
Optical Music Recognition and Data Import/Export

Music 253/ CS 275A

Eleanor Selfridge-Field

Optical Music Recognition

History of efforts from c. 1968

- CCARH survey in 1993-4: 37 projects, 7 responses

Why is optical recognition **difficult**?

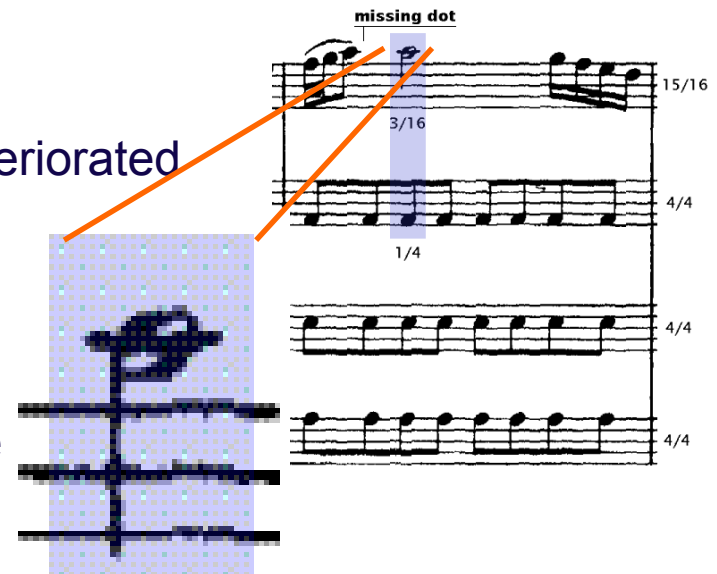
- **Semantic meaning** of many objects depends on graphical context more than shape

Sources and their **legibility**:

- *Manuscripts*: very irregular
- *Out-of-copyright prints*: images often deteriorated
- *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 formats**



Basic problems in optical data acquisition

- Image is **crooked**
- Elements of **layout unconventional**

2

MY OLD KENTUCKY HOME,
GOOD-NIGHT!

Words and Music by
STEPHEN C. FOSTER.

VOCE AD AGIO.

The sun shines bright in the
old Kentucky home, 'Tis summer, the darkies are gay, The

Entered according to Act of Congress AD 1868 by Firth, Pond & Co in the Clerk's Office of the District Court of the Southern District of New York.
1868

3

MY OLD KENTUCKY HOME, GOOD NIGHT
FOR THE GUITAR

Words and Music by STEPHEN C. FOSTER.

MODERATO.

They hunt no more for the pos-sum and the coon On the
The sun shines bright in the old Kentucky home, 'Tis
meadow, the hill, and the shore, They sing no more by the
summer, the darkies are gay, The corn top's ripe and the
glimmer of the moon, On the bench by the old cabin door. The
meadow's in the bloom, While the birds make music all the day. The

Entered according to Act of Congress AD 1868 by Firth, Pond & Co in the Clerk's Office of the District Court of the Southern District of New York.
1868

How does OMR work?

- **Separation** of lines and other matter
- **Isolation** of objects
- **Recognition** of objects
- **Export to a format** for
 - storage
 - printing
 - sound
 - data interchange

Valse
F Chopin.
Op. 69 No. 2 (1829)

The image displays a musical score for Frédéric Chopin's Valse Op. 69 No. 2. The score is presented in six systems, each consisting of a treble and bass staff. The key signature is one sharp (F#) and the time signature is 3/4. The notation includes various musical symbols such as notes, rests, and dynamic markings like *p* (piano), *rit* (ritardando), *f* (forte), and *sf* (sforzando). The score is a page from a larger document, as indicated by the 'page 1' label at the bottom.

page 1

Why is OMR difficult?

Problems of **image quality**:

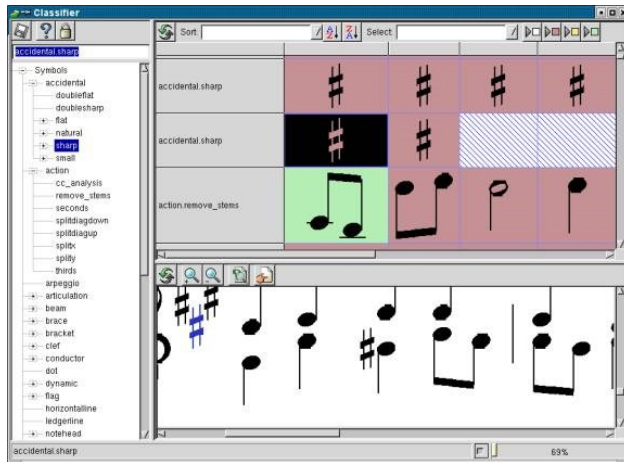
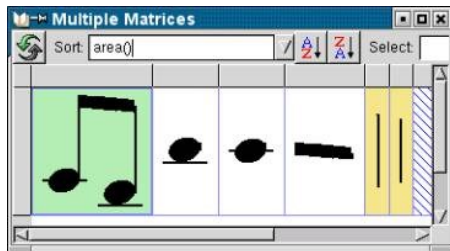
- **Ideally**
 - Staff lines are straight
 - Spacing is uniform
 - The scanned material is clean (unspotted)
 - Slurs are symmetrical
 - Beams are parallel
 - All lines are unbroken
- **Reality** is different!

Problems of **graphical context**

- Symbols affecting interpretation of **pitch**
 - Key signatures
 - Octave alterations 15^{ma} 8^{va}
- Symbols affect interpretation of **duration**
 - Meter signatures
 - Tempo indicators
 - Fermatas \frown
- Symbols relating to dynamics or **technique**
 - Dynamics marks pp sfz ff
 - Repetition of note-groups \times , of sections \oplus $\%$
 - Instrumental technique Ped. v \blacksquare

More difficulties

Multiple configurations for same objects



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**

	Input	Capture format	Post- Processing
Carter	00:20	SCORE	9:20
CCARH	2:30 + 7:05	MuseData	00:15

Samples from Library of Congress site

Random material from loc.gov

SCHUBERT'S, LA SERENADE
Arranged as a VOCAL DUETT by
LEOPOLD HOFFMAN.

pp

Murmur low and sweet,
Through the leaves the night-winds moving, Murmur low and sweet,
To thy chamber window re-ving, Love hath led my feet,
O! Love hath led my feet.

LA PALOMA.
THE DOVE.
English adaptation by LOUIS C. ELSON. **TRADIER.**
Allegretto.

1. I think . . . of the morn when I sailed a-way from thee. I
2. Ni na . . . when to shore re-turn-ing thy smile I see. My
3. At last . . . on the shore we're land-ing, and grief has flown. And
4. En fin . . . nous touchons la ter-re, mon Dieu, mor-tel.

said . . . pray to God for me pray to God for me. I
dis- . . . Pri-er Dieu, pri-er Dieu pour votre sa-fa- . . . I
love . . . for that time is yearning to com-fort thee. And
-rai . . . mais non plus en l'air, ma mè-re et toi. . . a
there . . . is my moth-er standing, but why a lone? Why
-ja . . . ja-perçois ma mè-re, ma mè-re seule. . . c'est . . . Pourquoi.

Samples 2:

More random material from loc.gov

WALZER.

1.

In Ländler-Tempo.

Johannes Brahms. Op. 52.

PIANO.

p *p dolce*

Revised according to first of Congress, in the year 1894, by G. Schirmer is the office librarian of Congress, at Washington, D. C.

1894

and the Office of the Librarian of Congress at Washington, D. C.

2.

Akademische Fest-Ouverture.

Secondo.

Johannes Brahms, Op. 80.

Allegro.

PIANO.

pp sempre a tutta voce

molto p

Revised according to first of Congress, in the year 1894, by G. Schirmer is the office librarian of Congress, at Washington, D. C.

1894

and the Office of the Librarian of Congress at Washington, D. C.

Close-up views of conventionally typeset music

Surface imperfections

1. Visual surface problems



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

Surface imperfections

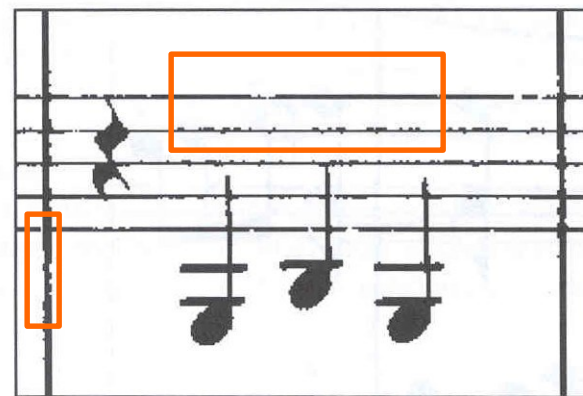


Figure 2. Surface imperfections: note the broken staff line at the top right and the variable width of both staff- and barlines.

