

A Retrospective
on
Optical Music Recognition Research

DDMAL



Ichiro Fujinaga

Music Technology Area, Schulich School of Music

McGill University

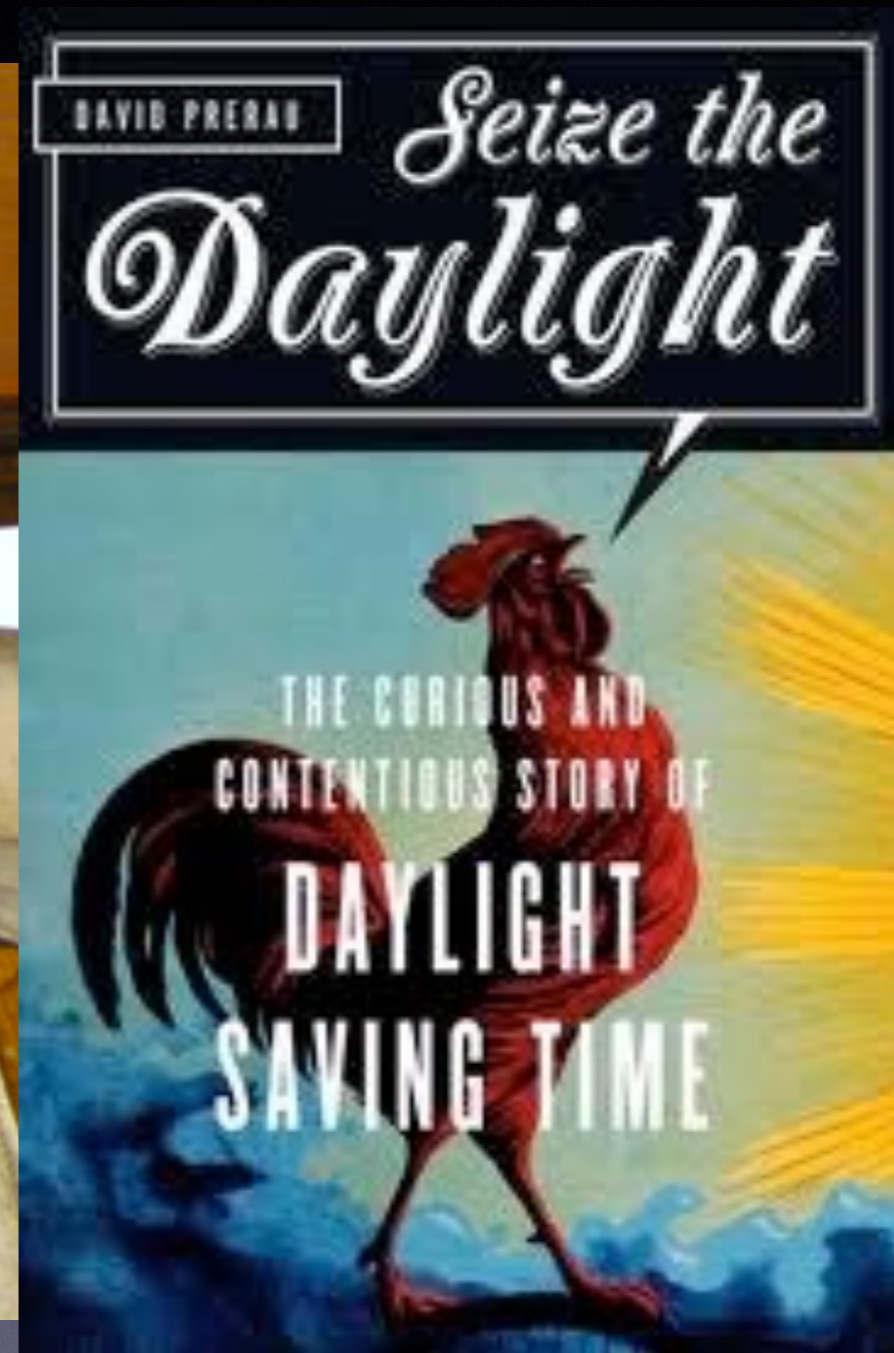
The Pioneers



Denis Bruslin AKA The Tool
David Prerau
Baker House, c. 1959

n of
MIT

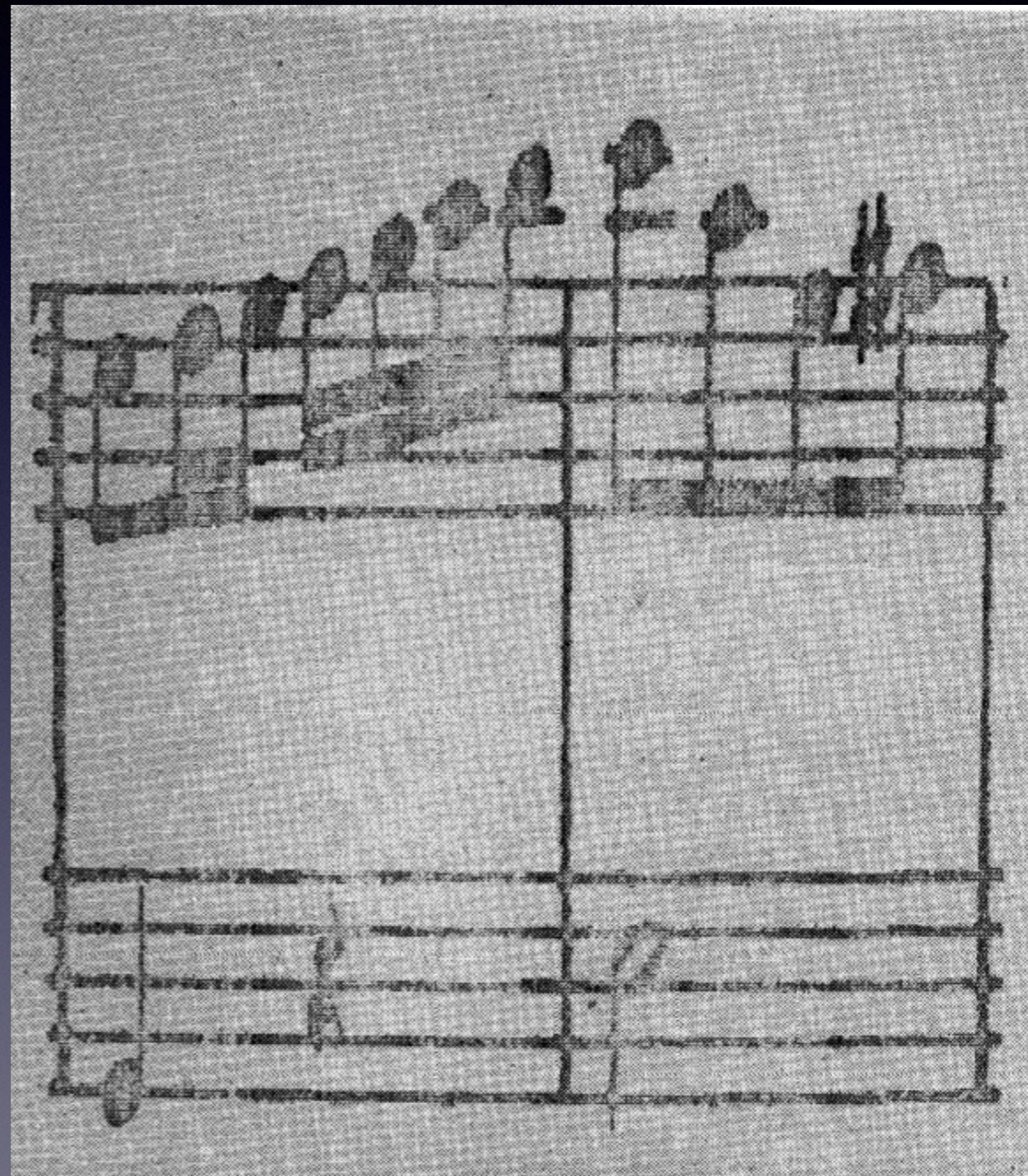
cogn
D. D



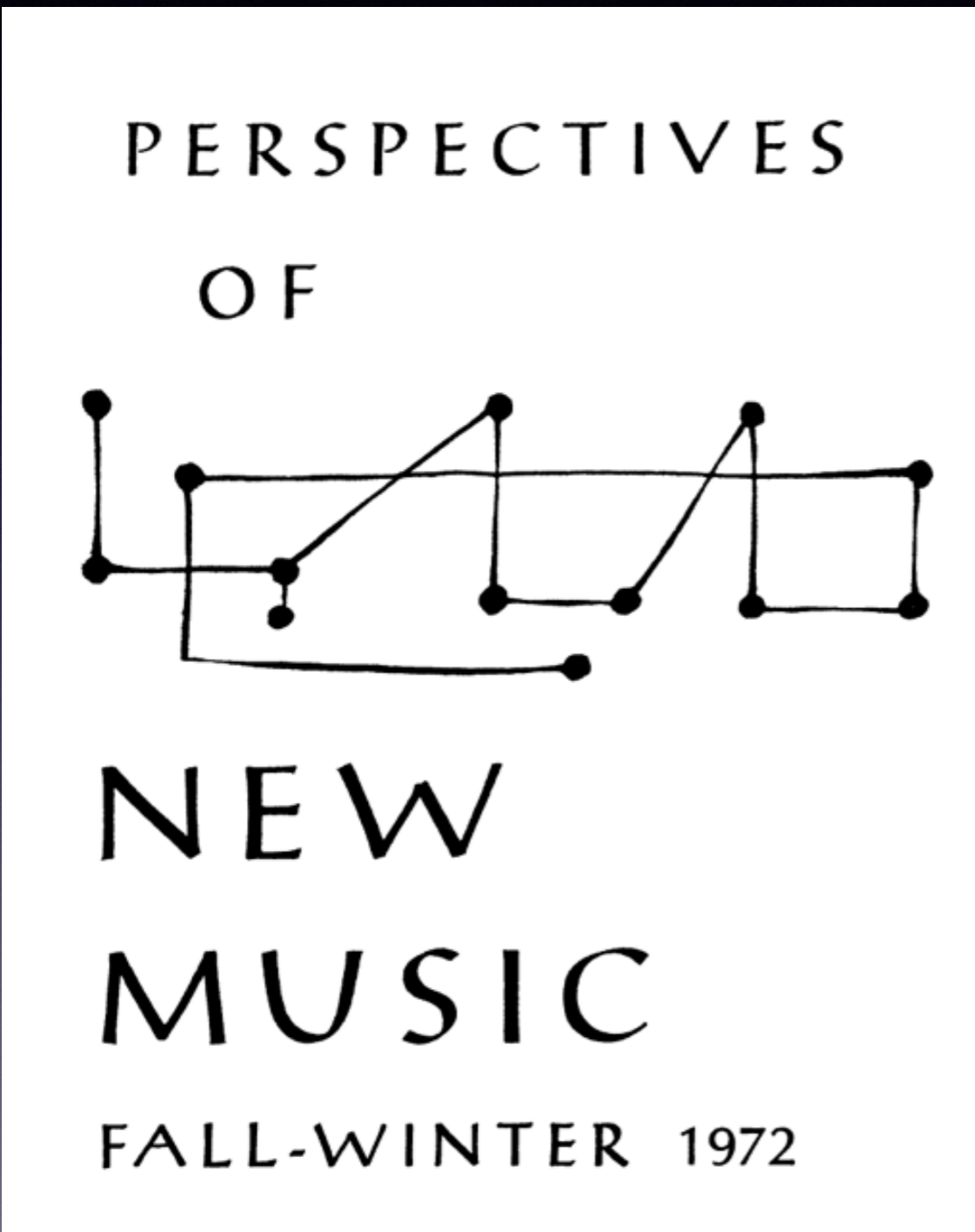
Denis Pruslin with
grandson Kevin, 2010/11

music

The first published digital scan of music (1970)



Review by Michael Kassler (1972)



Review: Optical Character-Recognition of Printed Music: A Review of Two Dissertations

