Optical Music Recognition and Data Import/Export

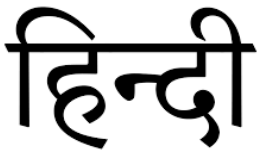
MUSIC 253/ CS 275A

ELEANOR SELFRIDGE-FIELD

Optical recognition for text

Reliable to roughly 96-99% for Roman alphabet

- **Good** when
  - Content is even and regular
  - Scanning is carefully fed
- **Less good** when
  - Text is uneven or irregular
  - Scanning is sloppy
- **Rarely (?) useful** for
  - Non-Roman texts (Cyrillic, Hindi, Mandarin, et al.)
  - Handwriting



Optical Recognition for Music

- Graphical imperfections in musical sources
- Layered contexts
- Output formats
  - MIDI
  - Other
- Evaluation techniques
- Diversity of musical textures and styles

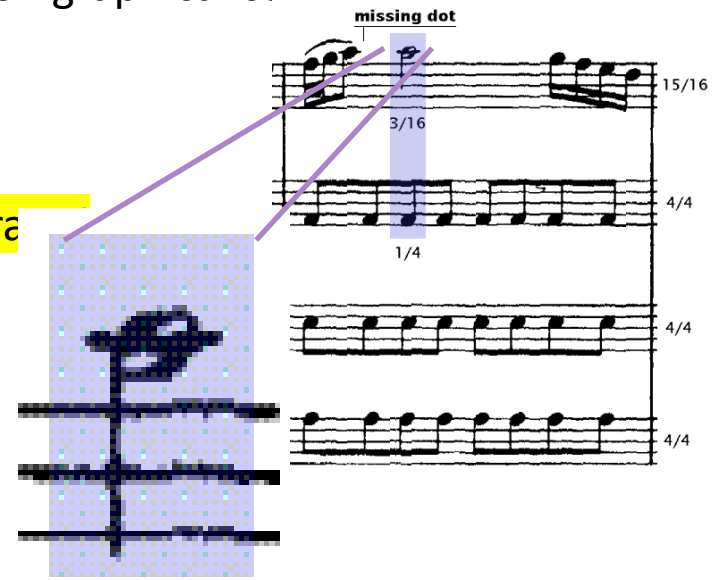
Optical Music Recognition (OMR)

History of efforts from c. 1968
- CCARH survey in 1993-4: 37 projects, 7 responses

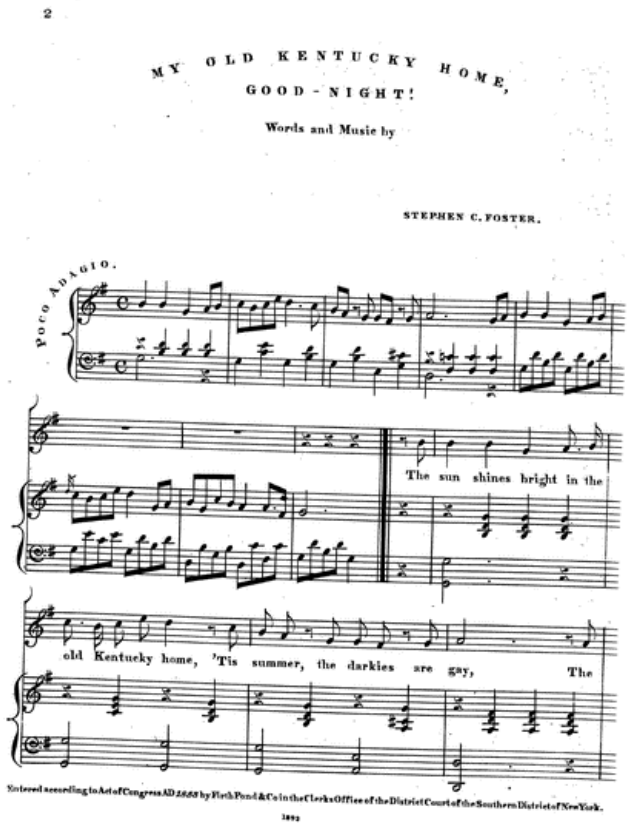
Why is optical recognition **difficult**?
- **Semantic meaning** of many objects depends on graphical context more than shape

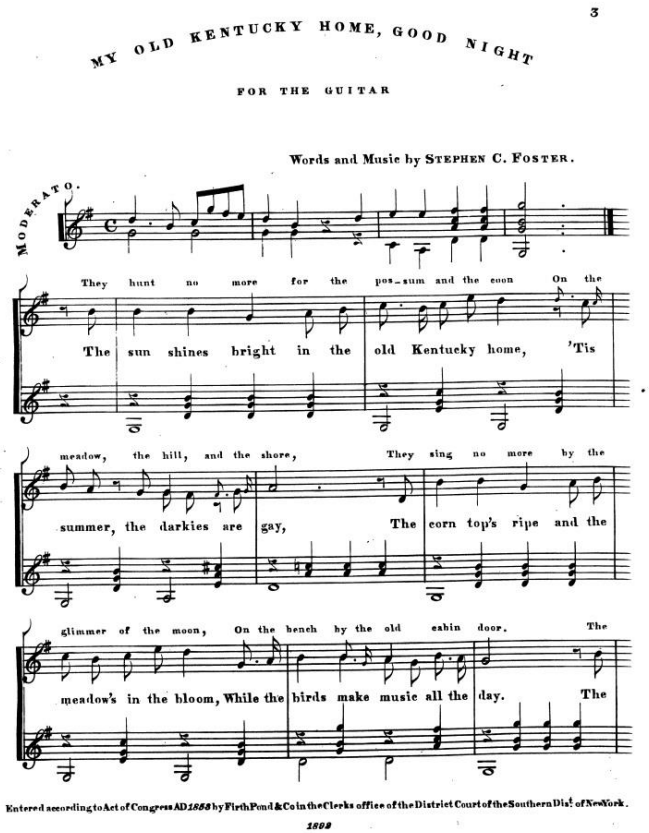
Sources and their **legibility**:
- *Manuscripts*: very irregular
- *Out-of-copyright prints*: images often deteriora
- ***In-copyright** prints*: not legal to copy
- **Errors** in source

Biggest problems for OMR developers
- **Superimposition** of objects in 2D image
- Constraints imposed by **output**



Basic problems in optical data acquisition





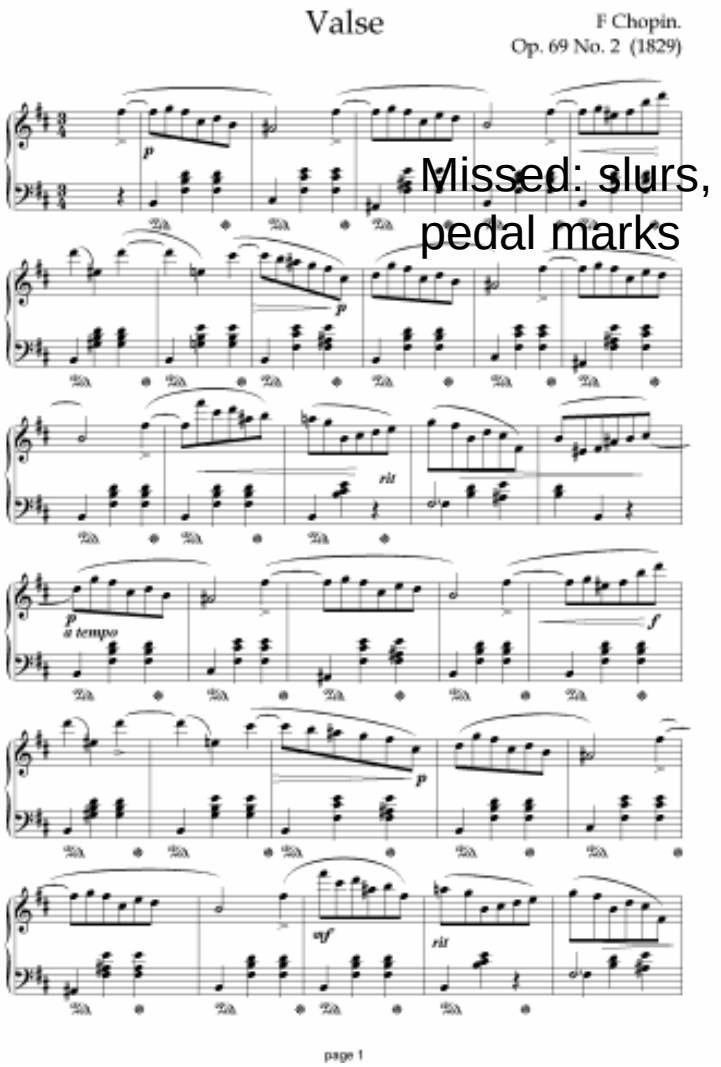
How does OMR work?