Close-up views (2)

Missing contextual information

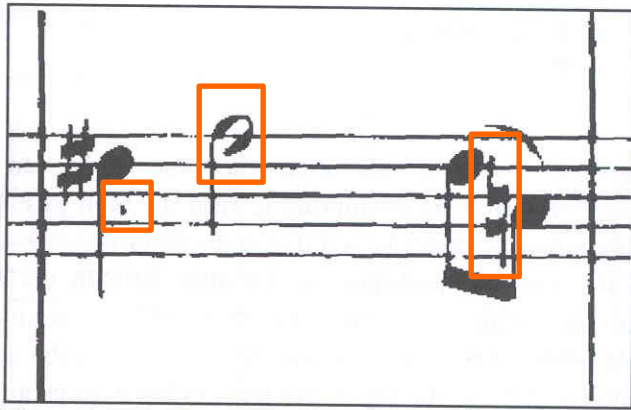


Figure 3. Insufficient information: the half note and the natural sign both lack closure. Compare the hypothetical white space in the half note with the actual white space bordered by the stem, the notehead, and the contingent flag in the tied octaves of Figure 4.

Graphic imperfections

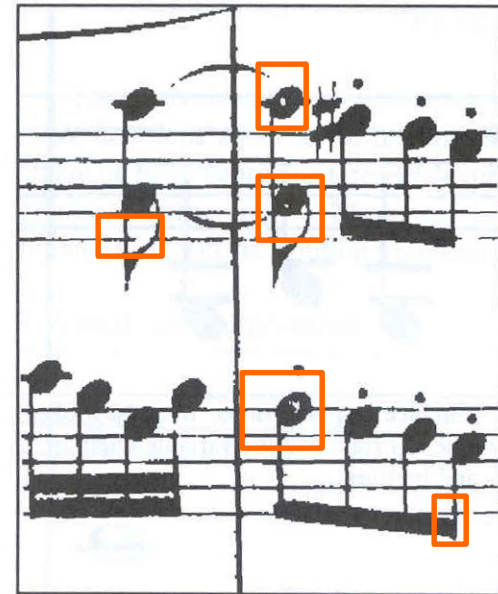


Figure 4. Flawed information: the eighth notes on the first beat are incompletely filled. Note the variable distance between the staccato dots and the notes to which they pertain.

Close-up views (3)

Dirt

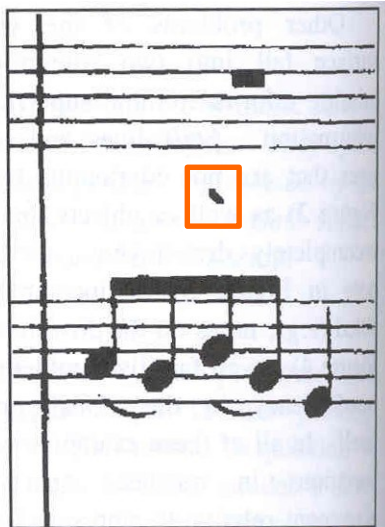


Figure 5. Superfluous information: dirt.

Variable appearance of equivalent objects

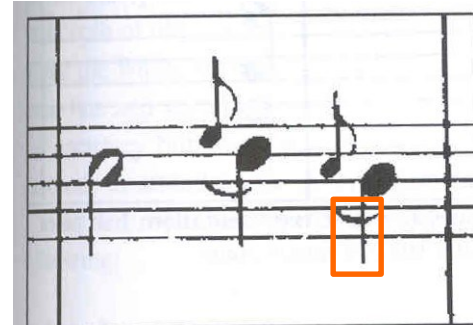


Figure 6a. Compare the stem lengths in this passage with those in Ex. 6b.

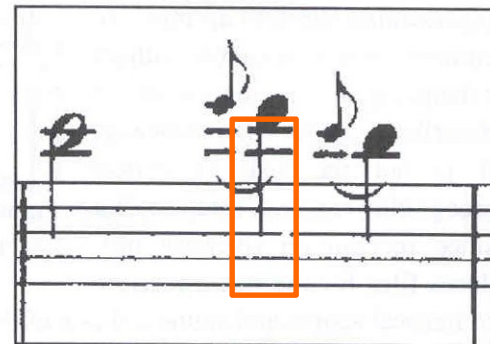


Figure 6b. Compare the stem lengths with those of Ex. 6a.