Michael Kassler

Page 250 of 250-254

OPTICAL CHARACTER-RECOGNITION OF PRINTED MUSIC: A REVIEW OF TWO DISSERTATIONS

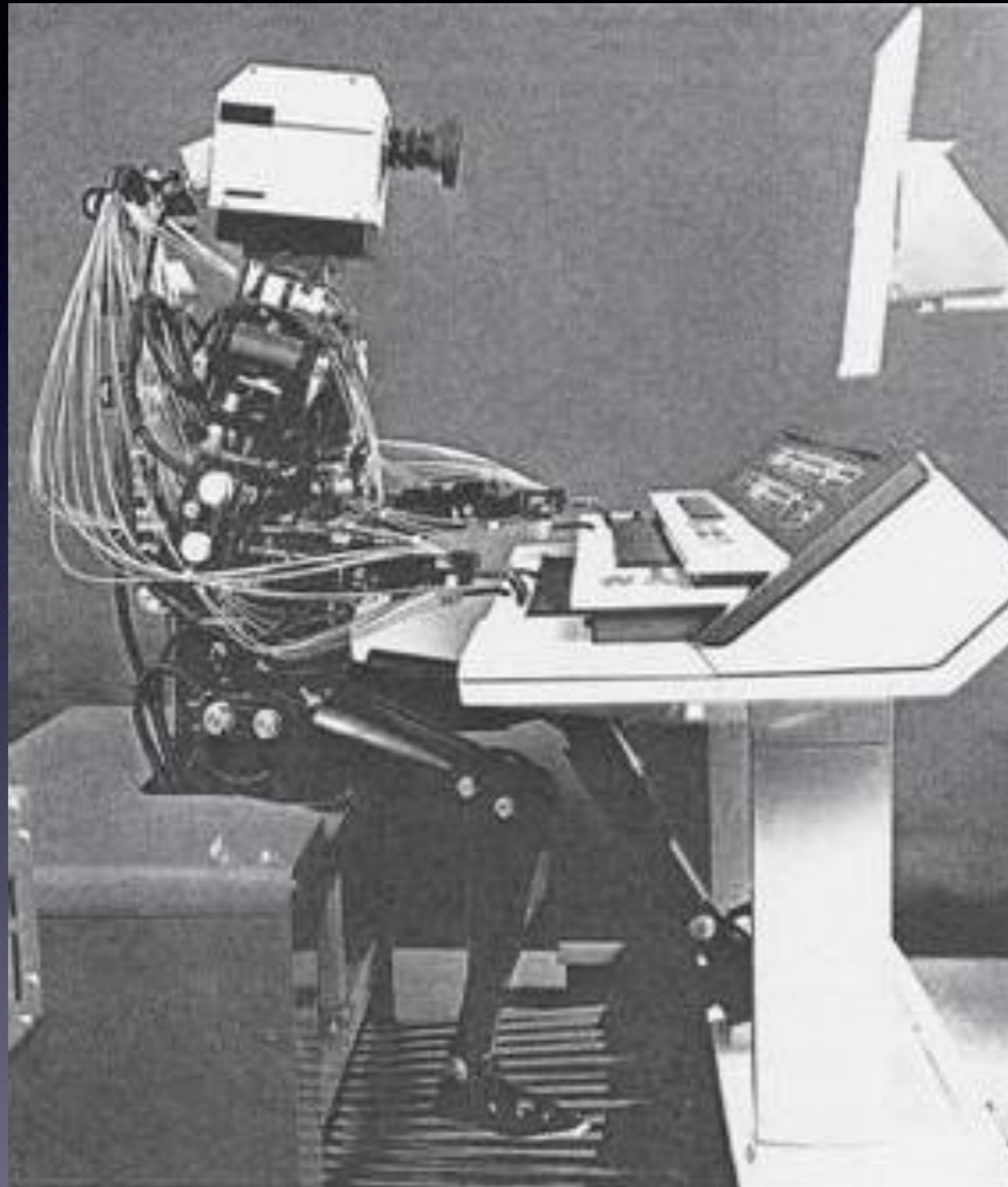
AUTOMATIC RECOGNITION OF SHEET MUSIC. By Dennis Howard Pruslin. Sc. D. Dissertation, Massachusetts Institute of Technology, 1966.

COMPUTER PATTERN RECOGNITION OF STANDARD ENGRAVED MUSIC NOTATION. By David Stewart Prerau. Ph. D. Dissertation, Massachusetts Institute of Technology, 1970.

Readers of *Perspectives* scarcely need be reminded of the pre-eminence of the written-musical domain (i. e., that domain of musical experience in which music is presented visually in one or another system of musical notation) in musicology: before Edison composers could not produce records of their work in the sounded-musical domain, and other domains of musical experience such as the tactile domain utilized in the Braille system have been employed comparatively infrequently; and even after Edison various extra-musical considerations (such as copyright law and the relatively high cost of sound-processing machinery) have joined with tradition to keep the written-musical domain a principal mode of non-transient musical communication. Within this domain various systems of musical notation have achieved various degrees of currency at various places and times, but of all these systems one—the current common musical notation ('CCMN' for short)—has dominated: virtually all music printed has been printed in one or another 'dialect' of CCMN: even music originally noted in another system generally has been transcribed into CCMN before printing.

In recent years digital computers have become more efficient and more prevalent, so that today, at least in computationally well-developed parts of the world, it no longer is unreasonable to delegate, or to plan to delegate, musical processes to electronic computing machinery. Of course, many musical processes do not involve previously recorded musical compositions: perhaps it is to the comparatively early success of a few such computer-mediated processes that an unfortunate synecdochic misidentification of 'computer music' with 'synthesizing sound through the use of a digital computer' has arisen.¹ But (and of this too readers will be well informed) central to musicology are processes that do involve prior musical compositions, and for the full delegation of these processes to computing machinery the relevant compositions must be put into computer-acceptable form. Human key-punchers can transcribe from CCMN onto (say) punch cards (at Princeton University the *Masses* of Josquin were so transcribed, at a rate of approximately 20 minutes per printed CCMN page), but as this task clearly requires no intelligence beyond that with which machines can be endued it is only natural to consider

1984: Wabot-2



<https://www.scaruffi.com/mind/ai/wabot.jpg>

OMR Thesis



- ❖ 1997 David Bainbridge
- ❖ 2006 Laurent Puig
- ❖ 2009 Alicia Forn
- ❖ 2012 Ana Rebelo
- ❖ 2014 Andrew Ha
- ❖ 2016 Jorge Calvo



2000: Gamera



- ❖ Fra
- rec
- ❖ De
- ❖ Im
- ❖ Do
- ❖ Sy
- ❖ Po

)

atch

Generalized Algorithms and Methods for Enhancement and Restoration of Archives

Gamera @ Peabody

- ❖ Designed and built by Karl McMillan and Michael Droettboom (started fall 2000)
- ❖ Master's students at Computer Music Department at Peabody Conservatory of Music, Johns Hopkins University
- ❖ Both worked at Digital Knowledge Centre, Johns Hopkins University Library
- ❖ Both graduated in 2002

Gamera developers

- ❖ Karl McMillan

- ❖ Expert of Security (Linux)
- ❖ Worked at Tru
- ❖ CTO at RAKI

- ❖ Michael Droer

- ❖ Science software Institute (NASA)
- ❖ Still active in



(Linux)

Flat

escope Science

Birth of Gamera

Gamera: A Structured Document Recognition Application Development Environment

Karl MacMillan, Michael Droettboom, and Ichiro Fujinaga

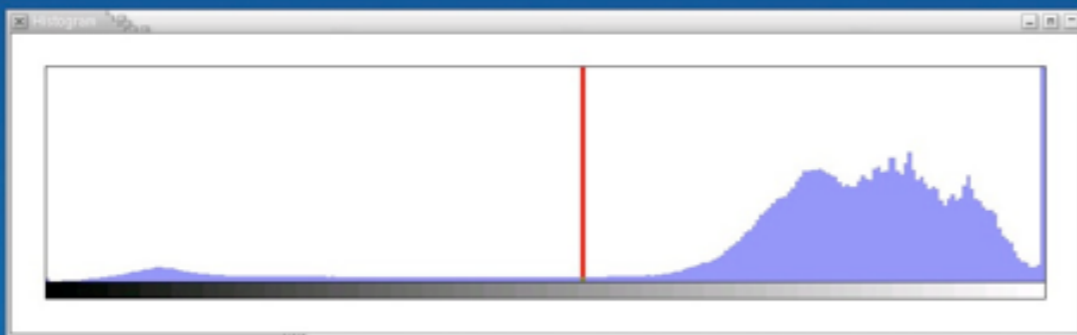
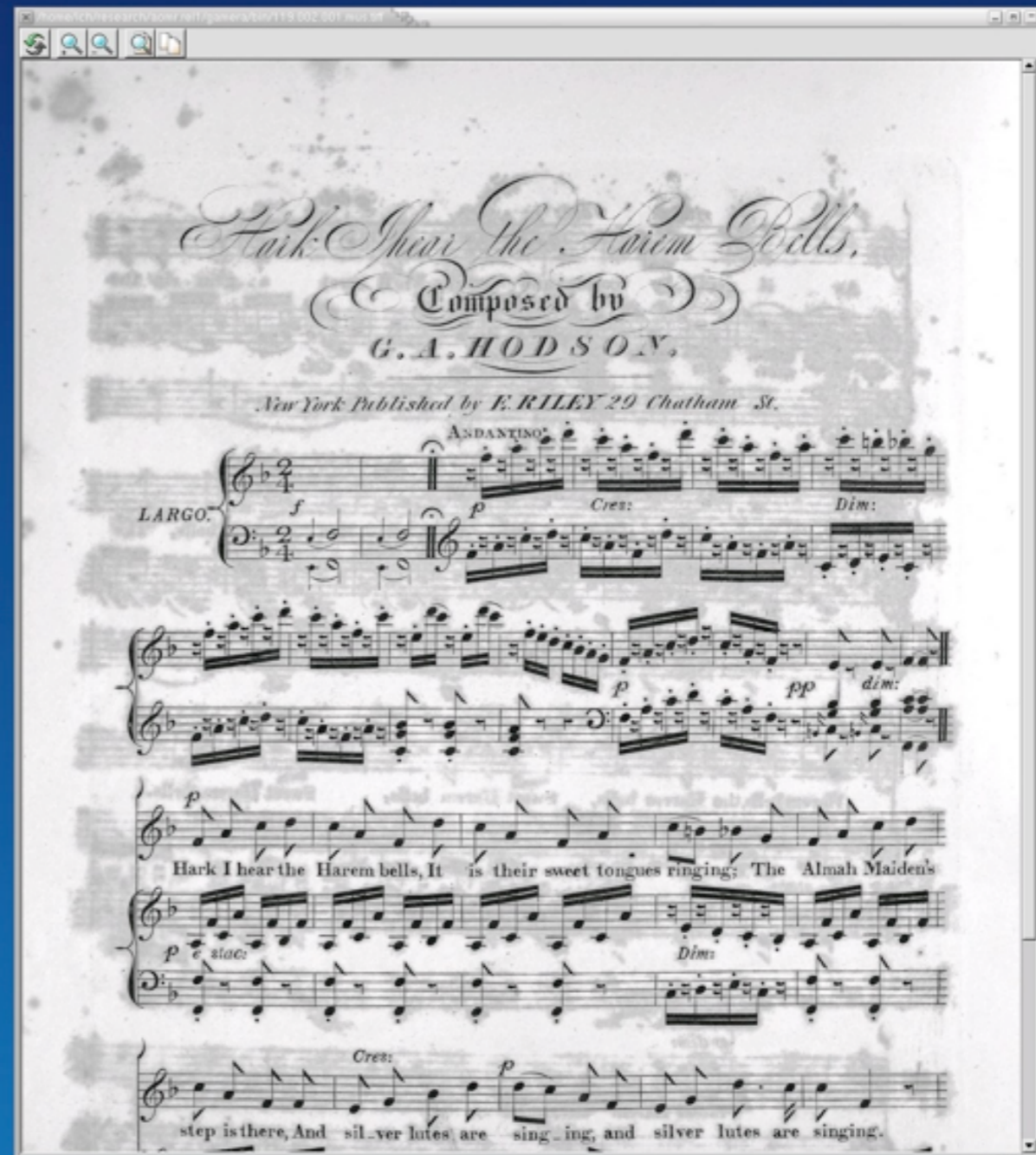
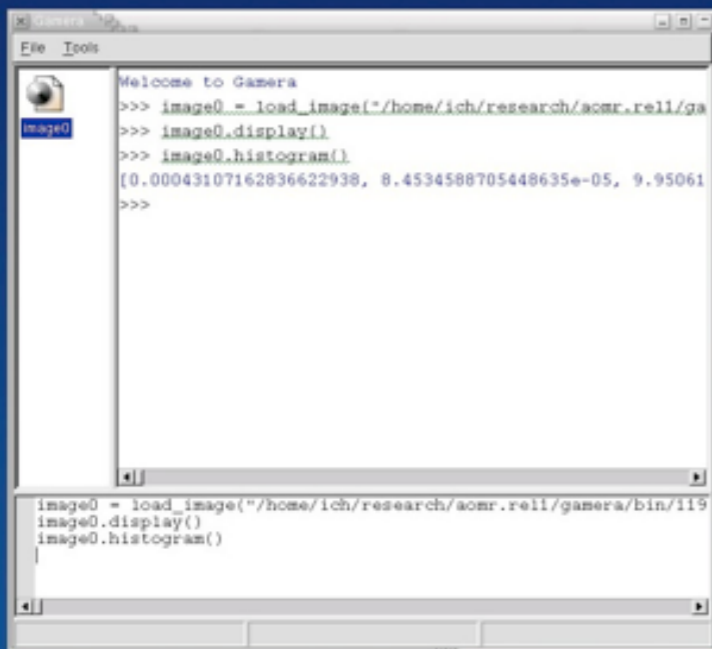
Peabody Conservatory of Music
Johns Hopkins University

1 East Mount Vernon Place, Baltimore MD 21202
email: {karlmac,mdboom,ich}@peabody.jhu.edu

- ❖ First paper presented at the 2nd International Symposium on Music Information Retrieval (ISMIR: October 2001) in Bloomington, IN


Early Gamera Screenshot (Linux)

ca. June 2002



Original Gamera Website

ca. December 2002

<p><i>rapid access</i></p> <ul style="list-style-type: none"> Introduction Authors Software Screenshots Papers Imaging Links 		<p>Software framework for the creation of domain-specific recognition applications</p>
<p>Introduction</p>	<p>This page describes the Gamera system, being developed at the Digital Knowledge Center at Johns Hopkins University. The project is funded in part by the National Science Foundation, the Institute for Museum and Library Services, and the Levy family. Please review the papers and software below for more information.</p>	<p>News flash! Large parts of Gamera are currently being rewritten. The major advantages of this are faster compilation- and run-times, portability to Windows, Mac OS-X and Unix, and a richer plugin system. We plan on making version 2.0 available once the code base stabilizes.</p>
<p>Authors</p>	<p>Ichiro Fujinaga Michael Droettboom Karl MacMillan</p>	
<p>Software</p>	<p>Download: Gamera 1.0 can be downloaded here (gzipped tarball).</p>	

2001–2008: Evolution of Gamera

❖ 2001
an
❖ 2002
re
❖ 2003
U
❖ 2004

Project: Gamera: Summary

Summary | Admin | Home Page | Tracker | Bugs | Support | Patches | RFE | News | CVS | Files |

Gamera is a framework for the creation of structured document analysis applications by domain experts. It combines a programming library with GUI tools for the training and interactive development of recognition systems.