Separation of HORIZONTAL lines from other matter

Isolation of objects

Recognition of objects

Export to a format for
- storage
- printing
- sound
- data interchange



Why are good results elusive?

Problems of image quality:

- Ideally
  - Staff lines are straight
  - Spacing is uniform
  - The scanned material is clean (unspotted)
  - Slurs are symmetrical
  - Beams are parallel
  - All lines are unbroken
- Reality is different!

Problems of graphical context

Unread symbols affect interpretation of **pitch**

- Key signatures
- Octave alterations

Symbols affect interpretation of **duration**

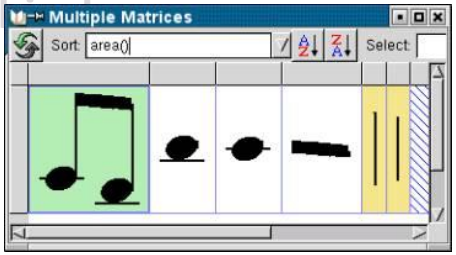
- Meter signatures
- Tempo indicators
- Fermatas

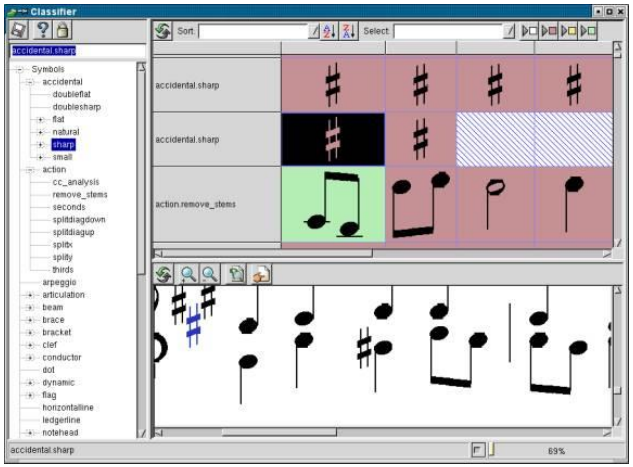
Symbols relating to dynamics or **technique**

- Dynamics marks
- Repetition of note-groups, of sections
- Instrumental technique

Other difficulties in CMN (common western notation)

Multiple configurations for same





Methods of evaluation and control

Musical **accuracy**?

Handicaps for **post-processing**

Controls for **input quality**

Comparison of **output formats**

Weighing speed against accuracy and usability

- **Work of Ichiro Fujinaga, McGill (c1988)**

Graphic flaws in conventionally typeset music

Surface imperfections

1. Visual surface problems

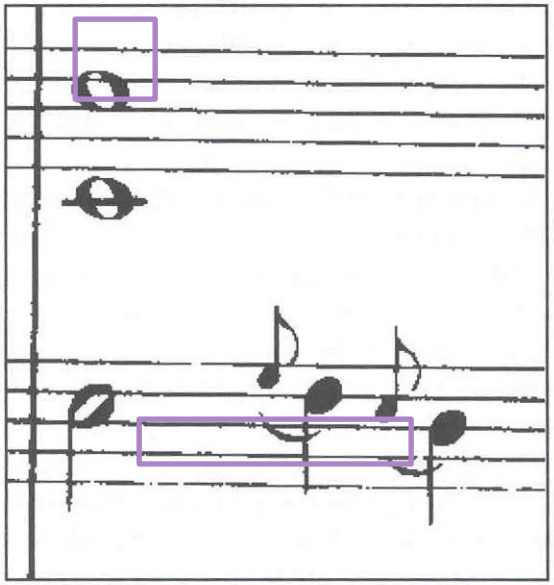


Figure 1. Surface imperfections: skewing and ambiguous positioning (uppermost note).

Surface imperfections

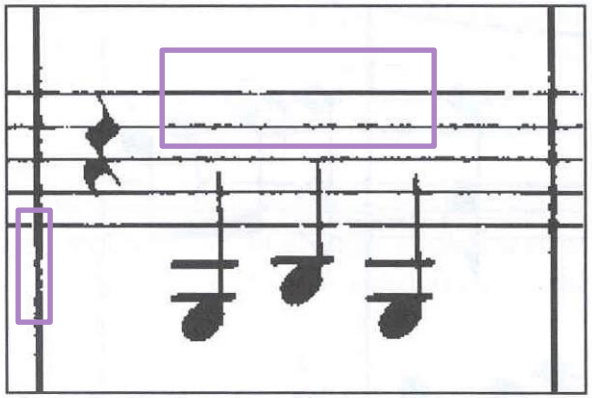


Figure 2. Surface imperfections: note the broken staff line at the top right and the variable width of both staff- and barlines.

Haydn: Symphony No. 1 (1895)
[out-of-copyright edition]

Close-up views (2)

Missing contextual information

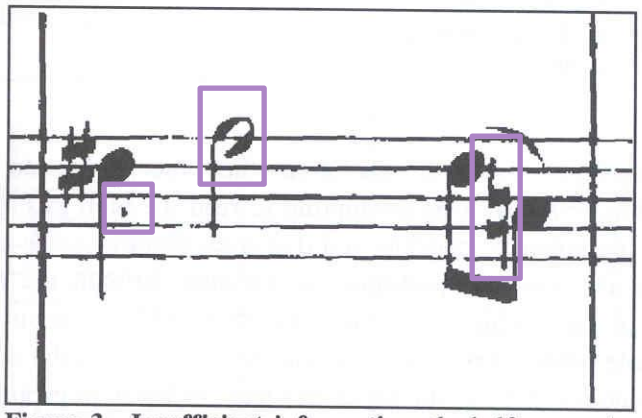


Figure 3. Insufficient information: the half note and the natural sign both lack closure. Compare the hypothetical white space in the half note with the actual white space bordered by the stem, the notehead, and the contingent flag in the tied octaves of Figure 4.

Graphic imperfections

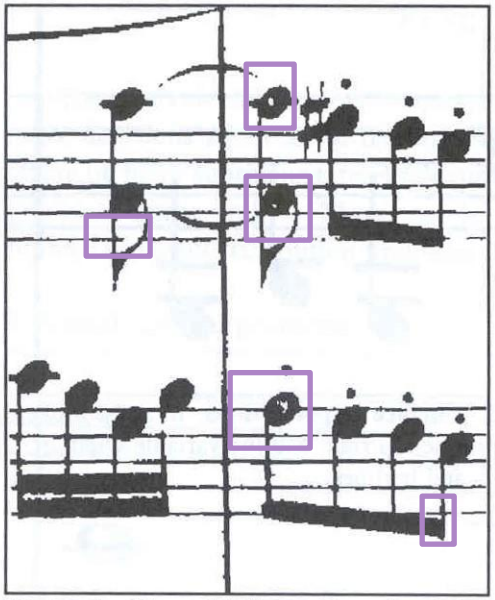


Figure 4. Flawed information: the eighth notes on the first beat are incompletely filled. Note the variable distance between the staccato dots and the notes to which they pertain.

