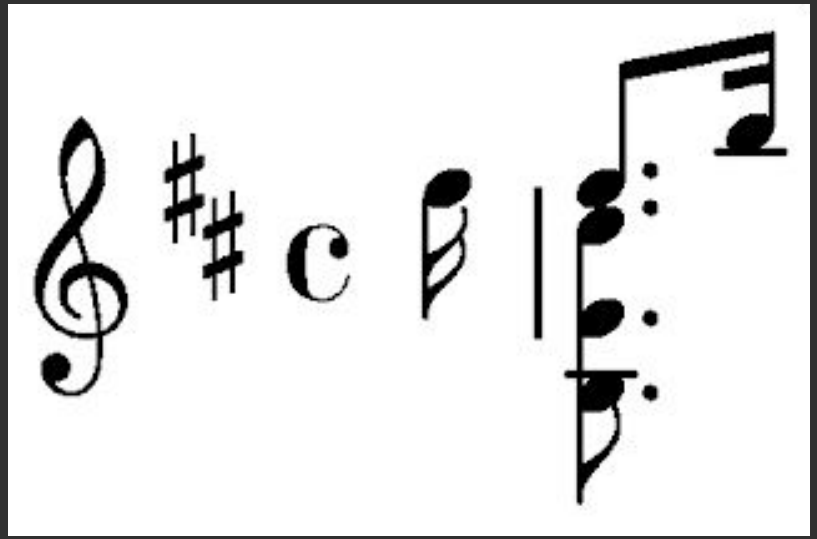


Staff Line Removal: Results

Clustering output

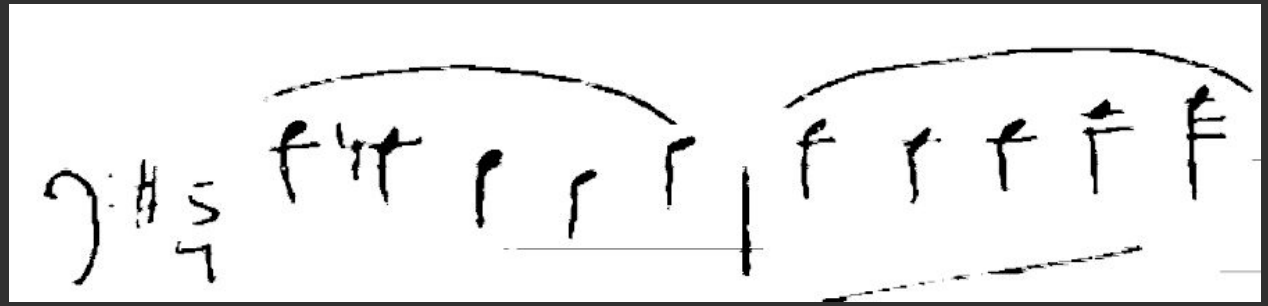


Ground truth staff-less



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.