:Python Foundry

- Development Status: 4 - Beta
- Environment: Cocoa (MacOS X), Win32 (MS Windows), X11 Applications
- Intended Audience: Developers, Education, Information Technology, Science/Research
- License: GNU General Public License (GPL)
- Natural Language: English
- Operating System: MacOS X, Windows NT/2000, Linux
- Programming Language: C++, Python
- Topic: Graphics Conversion, Artificial Intelligence, Information Analysis

Project UNIX name: gamera
Registered: 2004-01-12 10:13

Activity Percentile (last week): 63%

View project activity [statistics](#)
View list of [RSS feeds](#) available for this project

Developer Info

Project Admins:
[abrzeczk](#)
[mdboom](#)

Developers: 4
[\[View Members\]](#)

Latest File Releases

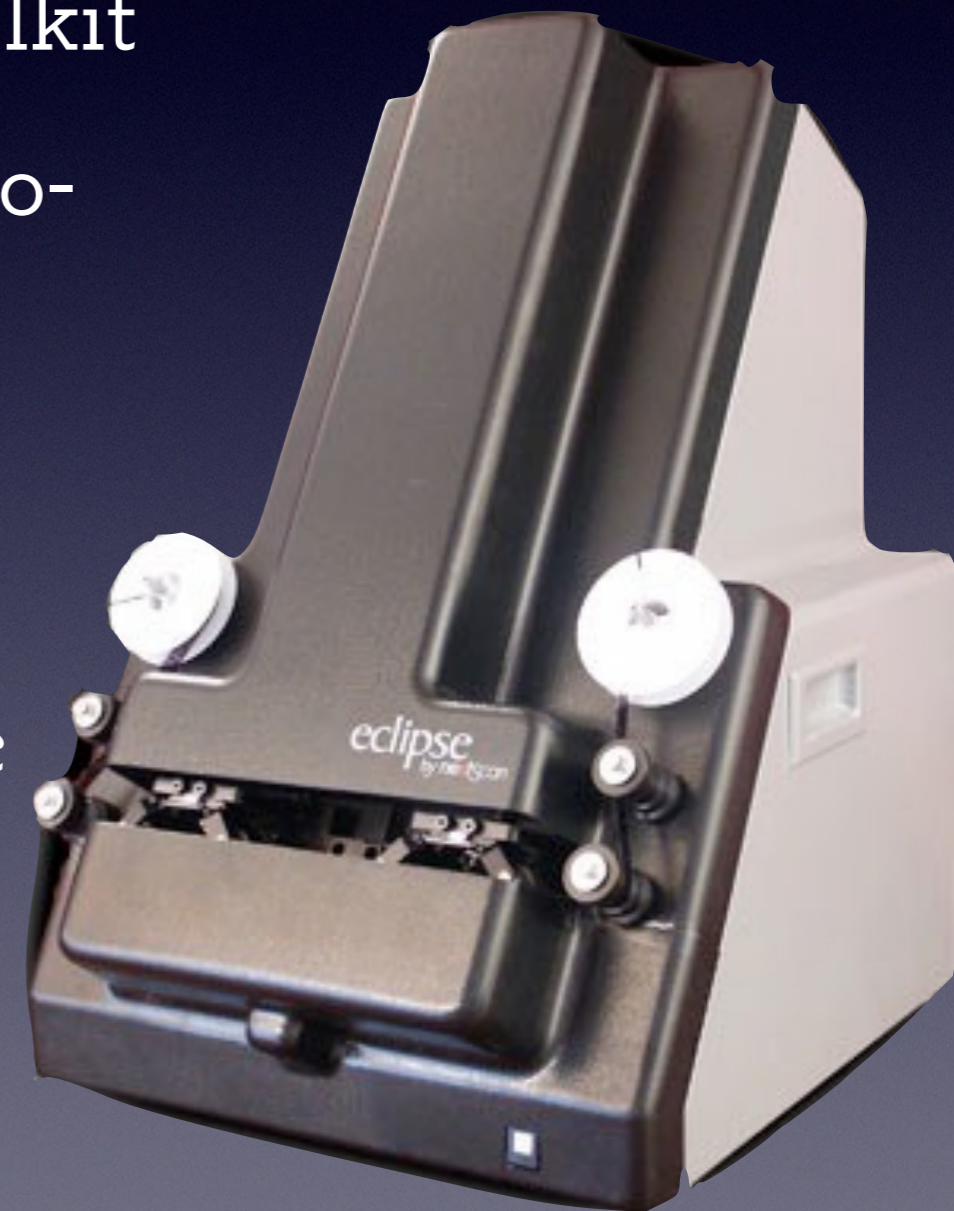
Package	Version	Date	Notes / Monitor	Download
gamera	gamera-2.2.0pre2	August 17, 2004	-	Download
gamera-video	1.0	January 13, 2004	-	Download
toolkit-skeleton	2.2.0pre2	September 8, 2004	-	Download
wxPython-Gtk-X11	wxPython-Gtk-X11-2.4.2.4	July 21, 2004	-	Download

[\[View ALL Project Files\]](#)

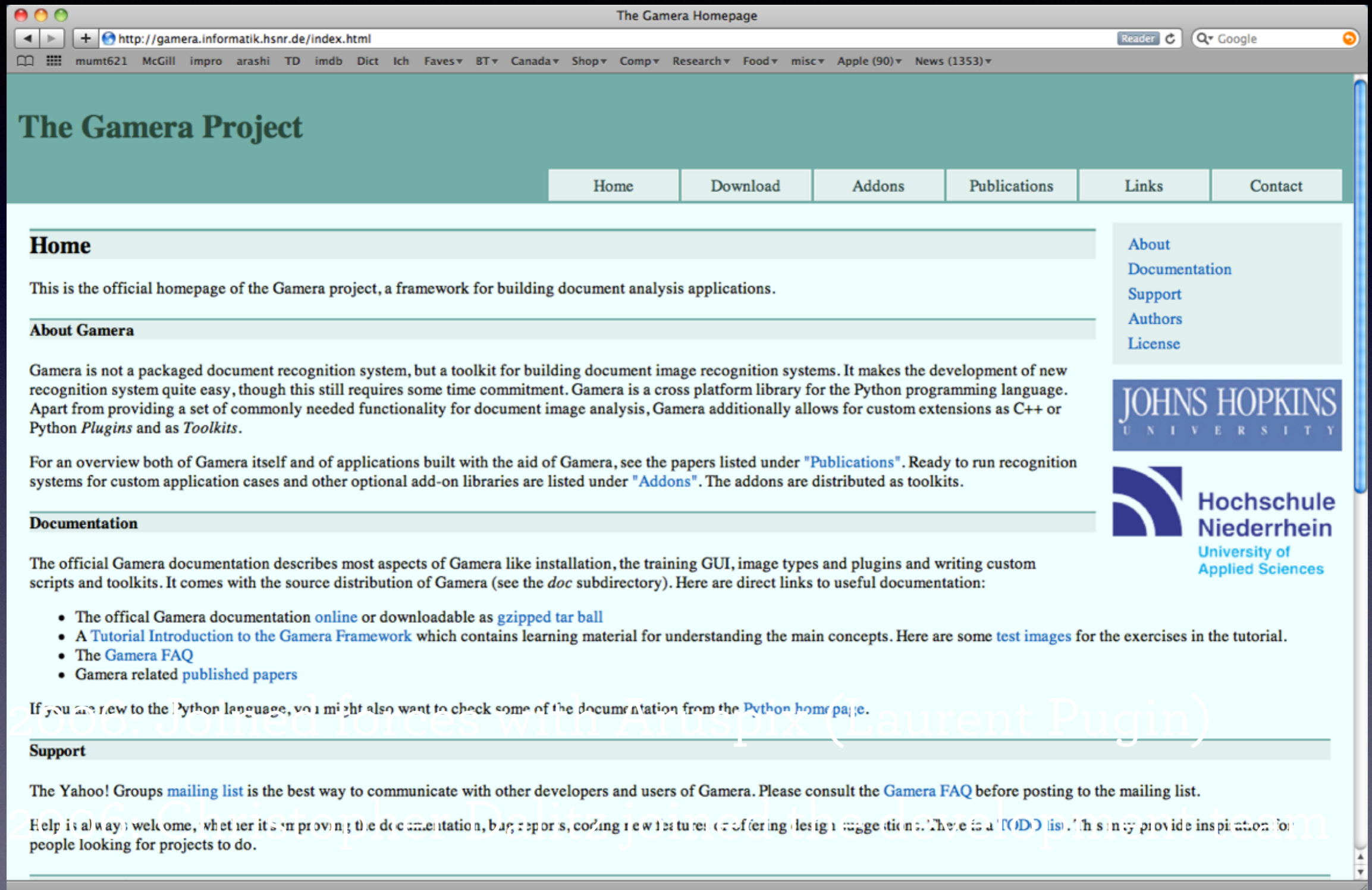
2005: GEMM

(Gamut for Early Music on Microfilms)

- ❖ Based on GAMUT: Gamera-based Automatic Music Understanding Toolkit
- ❖ Possibility of OMR for music on microfilms
- ❖ Almost all old Western music are on microfilms
- ❖ Efficient digitization using automatic microfilm scanner (Eclipse 500: 590ppm)



2001–2008: Evolution of Gamera



❖ 2008: Website moved to: <http://gamera.informatik.hsnr.de>

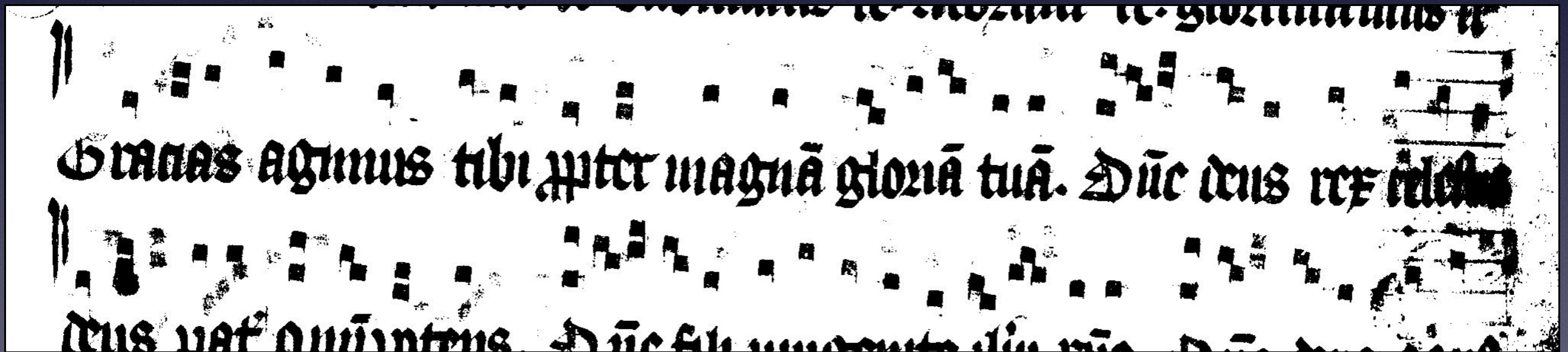
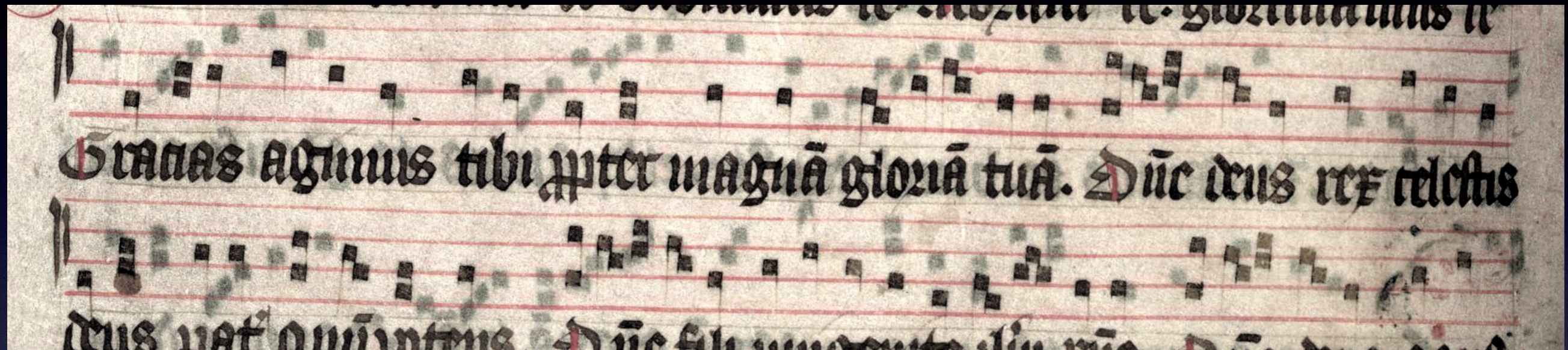
Some Features of Gamera c. 2008