Close-up views (3)

Dirt

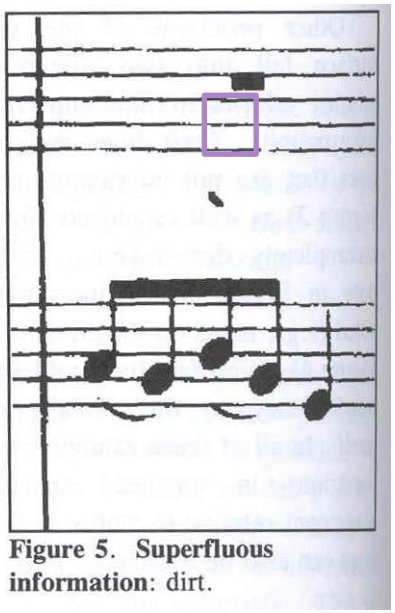


Figure 5. Superfluous information: dirt.

Variable appearance of equivalent objects

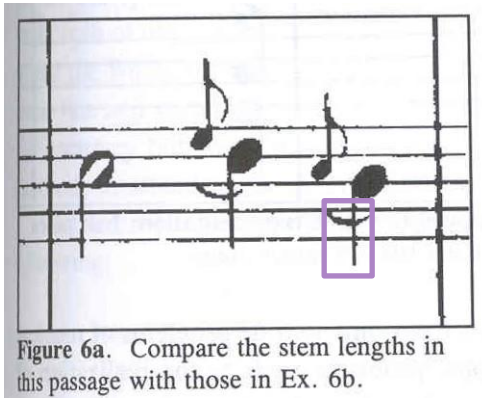


Figure 6a. Compare the stem lengths in this passage with those in Ex. 6b.

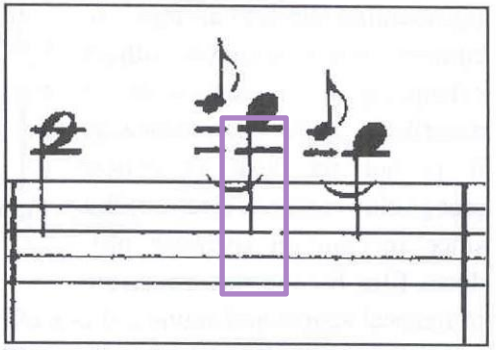


Figure 6b. Compare the stem lengths with those of Ex. 6a.

Close-up views (4)

Touching objects

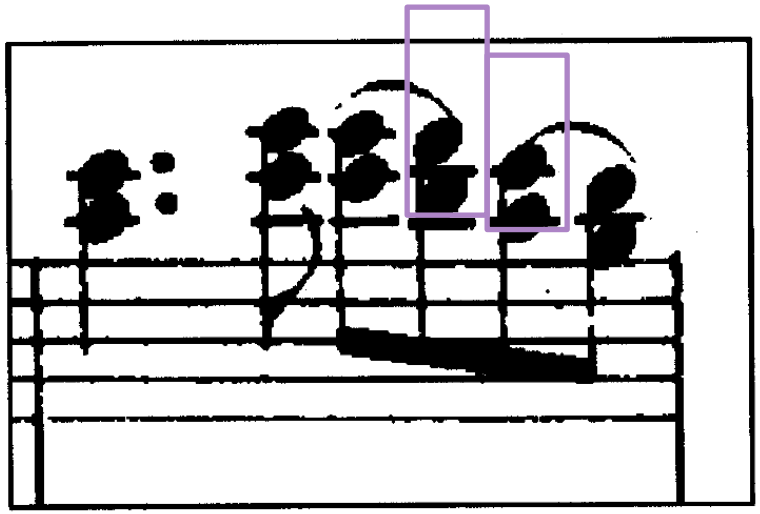


Figure 7. Superimposition: slurs touch noteheads. Note also that the flag of the first eighth note crosses a leger line.

Unconventional presentations

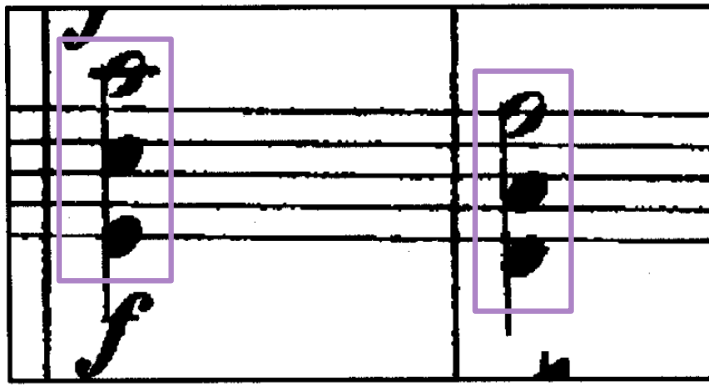
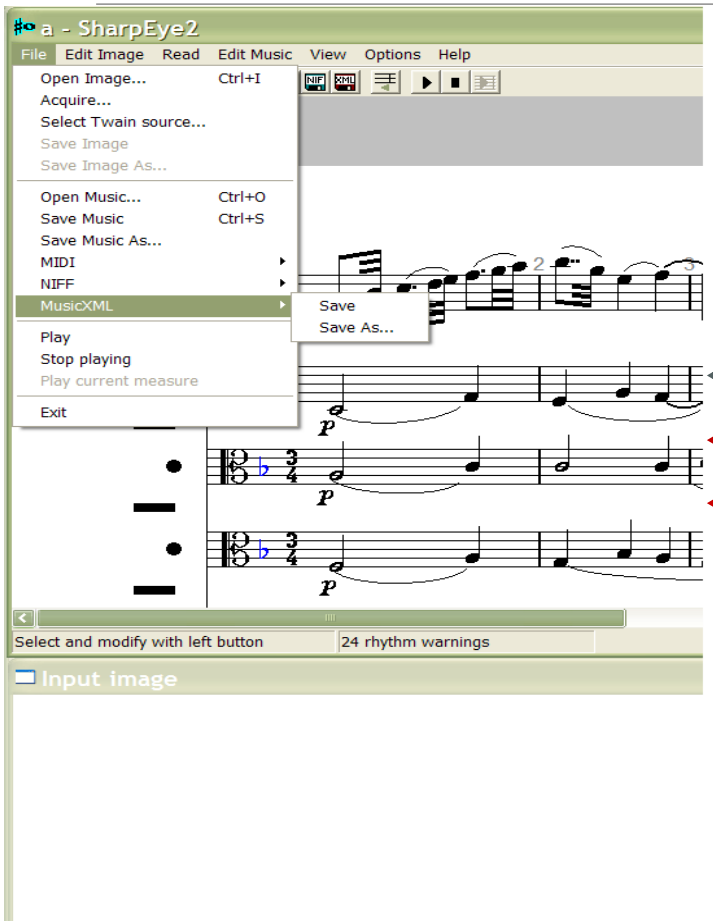


Figure 8. Issues in music representation:

SharpEye: File operations



Comes from Shetland Islands

Source code available

Exports to MusicXML

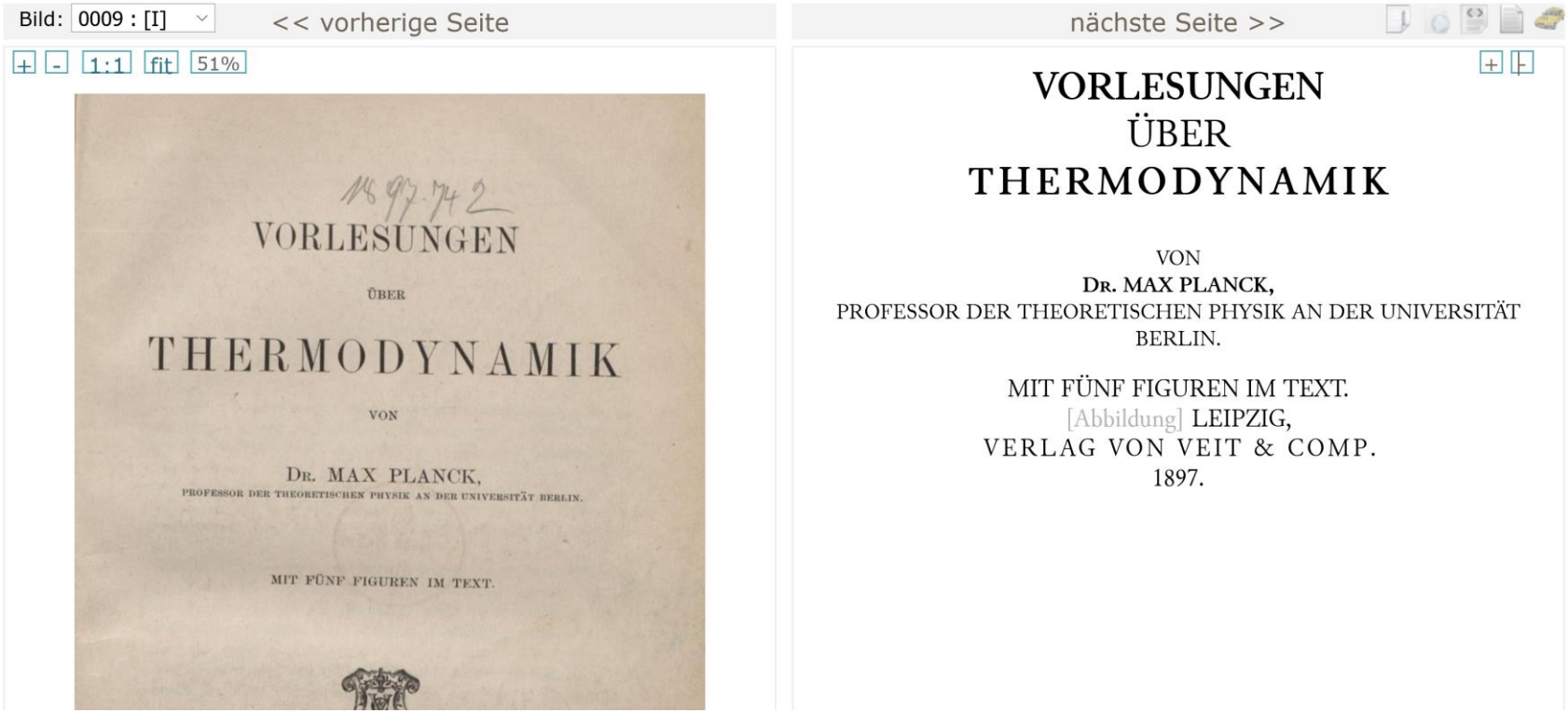
Four-step process

- **Capture** a page image
- **View** the auto-image
- **Correct** the image
- **Save/export** the result

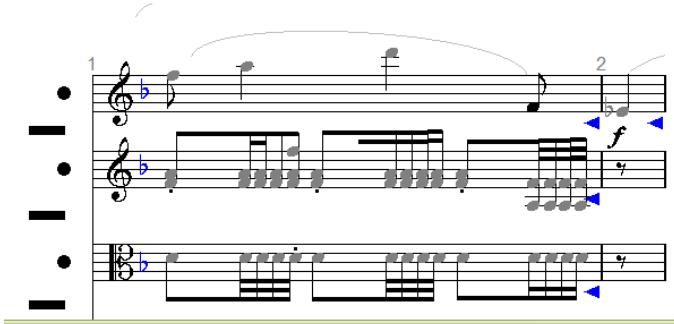
Vis-à-vis MuseData:

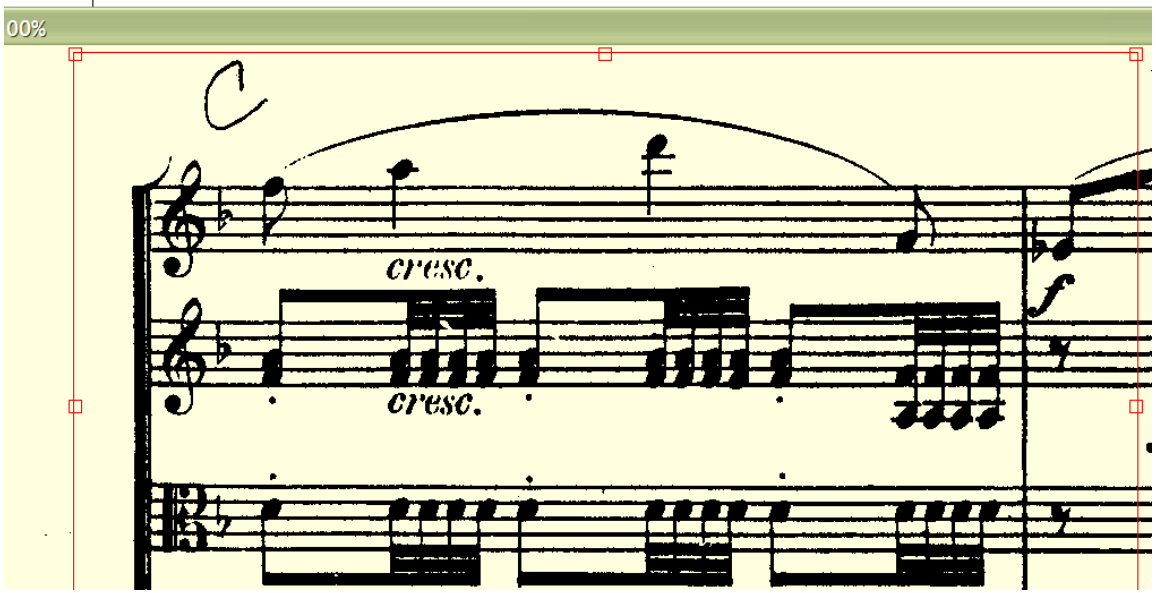
- SE: score-based
- MD: part-based

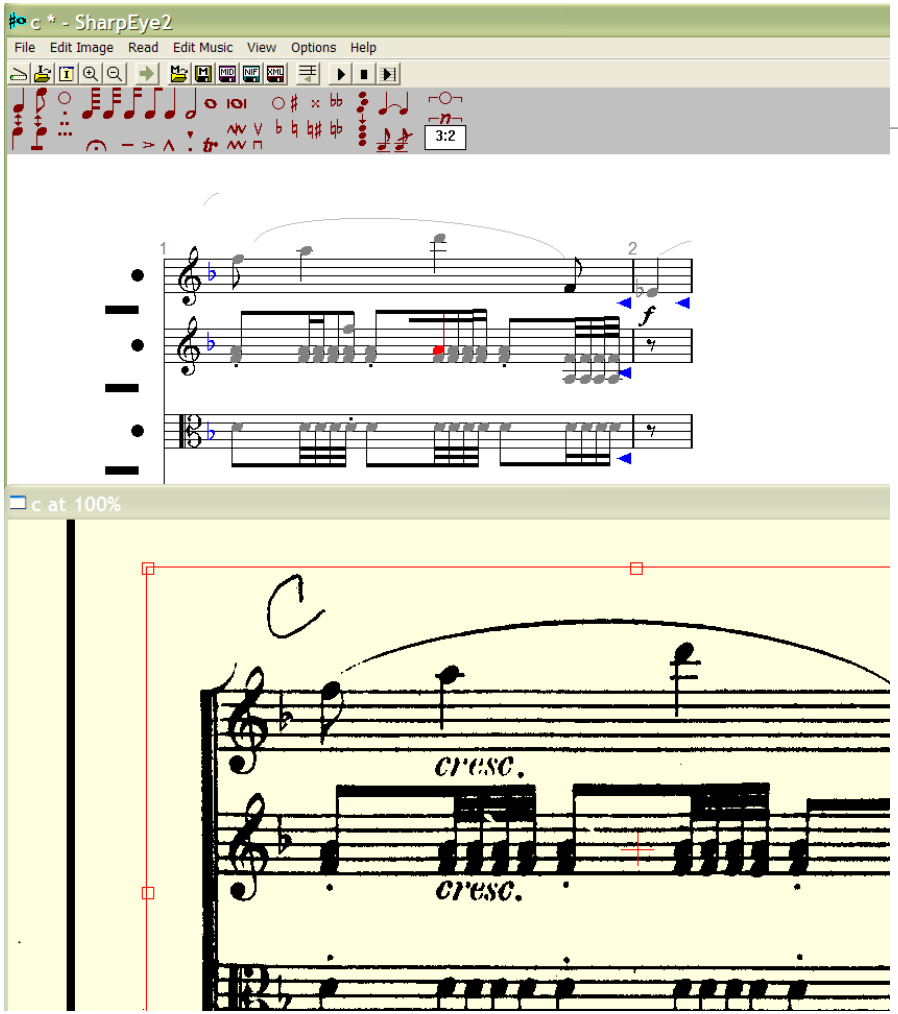
OTR benefits from side-by-side capture



Deutsche Textarchiv: Max Planck, *Readings on Thermodynamics* (1897)