Close-up views (4)

Touching objects

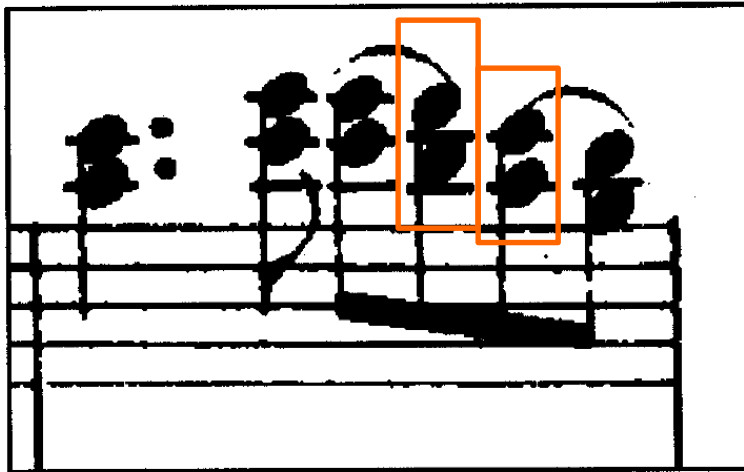


Figure 7. Superimposition: slurs touch noteheads. Note also that the flag of the first eighth note crosses a leger line.

Unconventional presentations

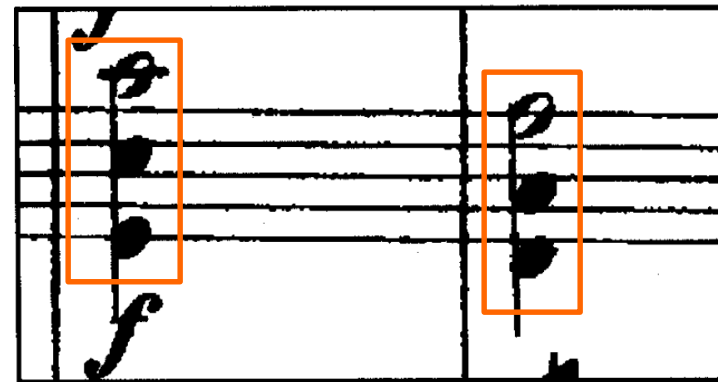
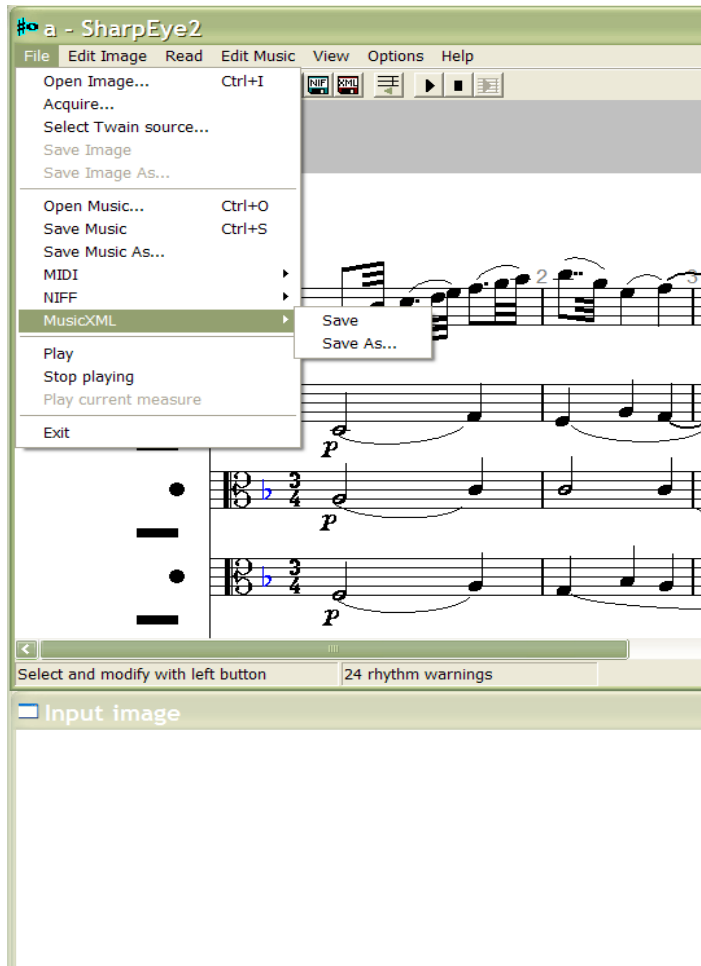