Preprocessing



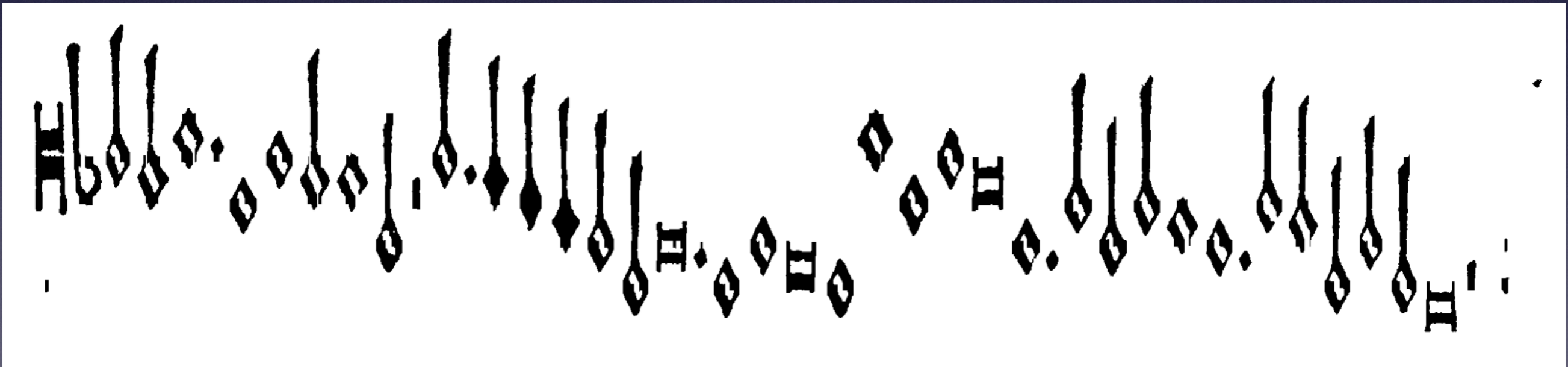
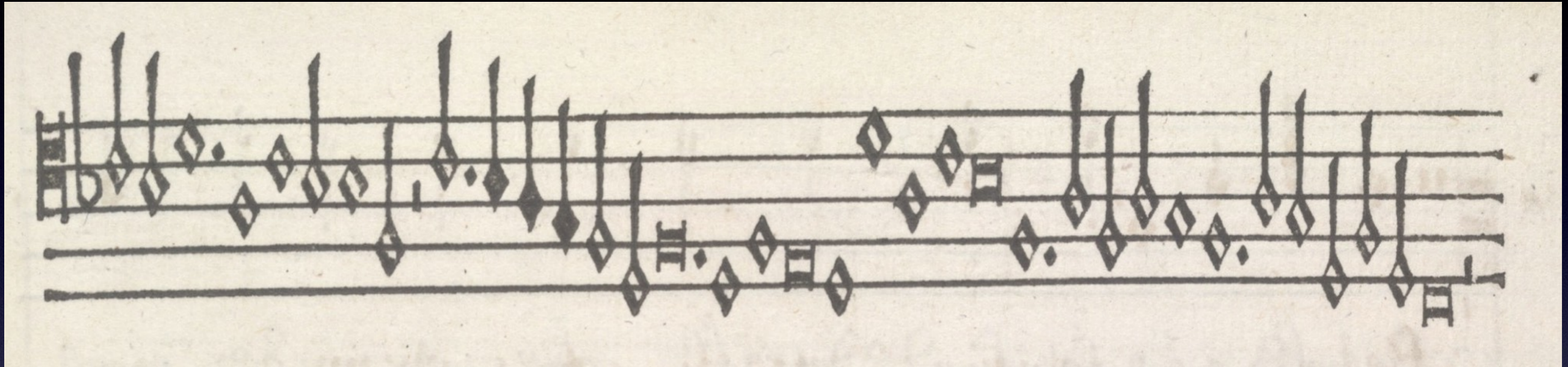
Brightness Enhancement

Preprocessing

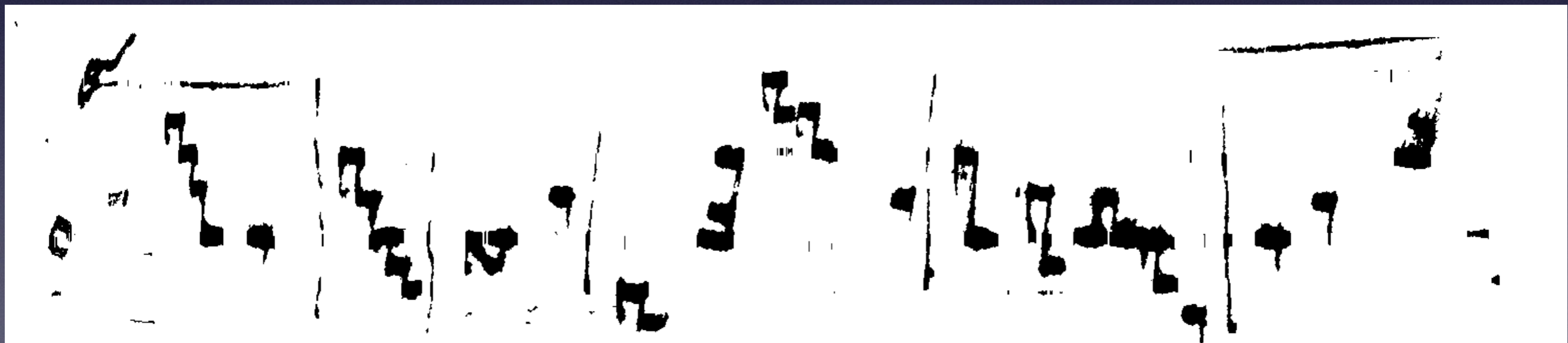


Thresholding

Staffline Removal

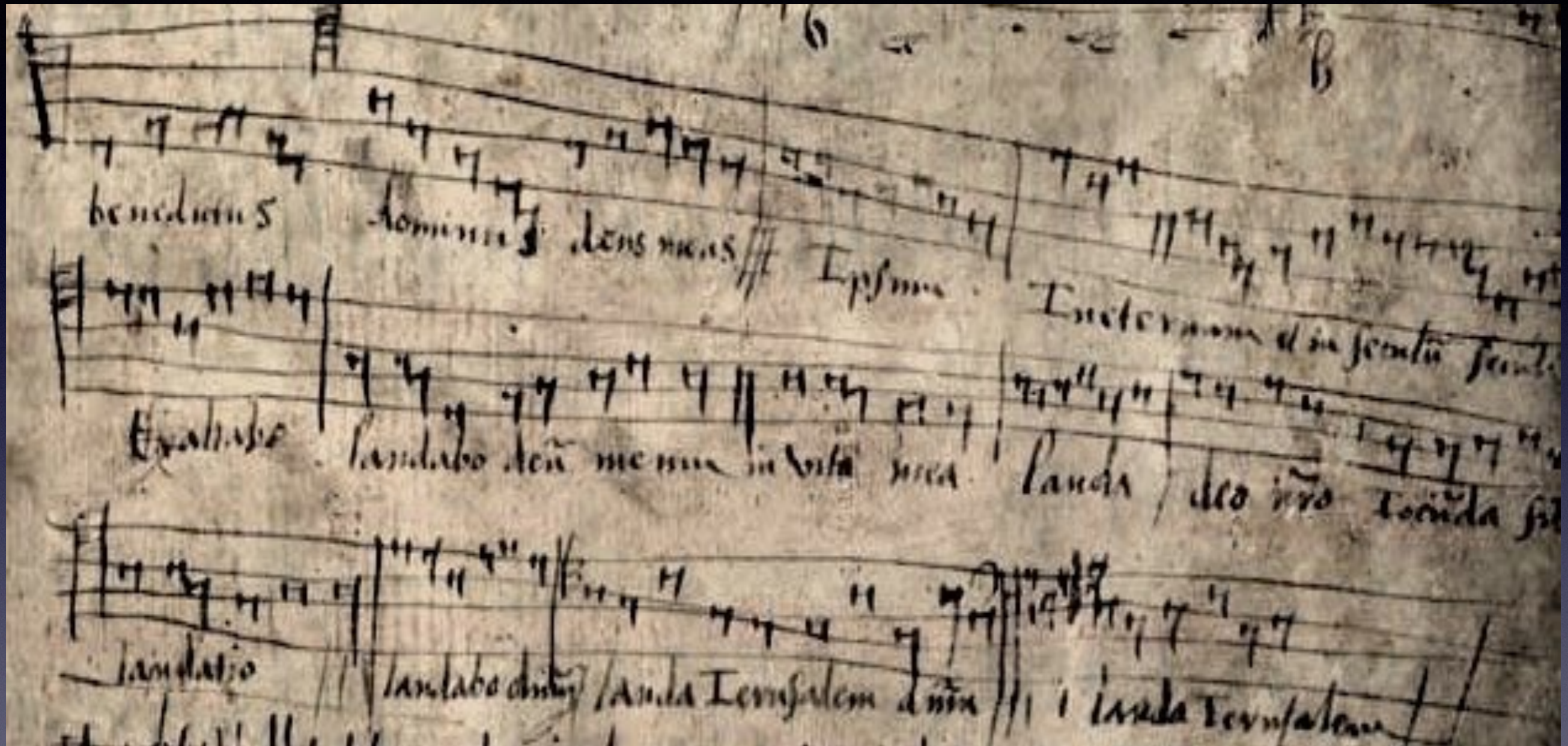


Staffline Removal



Four-line hand-drawn staff example

Staffline Removal



Difficult

Staffline Removal

Ræludium Laurencini.

The image displays two versions of a lute tablature score for 'Ræludium Laurencini'. The top version is the original manuscript, featuring five staves with letters (a, b, c, d, e) placed on the lines to indicate fret positions. The bottom version is a processed version where the stafflines have been removed, leaving only the letters and their relative positions on the staves. Both versions include a decorative initial 'R' and 'B' at the beginning of the piece.

Lute tablature

Symbol classifier / Gamera

Window: kNNInteractive Classifier

File Image Classifier Rules

UNCLASSIFIED

Search: x.get_main_id()

Left Panel: Symbols

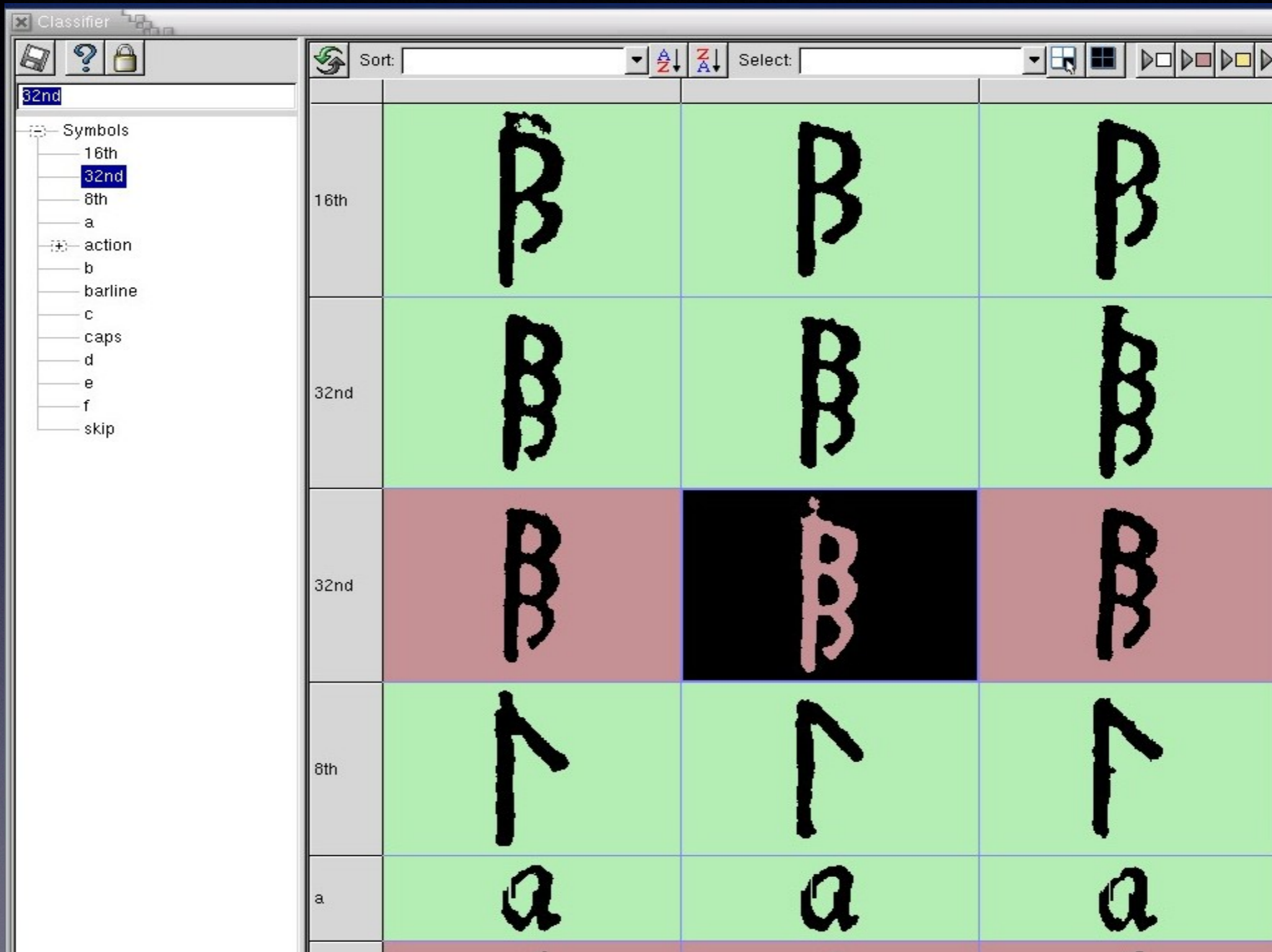
- group
- part
- split
- cc_analysis
- cc_and_clu:
- music_segm
- remove_ster
- splitx
- splitx_left
- splitx_max
- splitx_right
- splity
- splity_bottom
- splity_top
- gb_lbl_c35f10
- gb_lwa_102_a
- skip

Main View: Grid of 8x8 glyphs (musical symbols) with various classification markers.