1
2
f


00%
C
cresc.
cresc.
f



What SharpEye thought it saw

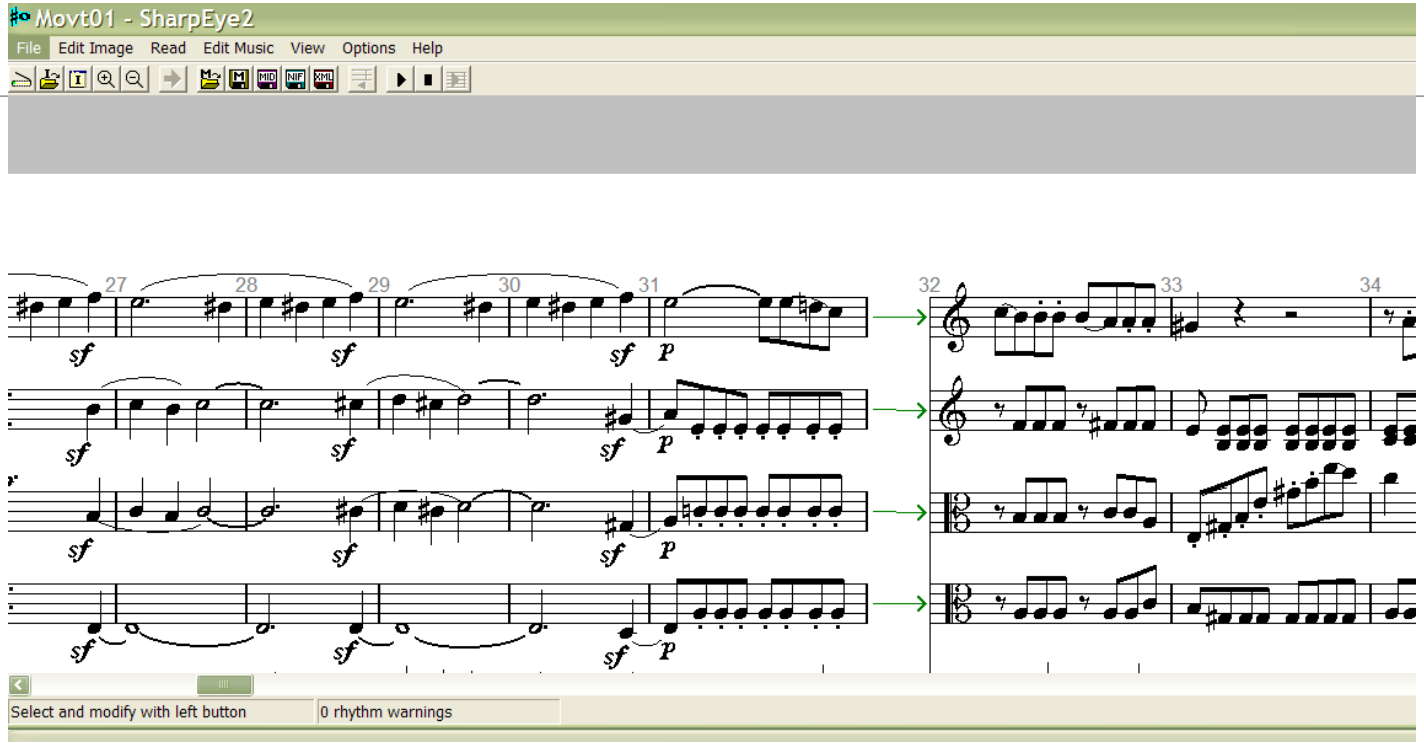
Edit mode:

- Captured image below
- Interpreted image above
- Live object in red
- Available symbols in red

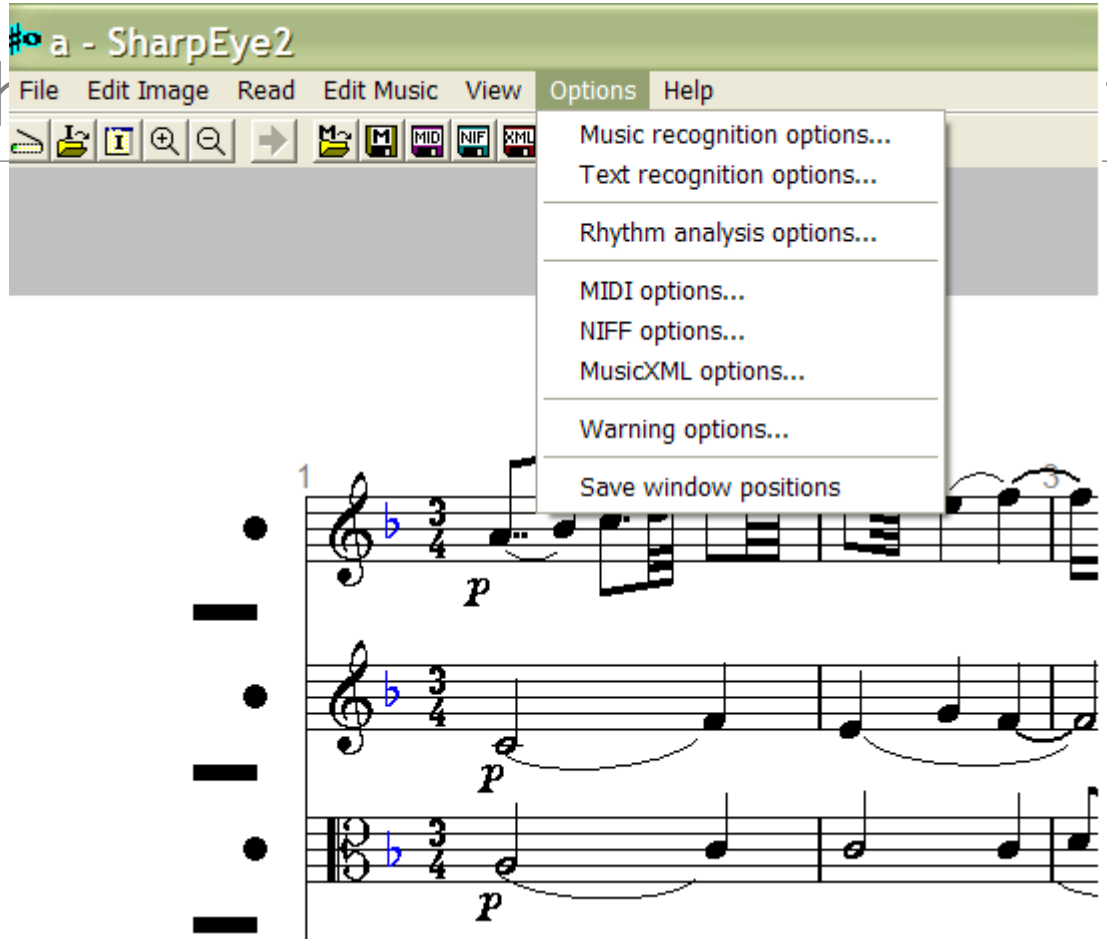
What SharpEye scanned

Step 1: Select a portion the score to edit

SharpEye: Scroll view



Sh tions



A word on NIFF:
Note Image FF
pixel-based
*c*1996 (MS)

Important questions about OMR software

http://www.wikicfp.com/cfp/servlet/event.showcfp?eventid=11836http://www.wikicfp.com/cfp/servlet/event.showcfp?eventid=1183633

What **output formats** are available?