Figure 8. Issues in music representation:

SharpEye: File operations



Comes from Shetland Islands

Source code available

Exports to MusicXML

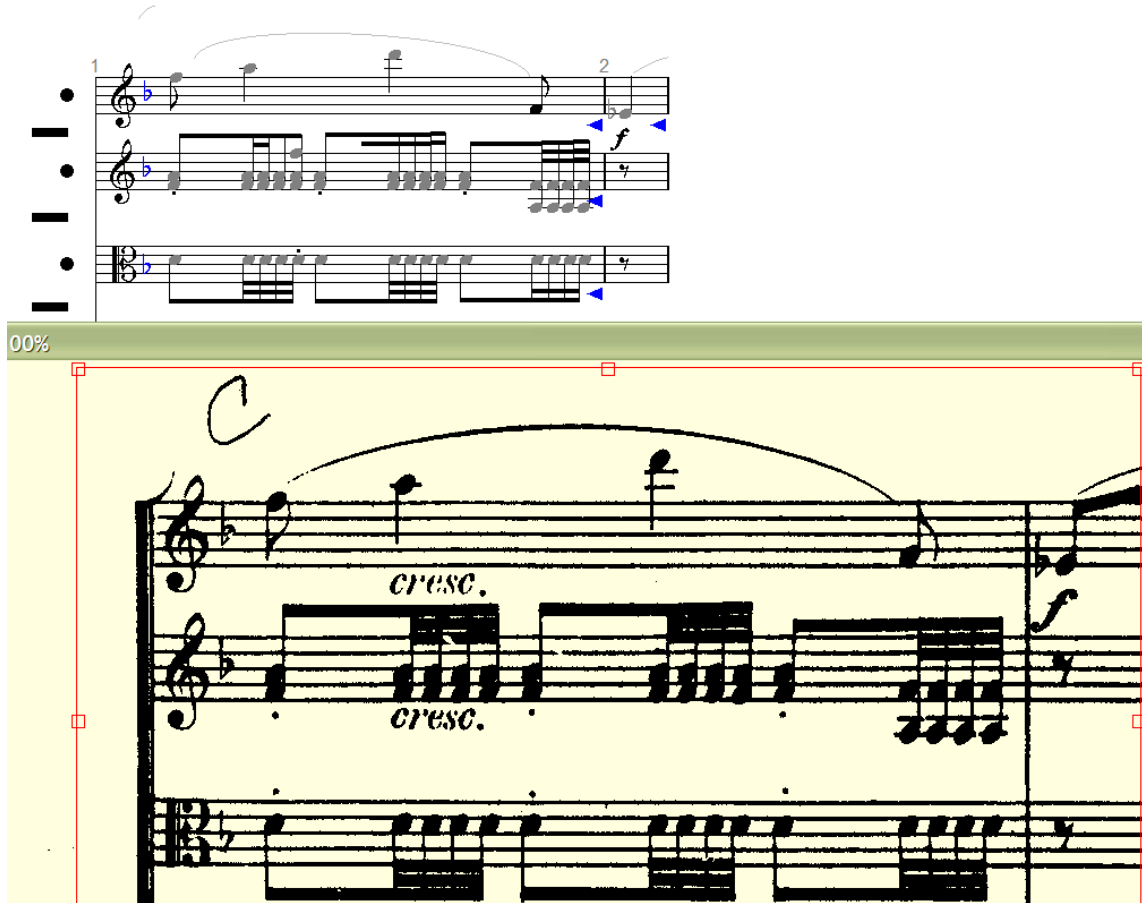
Four-step process

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

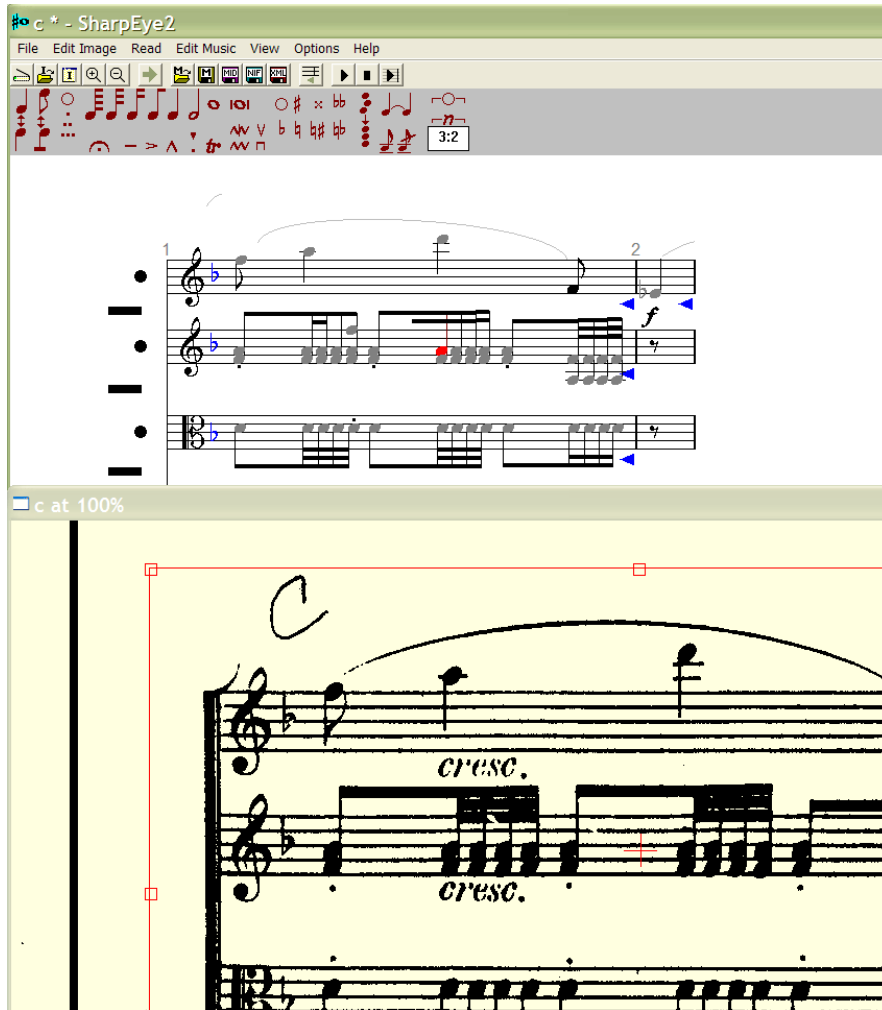