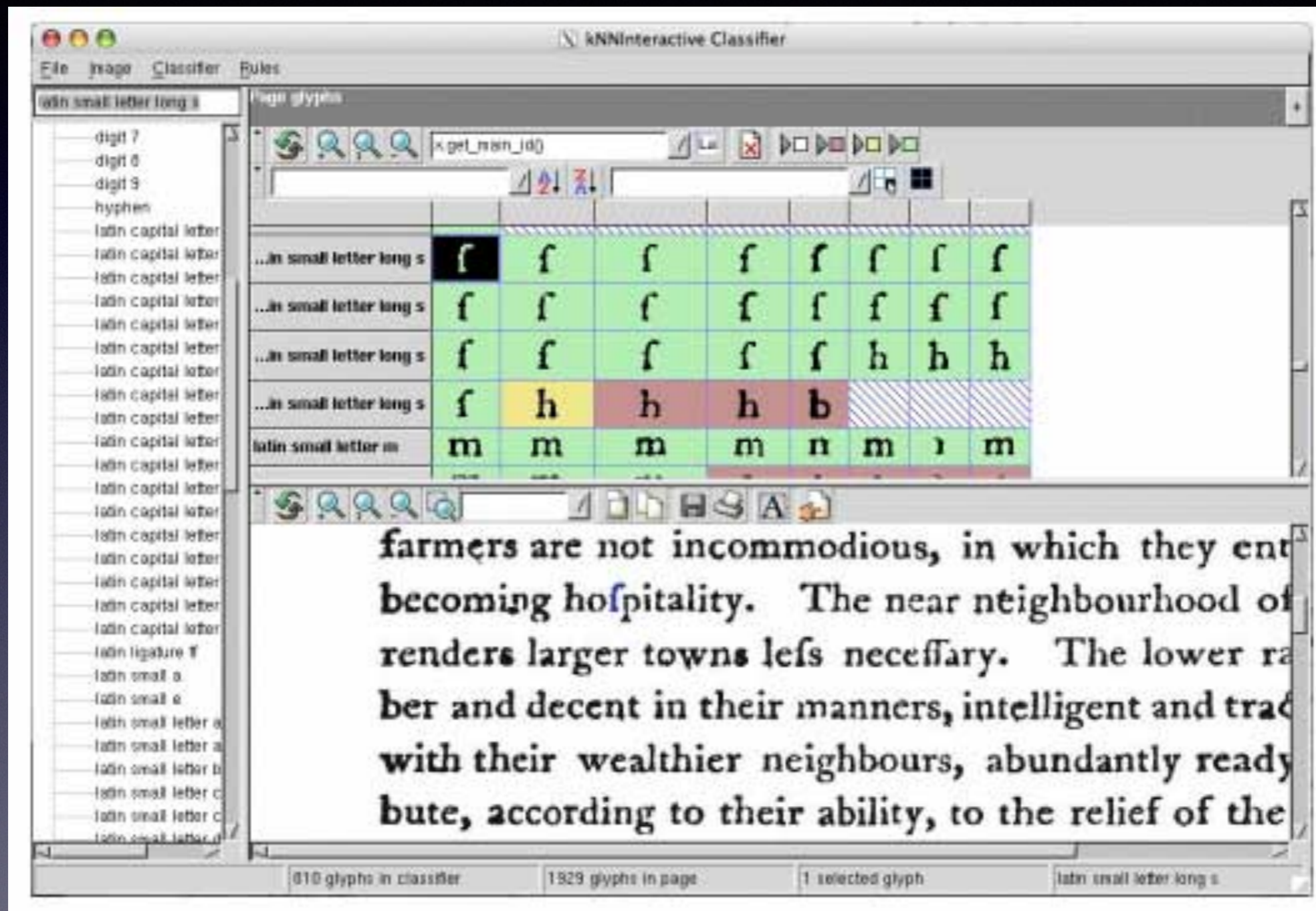
Bottom View: Row of 12 colored glyphs (musical symbols).

Status: 23 selected glyphs UNCLASSIFIED

Lute tablature symbol recognition

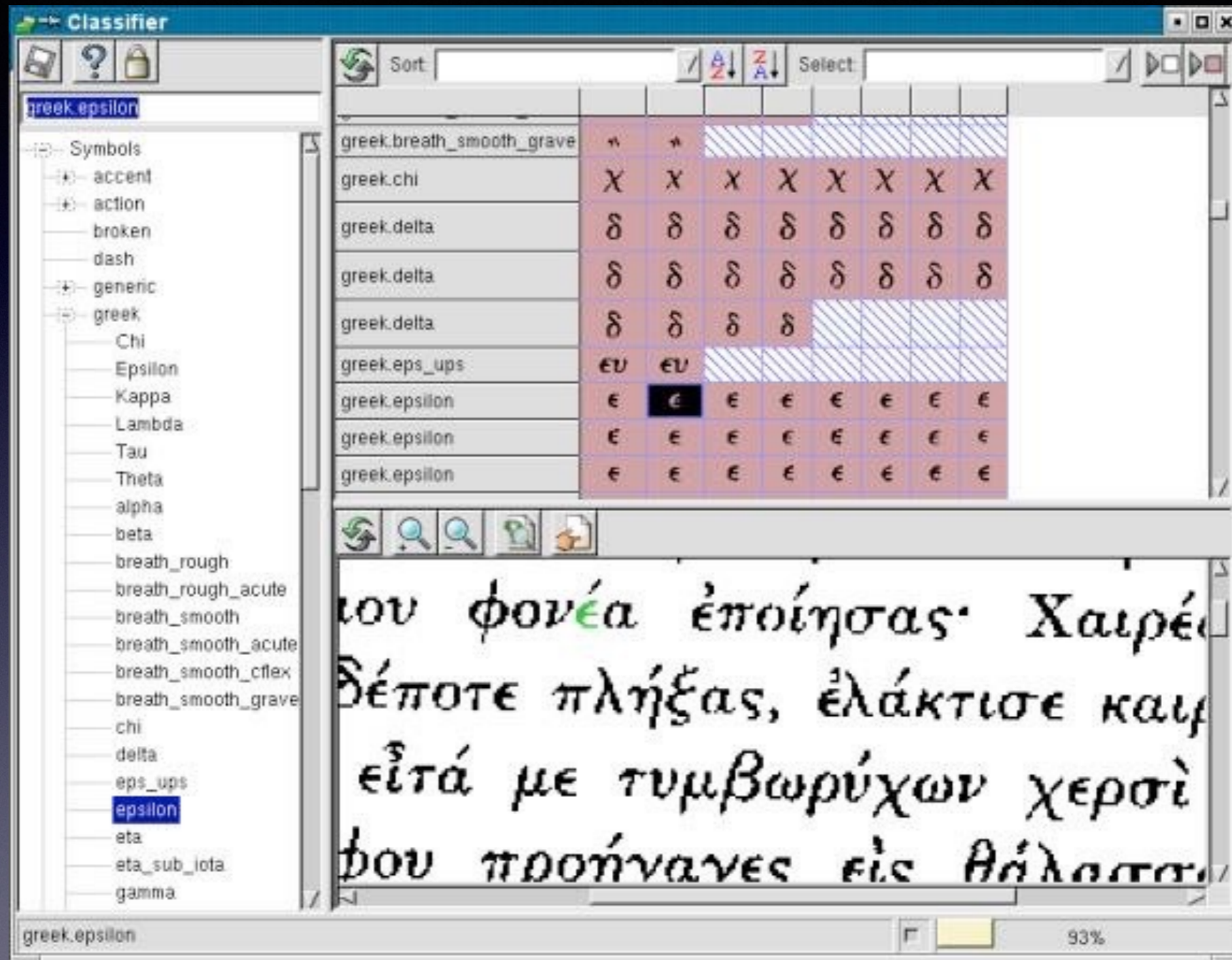


Other Applications



Early Modern English

Other Applications



Greek

Other Applications

Navajo language recognition

Other Applications

Du tout me tins a lon perde



*E oï la vieille ba a bel acueil.
La vieille illec pl' ne seïone.
 Le troï a bel acueil se tōne.
 Qui la toure par mal li retarde.
 E ar bien se souffrist de tel garde.
 Tant ba quel a encontre
 La toure ou tolt est entree*

Roman de la Rose (Bodelian MS Douce 195, 90v, 15th C.)

2008: Other Applications

Optical Recognition of Psaltic Byzantine Chant Notation

Christoph Dalitz · Georgios K. Michalakis · Christine Pranzas

2009: Other Applications

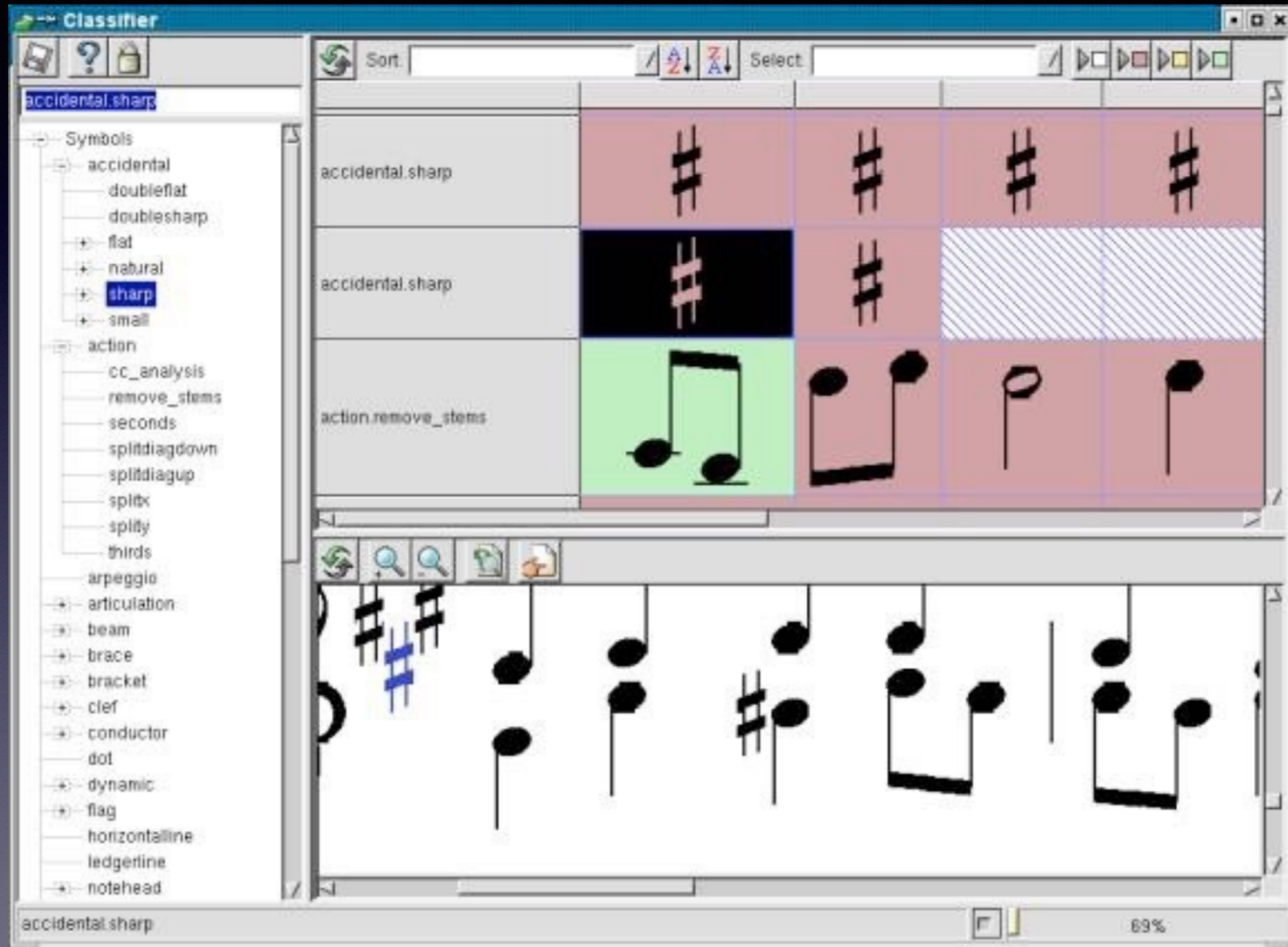
	F	F	F	F	F	F
a	a	b	a	b	a	a
a	a	b	a	b	a	a
a	a	a	c	b	a	c
c	c	a	b	a	c	c