- MIDI-level features only?
- Graphical position?
- Markup?

OMR forum: WoRMS

http://www.wikicfp.com/cfp/servlet/event.showcfp?eventid=118363

Music Reading Systems, incl. OMR

Jorge Calvo-Zaragosa et al., Alicante, ES

Innovations

- End-to-end document capture
  - https://archives.ismir.net/ismir2017/paper/000034.pdf

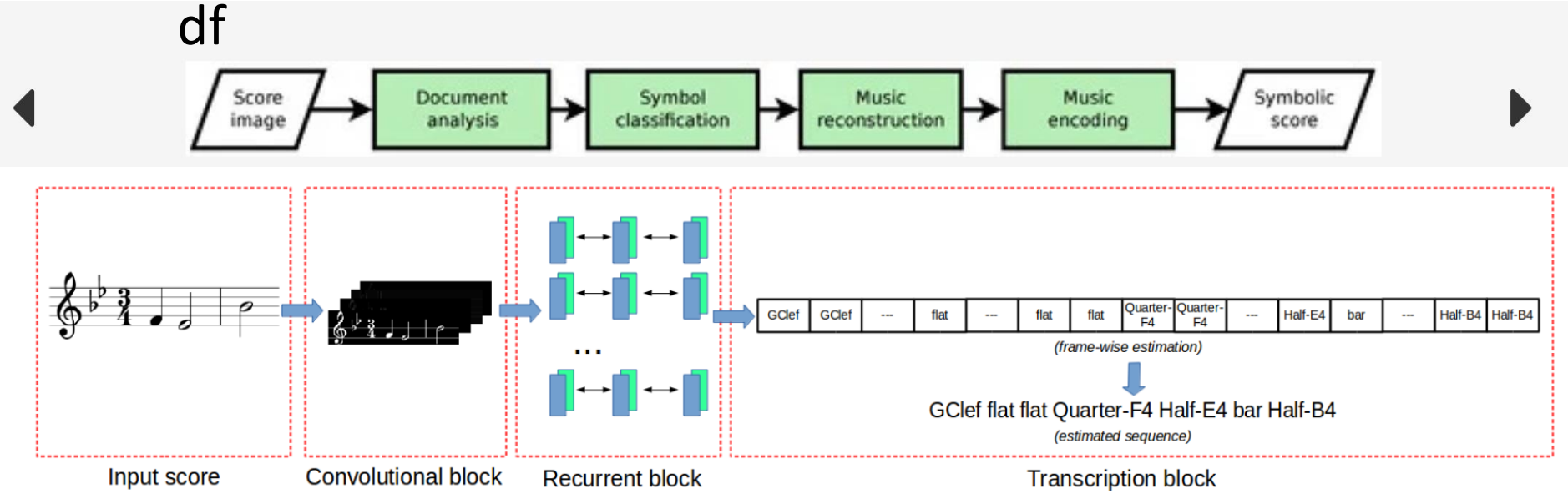
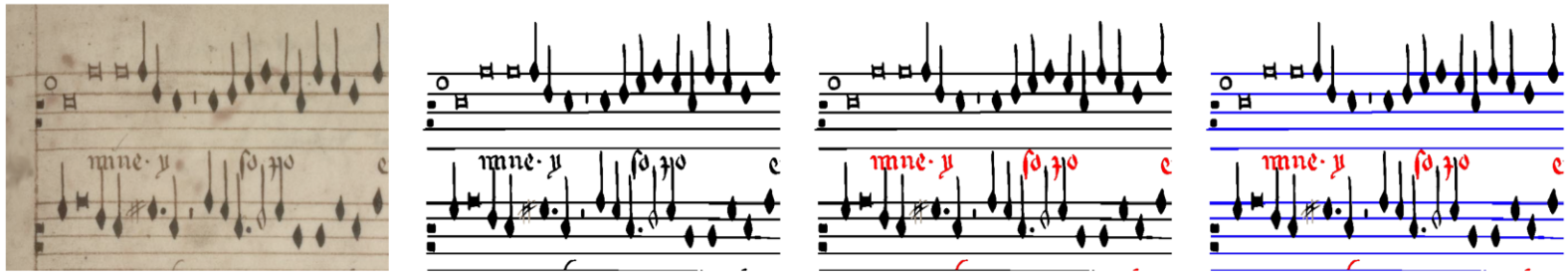


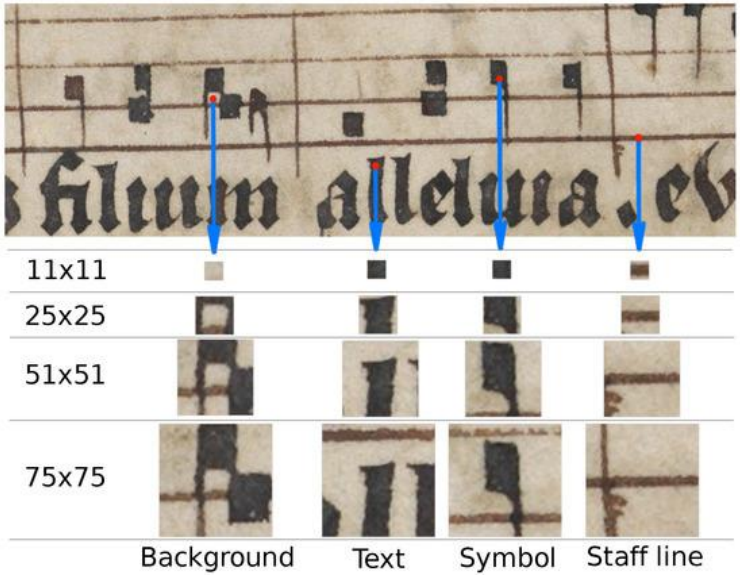
Figure 3. Conceptual scheme of the proposed approach. The input score is processed with a series of convolutional filters; the resulting features are then processed by the recurrent layers to model the temporal context of the piece; a frame-wise transcription using CTC is performed to obtain the estimation in an end-to-end fashion.