Vis-à-vis MuseData:

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

SharpEye: Raw Capture



SharpEye: Correcting the interpretation

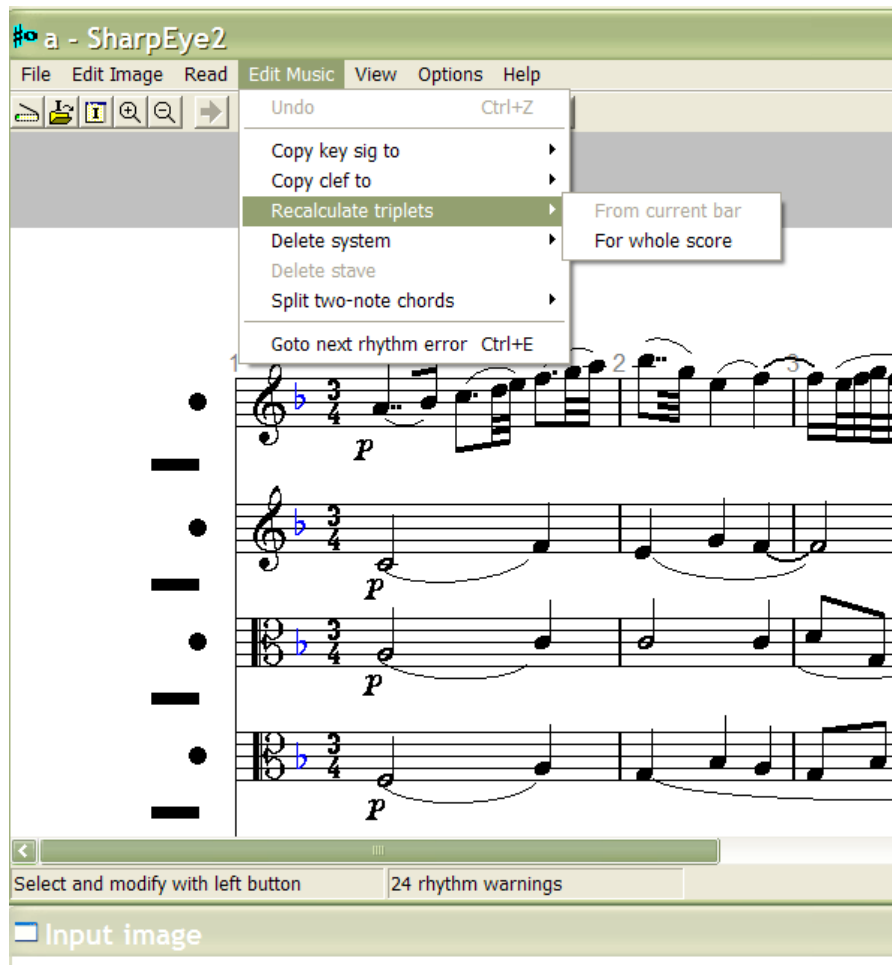


- Edit mode:
 - Captured image below
 - Interpreted image above
 - Live object in red
 - Available symbols in red