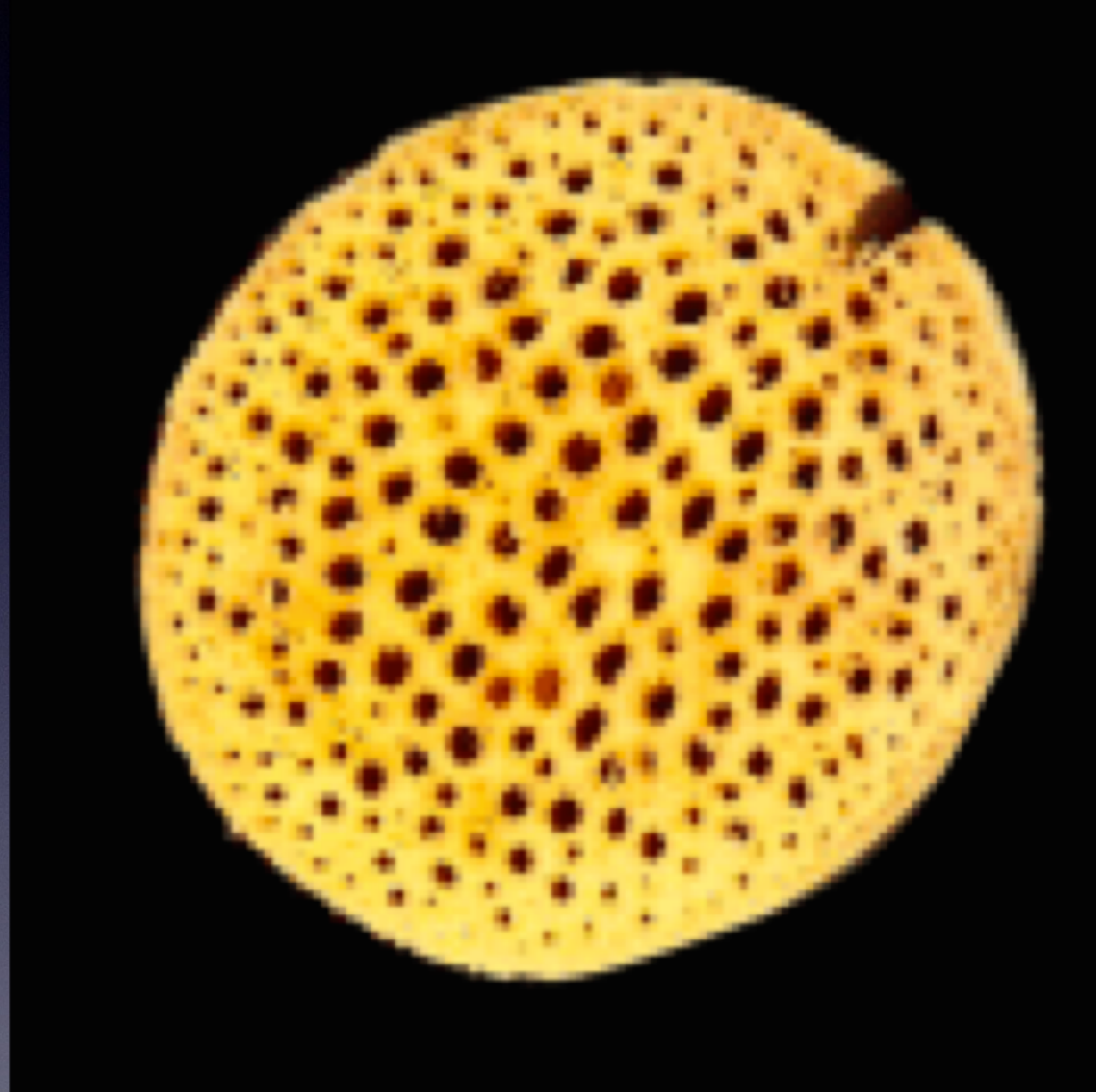
Optical Recognition of Lute Tablature

Christoph Dalitz · Thomas Karsten

Other Applications

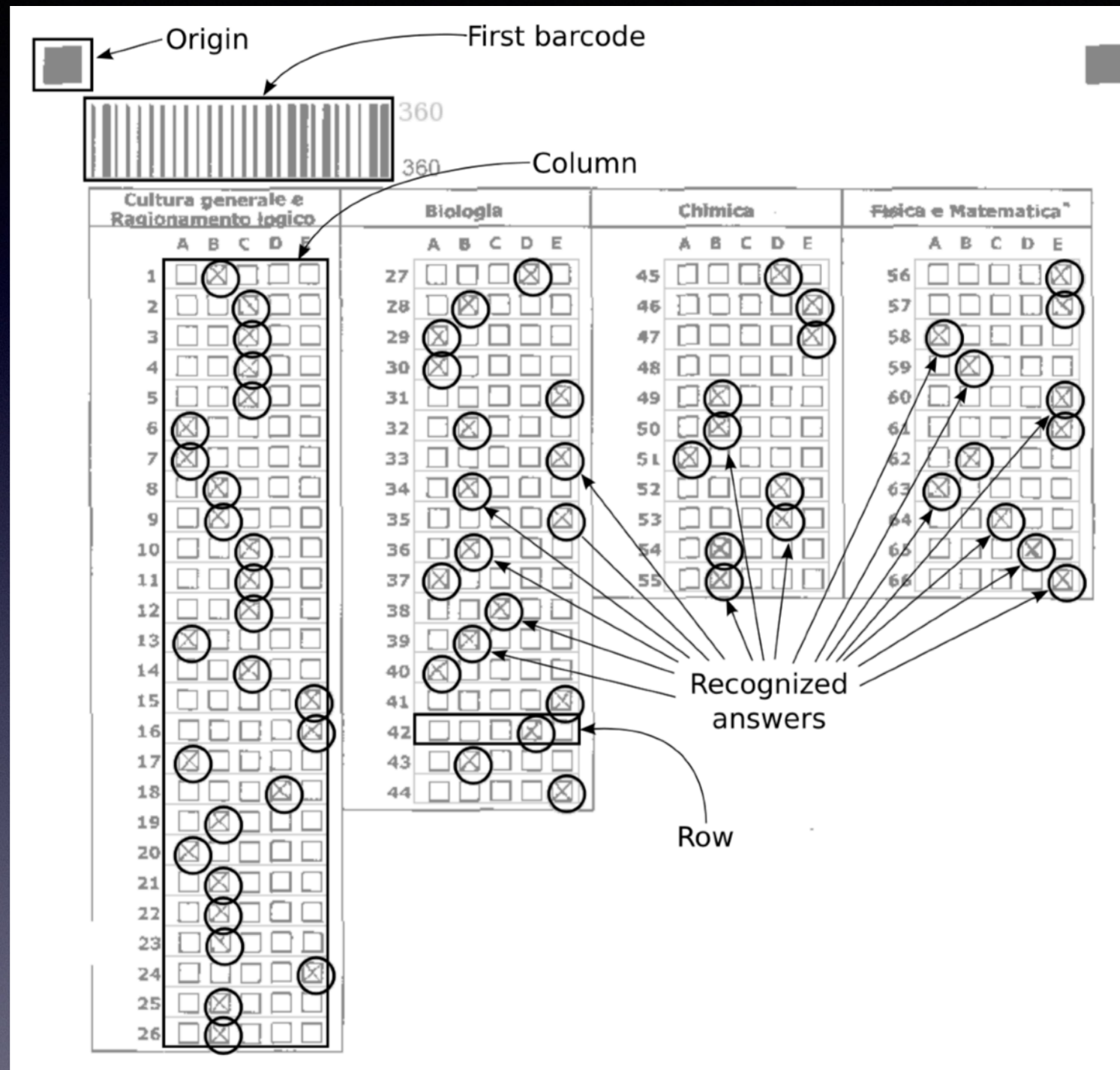


2005: Other Applications



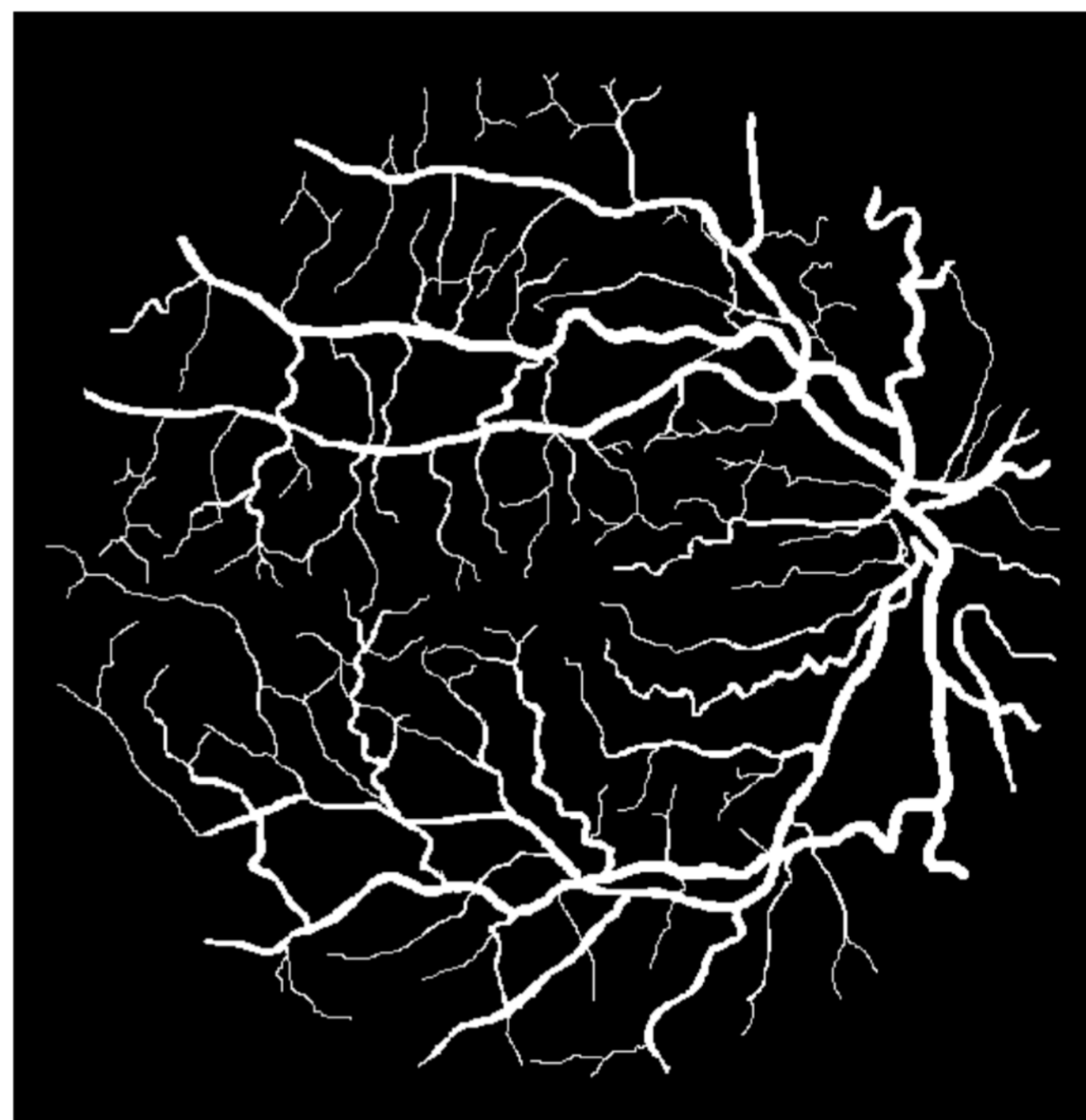
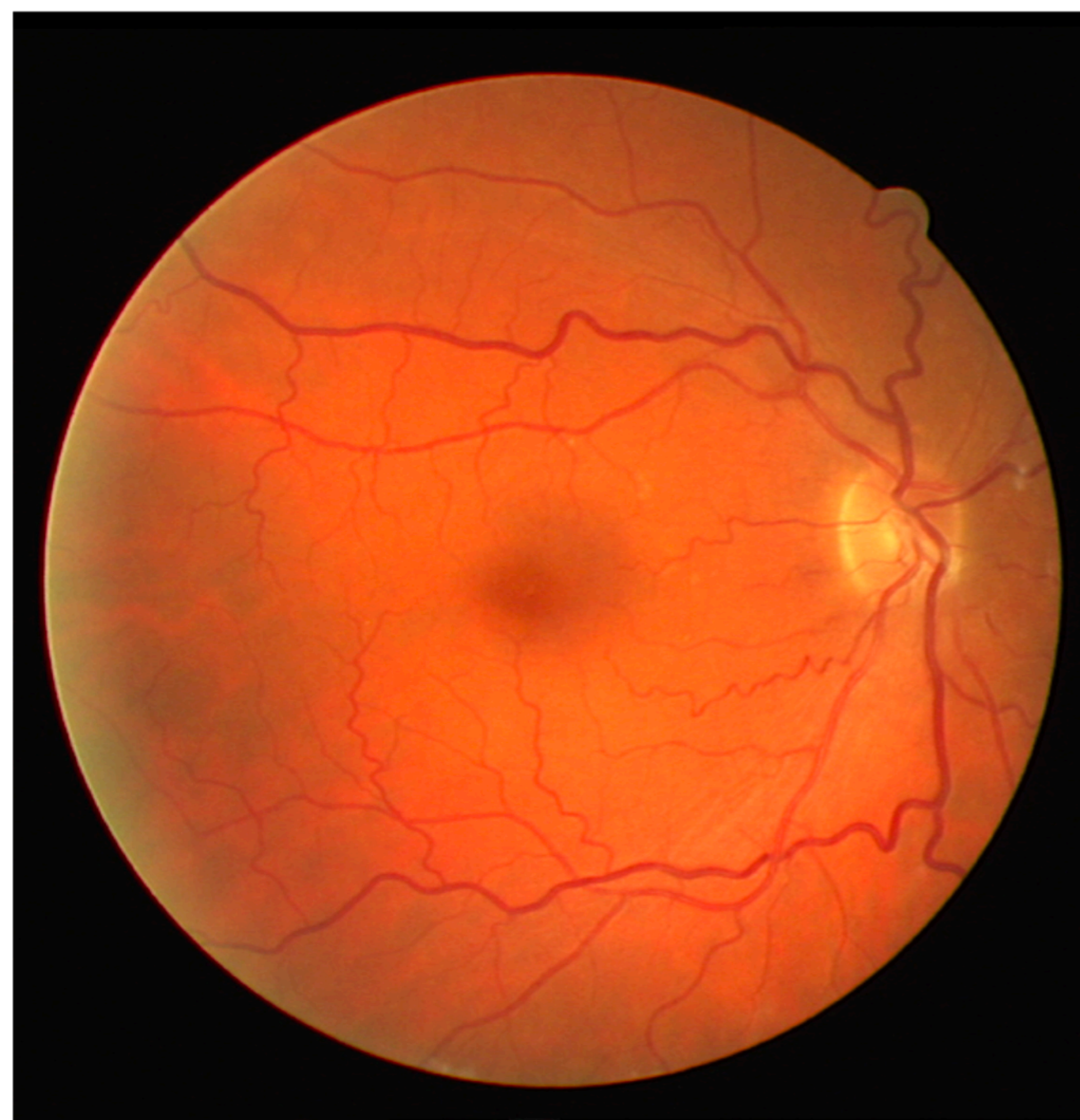
Vascular Anatomy of Plants (Alex Cobb, Harvard)

2009: Other Applications



Multiple-choice Test Recognition System (Spadaccini & Rizzo)

2015: Other Applications



Blood vessel extraction (Dalitz et al.)