Graphical-musical categories

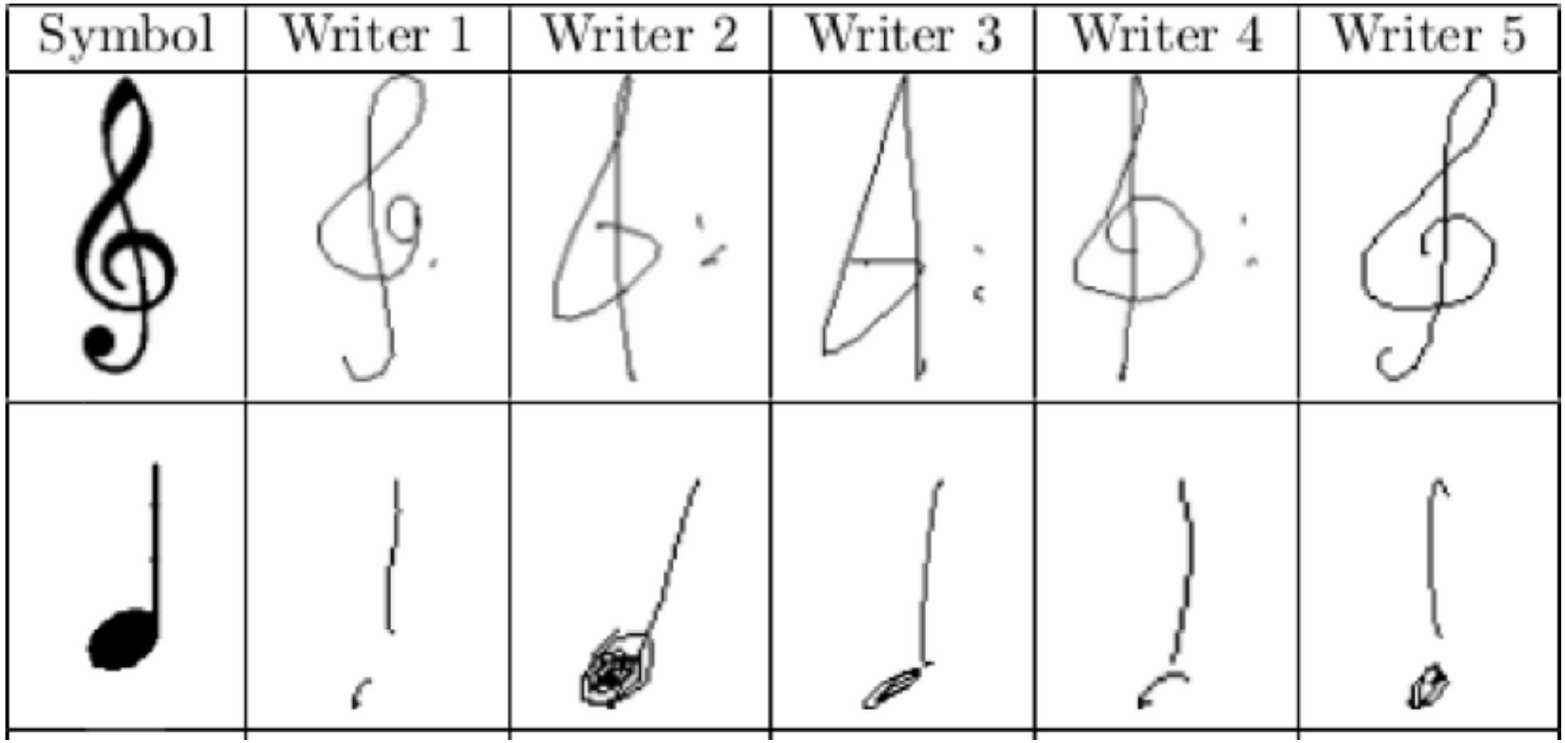
Pixel-level foreground/background differentiation



(a) Input image (b) Binarization (c) Text detection (d) Staff-line detection



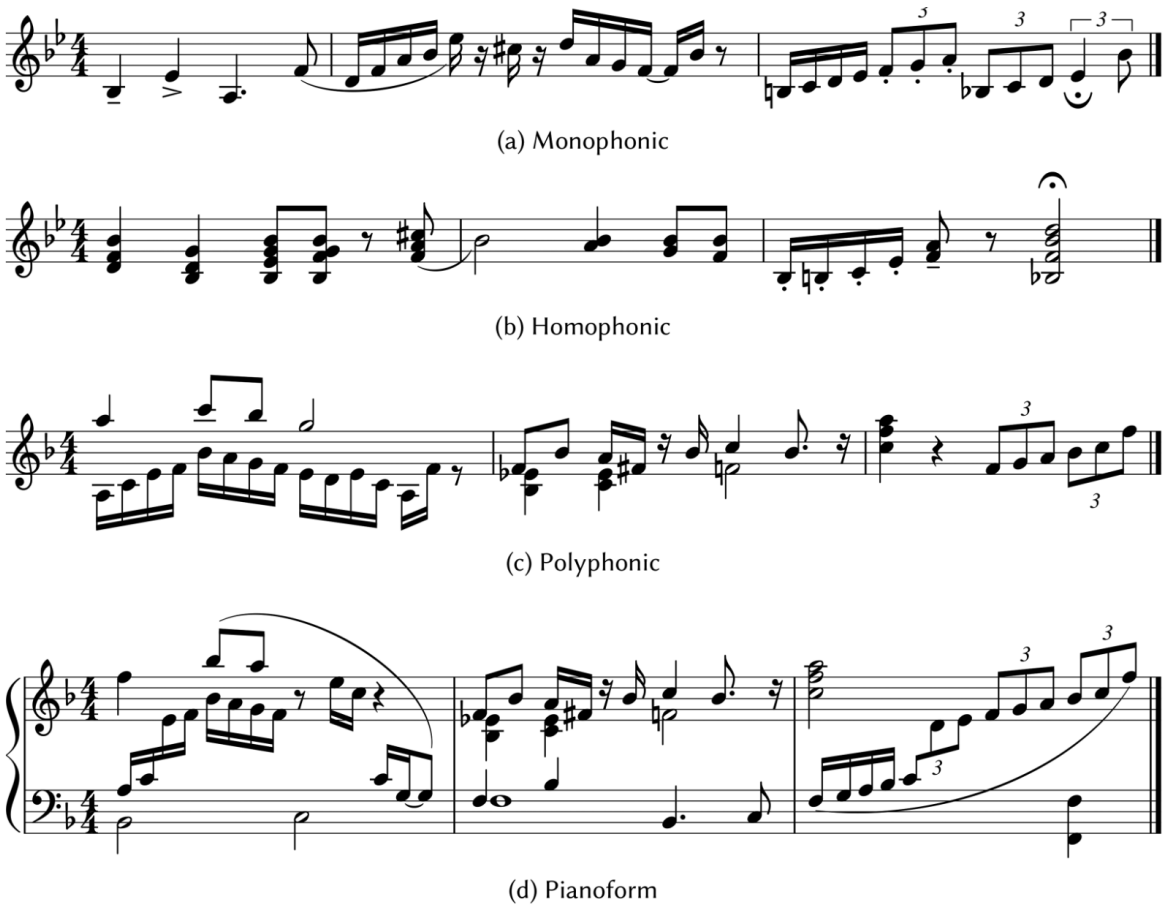
HOMUS dataset for hand-written music



HMM vs **NN**

IEEE Explore 2014

OMR Review article:
Levels of structural complexity (2020)