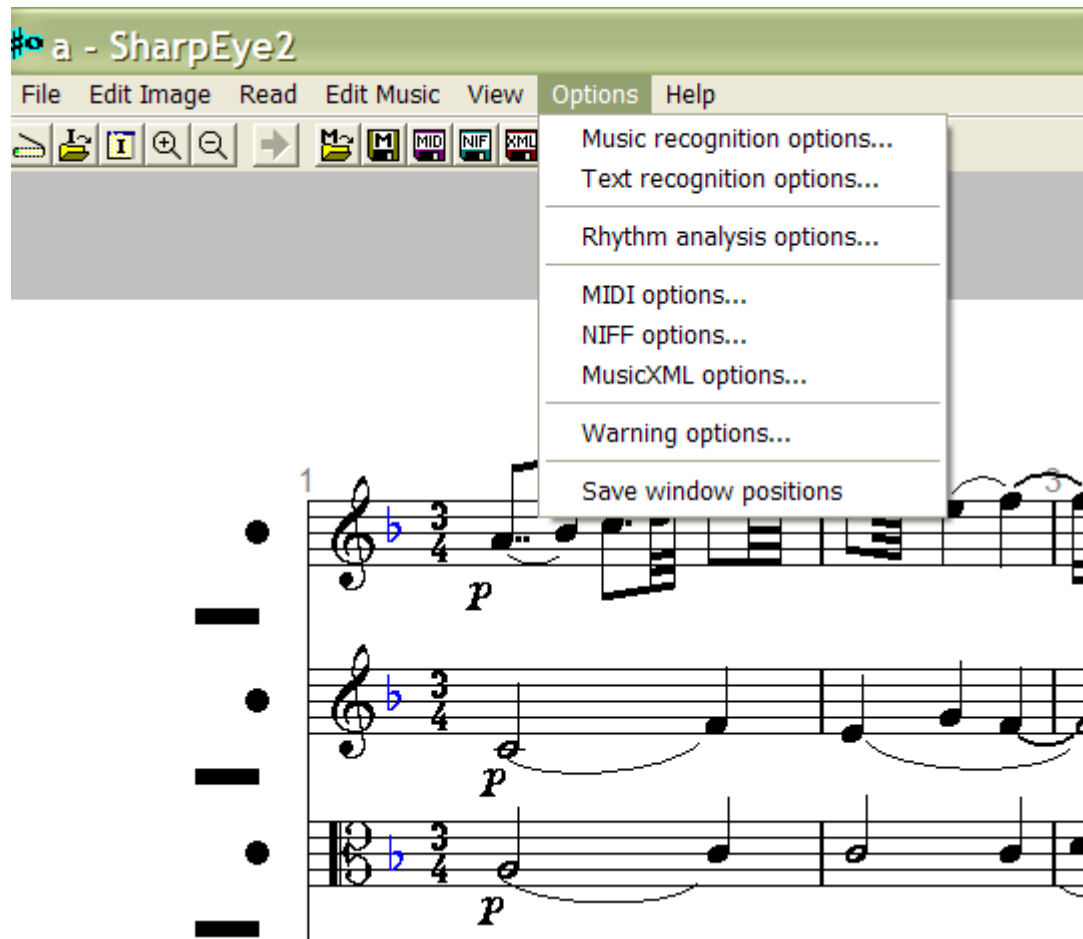
SharpEye: Scroll view

The screenshot displays the SharpEye2 software interface. At the top, a green title bar reads "Movt01 - SharpEye2". Below it is a menu bar with "File", "Edit Image", "Read", "Edit Music", "View", "Options", and "Help". A toolbar with various icons for editing and viewing is located below the menu bar. The main area shows a scroll view of musical notation. The notation is organized into four systems, each with four staves. The first system is numbered 27 to 31. The second system is numbered 32 to 34. The notation includes various musical symbols such as notes, rests, and dynamic markings like *sf* (sforzando) and *p* (piano). The interface also features a status bar at the bottom with a left arrow icon, the text "Select and modify with left button", and a counter showing "0 rhythm warnings".