2005–2016: VIPs



- ❖ 2005–2012 Ashley Burgoyne (PhD)
- ❖ 2006–2008 Laurent Pugin (Postdoc)
- ❖ 2007–2016 Andrew Hankinson (PhD/Postdoc)
- ❖ 23 publications on OMR between 2007–2016
- ❖ 18 publications on OMR between 2007–2012 (3 per year)

2002: Aruspix



Typographic Music: music set with type (single-impression)

2002: Aruspix



- ❖ Developed by Laurent Pugin
- ❖ Specialized for typographic music
- ❖ Uses HMM (Hidden Markov Model)
- ❖ Does not remove staff lines



SIMSSA | Single Interface for Music Score Searching and Analysis

- ❖ Similar to “Google Books” minus Google
 - ❖ OMR (optical music recognition) to enable full-text search
 - ❖ Sophisticated music analysis and query
- ❖ Access to digitized scores world-wide from a single website
- ❖ SSHRC-funded 11-year project: 2011–2021: \$4.4M

What would SIMSSA provide?

- ❖ Web-based OMR system with score editors
 - ❖ Rodan (Remote Online Document Analysis Network)
 - ❖ Gamera + Aruspix (a combination of existing OMR software)
 - ❖ Verovio (open-source music engraver) by Laurent Pugin
 - ❖ “Gradsourcing” to correct errors
 - ❖ Early music
- ❖ Web-based user interface to view, search, annotate, and analyze scores
 - ❖ MEI (Music Encoding Initiative) by Perry Roland and others
 - ❖ Diva.js (web-based IIIF-compatible document delivery system)
 - ❖ Humdrum / music21 (analytical tools)

The Vision: Global Music Library



Tools



Music Encoding Initiative

Humdrum

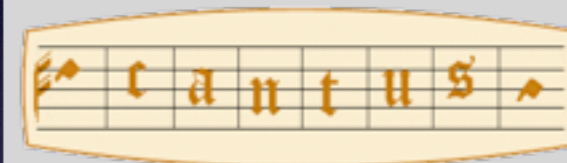
music21



Early Music Online

Music Treasures Consortium

Metadata & Text



SIMSSA Team

- ❖ Musicologists (20)
- ❖ Music Librarians (8)
- ❖ Music Technologists (11)
- ❖ Partners (23) including:
 - ❖ Bavarian State Library
 - ❖ Bibliothèque nationale de France
 - ❖ British Library
 - ❖ Harvard University Music Library
 - ❖ HathiTrust Research Center
 - ❖ New York Philharmonic Archives

Optical Music Recognition (OMR)

A process of converting images of music scores into a symbolic computer representation, such as MIDI, MusicXML, or MEI (Music Encoding Initiative).



```

<section xml:id="section-0000001229415468">
  <measure xml:id="measure-L6" n="1">
    <staff xml:id="staff-L6F2N1" n="1">
      <layer xml:id="layer-L6F2N1" n="1">
        <rest xml:id="rest-L7F2" dur="2" fermata="above" />
        <note xml:id="note-L11F2" dur="2" oct="4" pname="g" accid.ges="n" />
      </layer>
    </staff>
    <staff xml:id="staff-L6F1N1" n="2">
      <layer xml:id="layer-L6F1N1" n="1">
        <note xml:id="note-L7F1" dots="1" dur="4" oct="3" pname="c" accid.ges="n" />
        <tuplet xml:id="tuplet-L8F1-L10F1" num="3" numbase="2" num.format="count">
          <beam xml:id="beam-L8F1-L10F1">
            <note xml:id="note-L8F1" dur="16" oct="3" pname="d" accid="s" />
            <note xml:id="note-L9F1" dur="16" oct="3" pname="e" accid="ff" />
            <note xml:id="note-L10F1" dur="16" oct="3" pname="f" accid="x" />
          </beam>
        </tuplet>
        <note xml:id="note-L11F1" dur="4" oct="3" pname="a" accid.ges="n" />
        <note xml:id="note-L12F1" dur="4" oct="3" pname="a" accid.ges="n" />
      </layer>
    </staff>
    <fermata xml:id="fermata-L10F1" staff="2" startid="#note-L10F1" place="above" />
    <tie xml:id="tie-L11F1-L12F1" startid="#note-L11F1" endid="#note-L12F1" />
    <slur xml:id="slur-L11F2-L16F3N1" staff="1" startid="#note-L11F2" endid="#note-L16F3" />
    <tie xml:id="tie-L12F1-L15F1" startid="#note-L12F1" endid="#note-L15F1" />
  </measure>
  <measure xml:id="measure-L13" n="2">
    <staff xml:id="staff-L13F2N1" n="1">
      <layer xml:id="layer-L13F2N1" n="1">
        <note xml:id="note-L15F3" dur="2" oct="4" pname="b" accid.ges="n" />
        <note xml:id="note-L16F3" dur="2" oct="5" pname="d" accid.ges="n" />
      </layer>
    </staff>
    <staff xml:id="staff-L13F1N1" n="2">
      <layer xml:id="layer-L13F1N1" n="1">
        <note xml:id="note-L15F1" dur="1" oct="3" pname="a" accid.ges="n" />
      </layer>
      <layer xml:id="layer-L15F2N2" n="2">
        <note xml:id="note-L15F2" dur="2" oct="2" pname="b" accid.ges="n" />
        <note xml:id="note-L16F2" dur="2" oct="2" pname="a" accid.ges="n" />
      </layer>
    </staff>
    <fermata xml:id="fermata-L16F3" staff="1" startid="#note-L16F3" place="above" />
  </measure>
  <measure xml:id="measure-L18" right="end" n="3">
    <staff xml:id="staff-L18F2N1" n="1">
      <layer xml:id="layer-L18F2N1" n="1">
        <note xml:id="note-L19F2" dur="4" oct="5" pname="c" accid.ges="n">
          <artic xml:id="artic-L19F2" artic="marc" />
        </note>
        <note xml:id="note-L20F2" dur="4" oct="5" pname="d" accid.ges="n">

```

Steps Involved in OMR

Digitized Score



Image Preprocessing

- Binarization
- Noise Removal
- Structural Analysis
- Image Segmentation



Music Symbol Recognition

- Staves Processing
- Symbol Segmentation
- Symbol Classification



Music Notation Reconstruction

- Symbol Combination
- Semantic Assignment (pitch, value)
- Musical Structure Reconstruction



Final Output

```

<pre>
<!--
  Copyright 2012-2013, Simon Steffenhagen.
  All rights reserved.
  See LICENSE for more details.
-->
<?xml version="1.0" encoding="UTF-8" ?>
<score>
  <part>
    <staff>
      <note>
        <pitch>4</pitch>
        <value>4</value>
      </note>
    </staff>
  </part>
</score>

```

2011: *Liber Usualis* Project


FEAST OF THE BLESSED SACRAMENT.
CORPUS CHRISTI.
Double of the First Class with Octave.
AT FIRST VESPERS.
All as at second Vespers, p. 956, except the following :
At Magn. Ant. 6. F
O quam su-á-vis est, * Dó- mi-ne, spí-
ri-tus tú- us! qui ut dulcédi-nem tú-am in ff-li-os de-
monstrá- res, páne su-a-vís- simo de caélo
praésti-to, esu-ri- éntes réples bónis, fa- stidi- ó- sos dí-
vi- tes dimít-tens in- á-nes. E u o u a e.
Cant. Magnificat. 6. F. p. 211 or p. 213. Prayer. Deus, qui nobis. p. 943.
At Compline and the Little Hours, the psalms of Sunday are said; the Hymn, today and throughout the Octave, is sung in the tone of Christmas, p. 367, with the doxology Jesu tibi sit glória, Qui natus es de Virgine. All hymns of the same metre end with this doxology throughout the Octave, even on the feasts of Saints, unless the contrary be indicated.
AT MATINS. †
Pater. Ave María. Credo.
D Omine, lábi- a mé-a apé-ri- es. R̄. Et os mé- um an-
† The Chant for Matins is taken from the publications of Solesmes.

Full-text search of 2,000 pages of Latin text and square notation



Search the Liber usualis

What is this? [Find out more about what we are trying to do.](#)

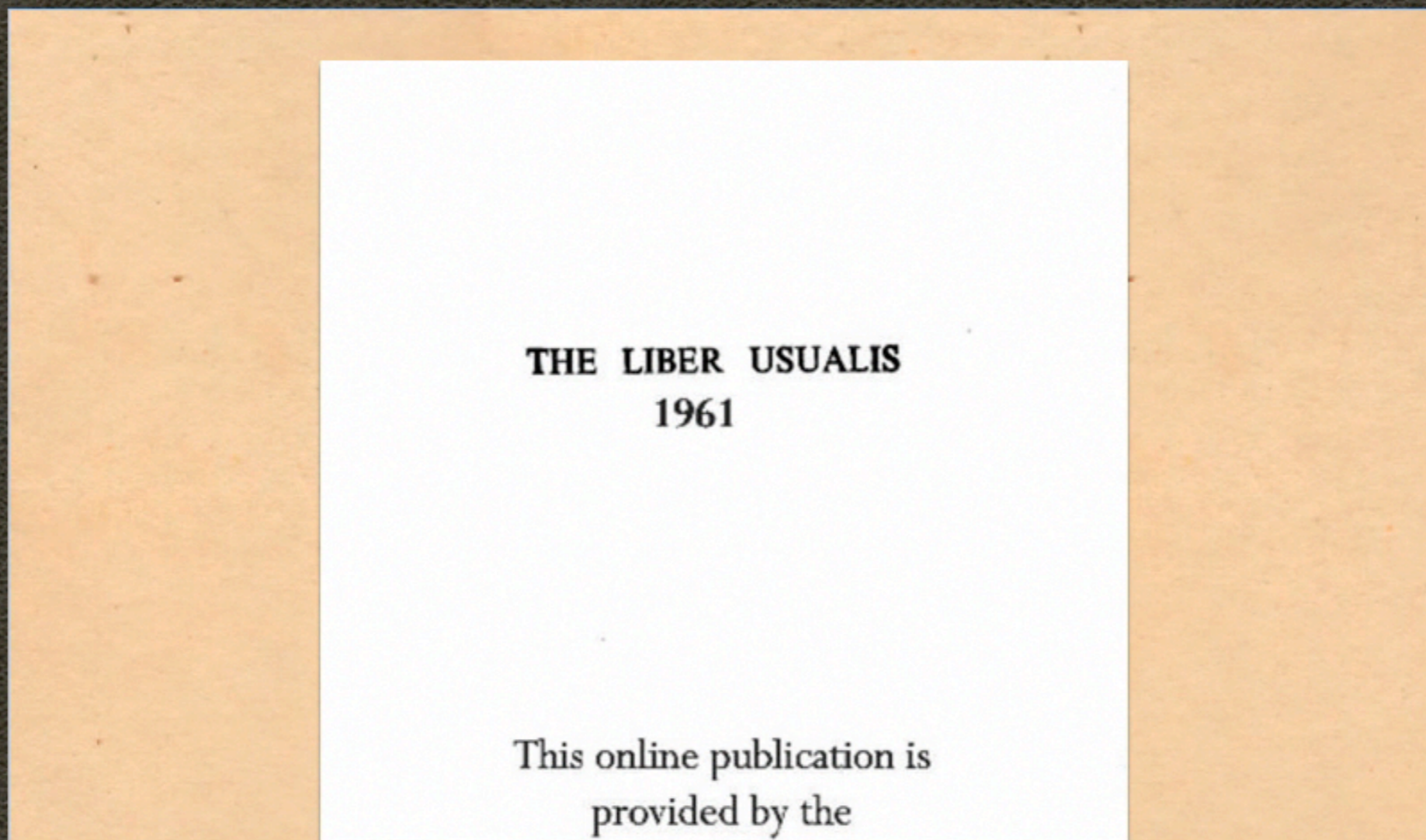
Strict pitch sequence 

Enter a search query



Go to page

Current page: 1 of 2340



Preprocessing: Aruspix

FEAST OF THE BLESSED SACRAMENT.
CORPUS CHRISTI.
Double of the First Class with Octave.
AT FIRST VESPERS.
All as at second Vespers, p. 956, except the following :

At Magn. Ant. 6. F

O quam su-á-vis est, * Dó- mi-ne, spí-
 ri-tus tú- us! qui ut dulcédi-nem tú-am in ff-li-os de-
 monstrá- res, páne su-a-vís-simo de caélo
 praesti-to, esu-ri-éntes réples bónis, fa- stidi- ó- sos dí-
 vi- tes dimít-tens in- á-nes. E u o u a e.