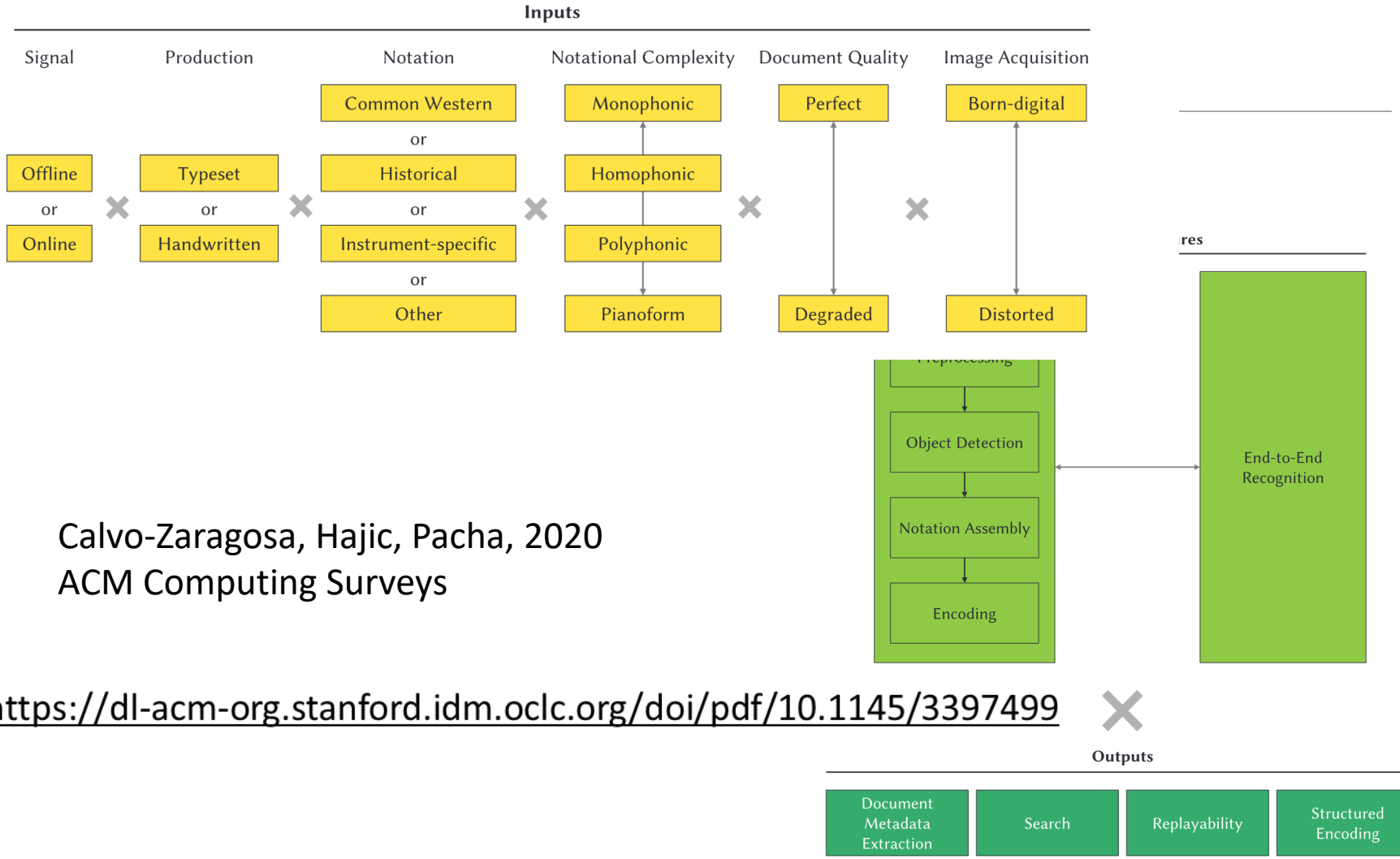
(a) Monophonic

(b) Homophonic

(c) Polyphonic

(d) Pianoform

Inputs and Outputs



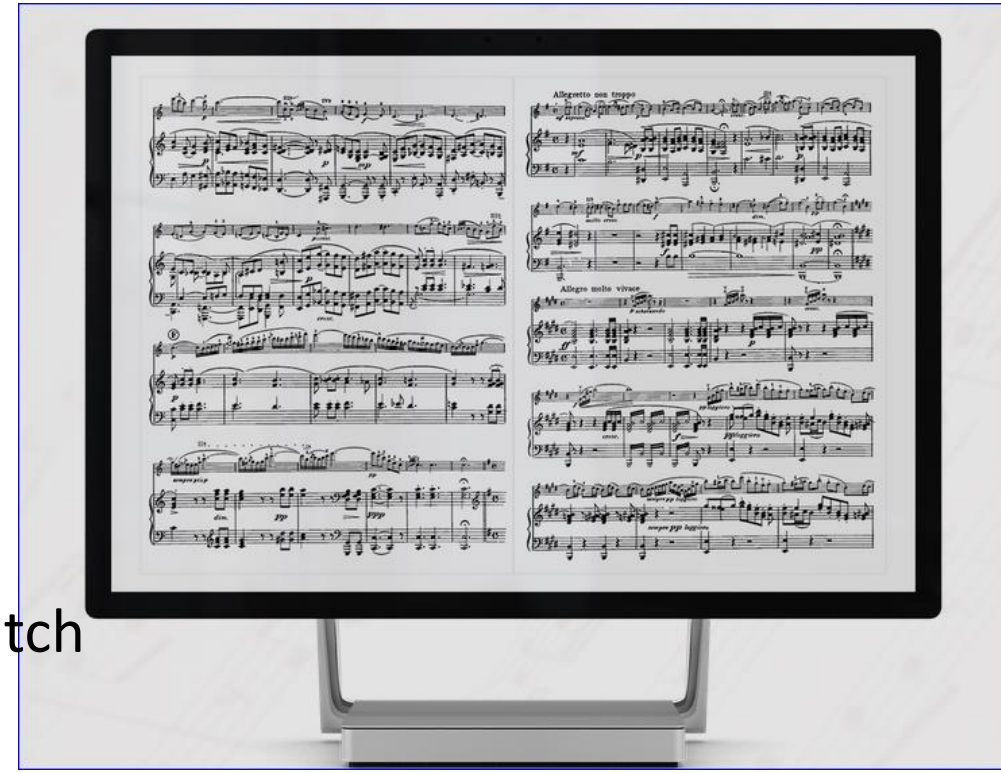
Calvo-Zaragosa, Hajic, Pacha, 2020
ACM Computing Surveys

https://dl-acm-org.stanford.idm.oclc.org/doi/pdf/10.1145/3397499

PlayScore 2 for mobile devices; Maestria

PlayScore 2 Sales pitch: https://www.playscore.co/

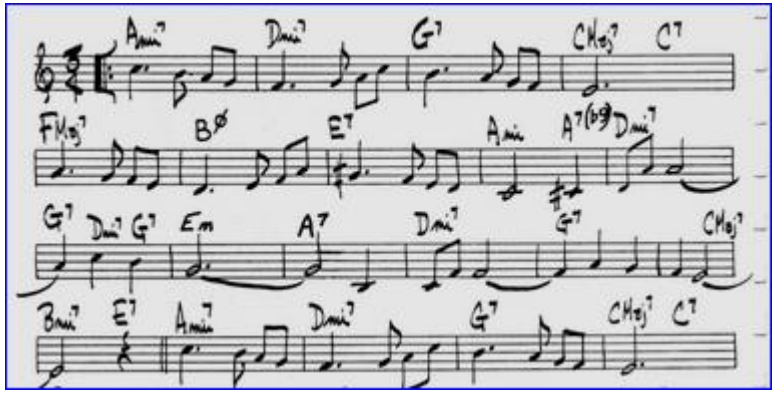
v. 2.8 in beta can currently be used for free.

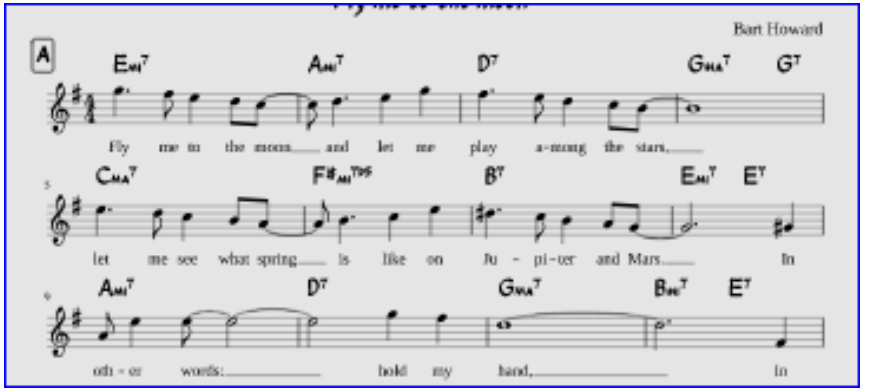


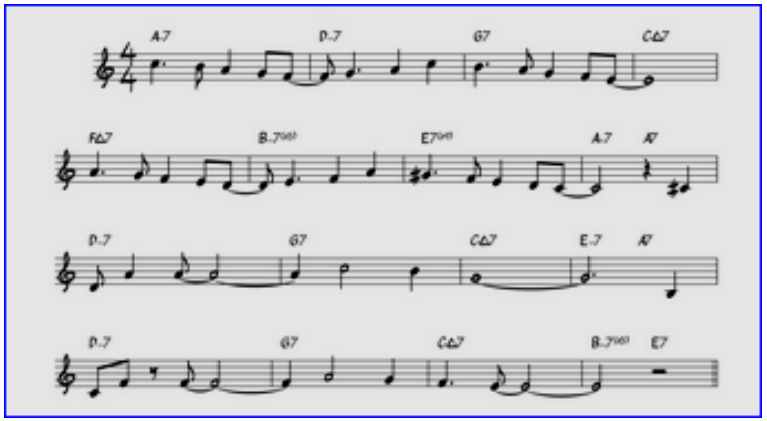
Maestria (from Newzik) prepub pitch

https://newzik.com/en/maestria/

Addendum: “Fly me to the moon”







Written score vs. performance

https://www.youtube.com/watch?v=HJuZUBJtWUo