SharpEye: System edits

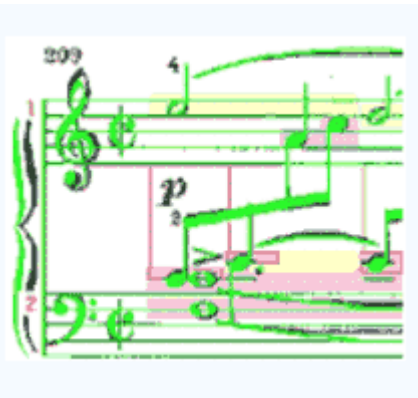


SharpEye: Data-interchange options

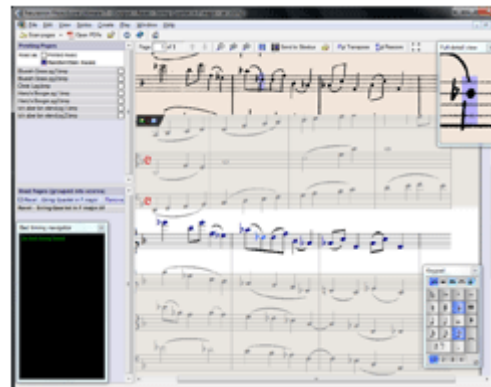


Other OMR Software

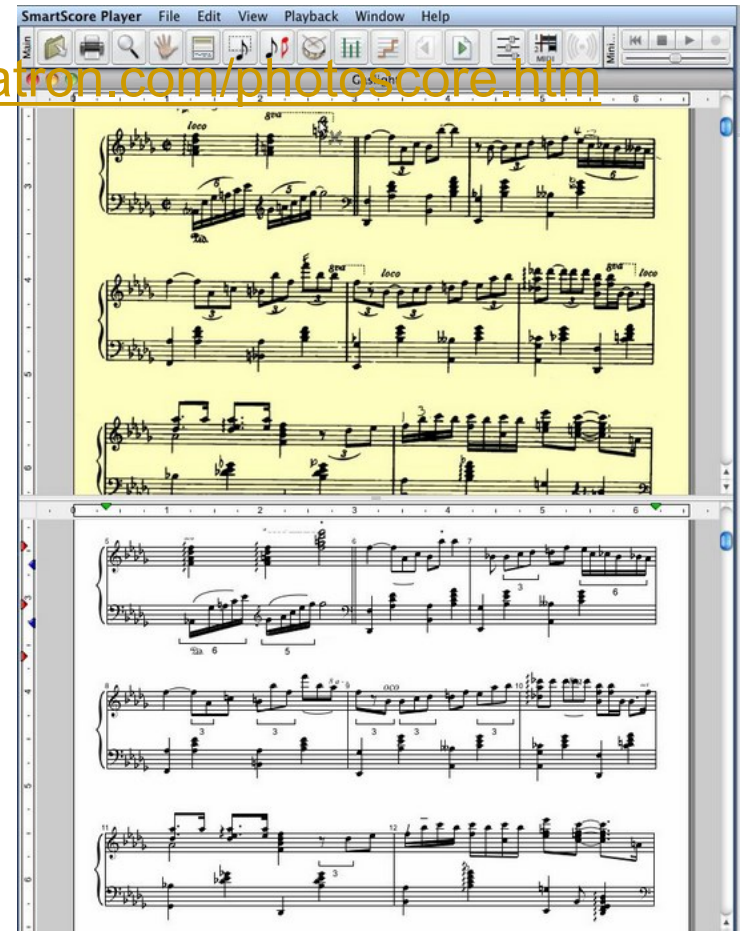
- Capella-scan: <http://www.capella.de/us/index.cfm/products/capella-scan/info-capella-scan/> [capella]
- Neuratron PhotoScore: <http://www.neuratron.com/photoscore.htm> [Sibelius]
- SmartScore: <http://www.musitek.com/>



CapellaScan



PhotoScore



SmartScore

Important questions about OMR software

- What does “accuracy” mean?
 - Text recognition optimal error rate: 40/2000 chars
- What kinds of errors?
 - Global variables?
 - Local events?
 - Non-MIDI objects
- What output formats are available?
 - MIDI-level features only?
 - Graphical position?
 - Markup?