Cant. Magnificat. 6. F. p. 211 or p. 213. Prayer. Deus, qui nobis. p. 943.
At Compline and the Little Hours, the psalm of Sunday are said, the Hymn, today and throughout the Octave, is sung in the tone of Christmas, p. 367, with the doxology Jesu tibi sit glória, Qui natus es de Virgine. All hymns of the same metre end with this doxology throughout the Octave, even on the feasts of Saints, unless the contrary be indicated.

AT MATINS. †
 Pater. Ave María. Credo.

D Omine, lábi- a mé-a apé-ri- es. R̄. Et os mé- um an-

† The Chant for Matins is taken from the publications of Solesmes.

FEAST OF THE BLESSED SACRAMENT.
CORPUS CHRISTI.
Double of the First Class with Octave.
AT FIRST VESPERS.
All as at second Vespers, p. 956, except the following :

At Magn. Ant. 6. F

O mi-ne, spí-
 ri-tus in ff-li-os de-
 monstrá- res, páne su-a-vís-simo de caélo
 praesti-to, esu-ri-éntes réples bónis, fa- stidi- ó- sos dí-
 vi- tes dimít-tens in- á-nes. E u o u a e.

Cant. Magnificat. 6. F. p. 211 or p. 213. Prayer. Deus, qui nobis. p. 943.
At Compline and the Little Hours, the psalm of Sunday are said, the Hymn, today and throughout the Octave, is sung in the tone of Christmas, p. 367, with the doxology Jesu tibi sit glória, Qui natus es de Virgine. All hymns of the same metre end with this doxology throughout the Octave, even on the feasts of Saints, unless the contrary be indicated.

AT MATINS. †
 Pater. Ave María. Credo.

D Omine, lábi- a mé-a apé-ri- es. R̄. Et os mé- um an-

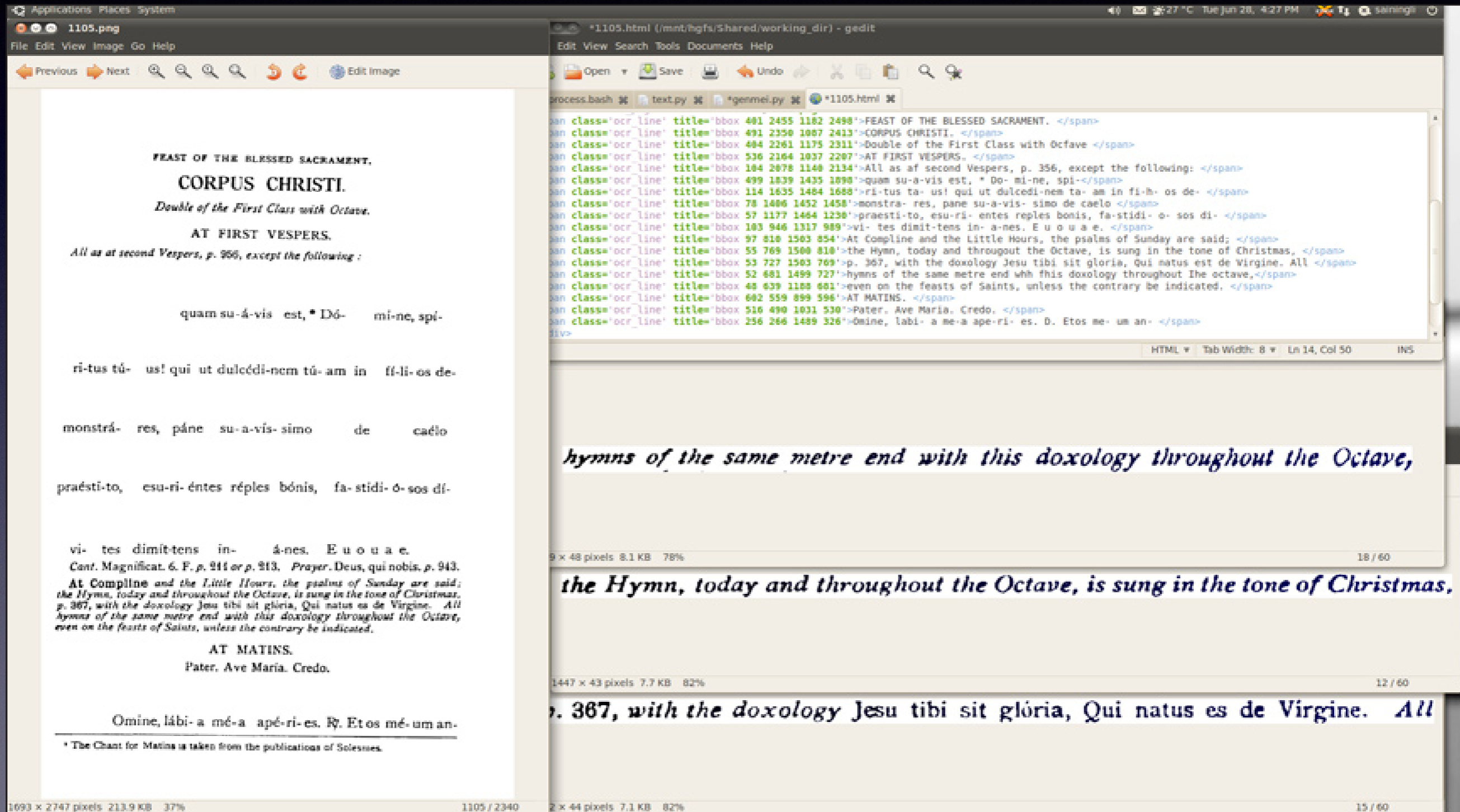
† The Chant for Matins is taken from the publications of Solesmes.

Legend:

- In staff music symbols
- Title elements
- Lyrics or inter staves music symbols
- Ornate letter
- In staff text elements
- Blank

Music recognition: *Gamera*


Text recognition: Ocropus



The image shows a workflow for text recognition using Ocropus. It consists of three main parts:

- Image Input:** A scanned page of a liturgical document titled "FEAST OF THE BLESSED SACRAMENT, CORPUS CHRISTI." The text includes instructions for "AT FIRST VESPERS" and "AT MATINS", along with Latin lyrics such as "quam su-a-vis est, * Dó-mi-ne, spi-ri-tus tú-us!" and "monstrá-res, páne su-a-vis-simo de caelo".
- HTML Output:** A window showing the HTML output of the Ocropus process. Each line of text from the image is wrapped in a `` tag, where the `title` attribute contains the bounding box coordinates for that line.
- Text Selection:** A zoomed-in view of the HTML output where specific text segments are highlighted. These segments correspond to instructions in the original document:
 - hymns of the same metre end with this doxology throughout the Octave,*
 - the Hymn, today and throughout the Octave, is sung in the tone of Christmas,*
 - p. 367, with the doxology Jesu tibi sit glória, Qui natus es de Virgine. All*

Pitch correction: *Aruspix*



Aruspix - Recognition - untitled (91X)

qu^áam su-á-vis est, * Dó-mi-ne, spí-ri-tus tú-us! qui ut dulcédi-nem tú-am in fí-li-os de-monstrá-res, páne su-a-vís-simo de caélo

Neumes Symbols

Web interface: *Diva.js*

FEAST OF THE BLESSED SACRAMENT.


CORPUS CHRISTI.

Double of the First Class with Octave.

AT FIRST VESPERS.

All as at second Vespers, p. 956, except the following :

At Magn. Ant. 6. F



FGACA Strict pitch sequence Search Clear

previous Found 77 results for FGACA next

Cantus Ultimus

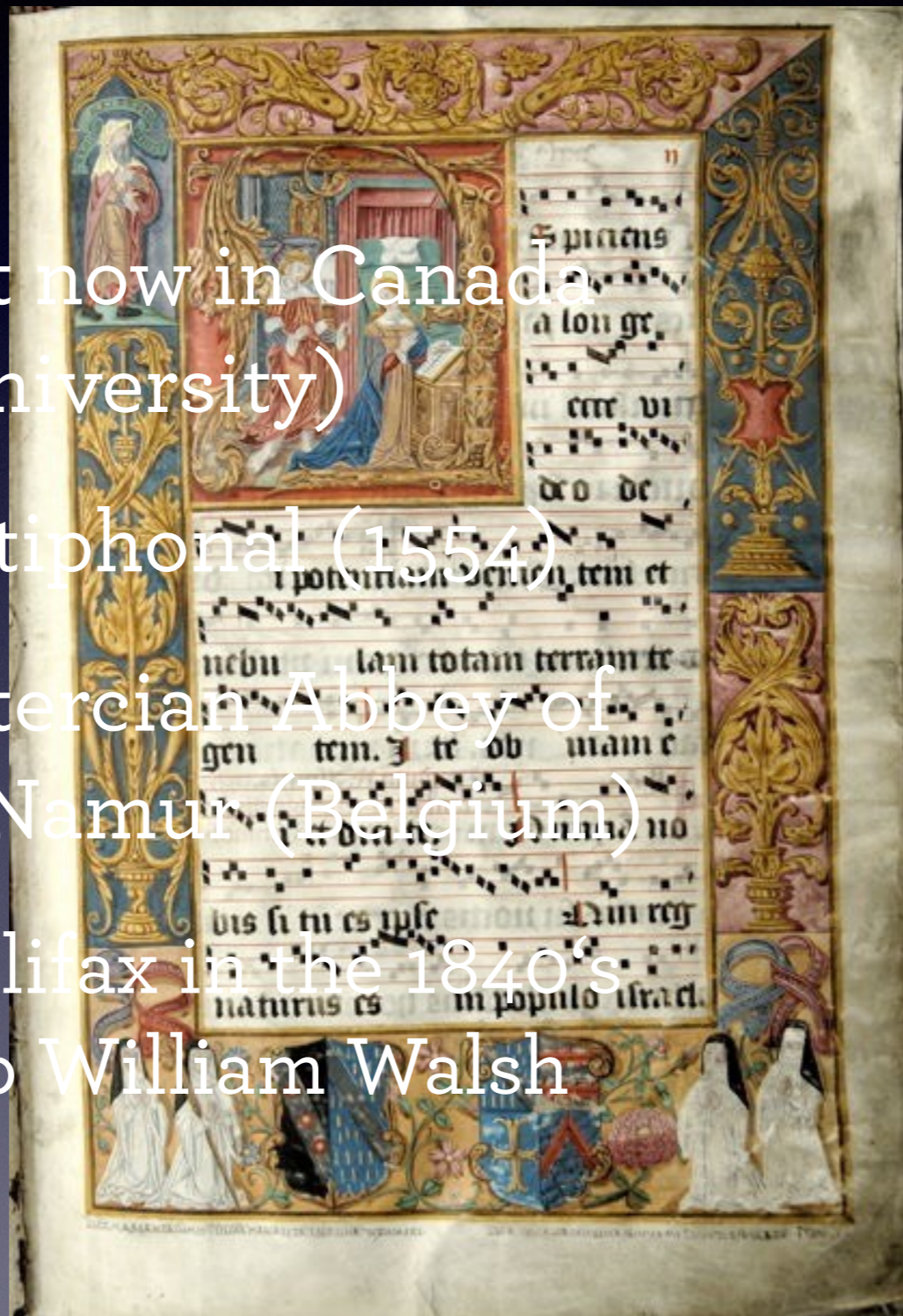
Main Goals of Cantus Ultimus



Roman?
Take a look through presentation slides

2012: *Salzennes* Project

- ❖ A manuscript now in Canada (St. Mary's University)
- ❖ *Salzennes Antiphonal* (1554)
- ❖ From the Cistercian Abbey of *Salzennes* in Namur (Belgium)
- ❖ Brought to Halifax in the 1840's by Archbishop William Walsh



2012: Rodan

Andrew Hankinson

- ❖ Remote Online Document Analysis Network

- ❖ We
rec

- ❖
- ❖
- ❖
- ❖
- ❖
- ❖
- ❖
- ❖
- ❖
- ❖



Rodan: OMR Workflow Management System

Rodan Project Workflow Window Help

Status Users Pages Designer Jobs Results

Workflows	Runs	Pages																																					
Crop, greyscale, Niblack, despeckle	Mon Jun 16 2014 10:33:39 G...	csg-0390_017.jpg	<h3>Workflow Run</h3> <p>URL: Mon Jun 16 2014 11:50:41 GMT-0400 (EDT) Created on: Mon Jun 16 2014 10:33:39 GMT-0400 (EDT) Updated on: http://rodan.simssa.ca/workflowrun/139ed692304d4402a34e8059c820f129/</p> <p>Results Packages Cancel Workflow Run</p> <h3>Run Jobs</h3> <p>Start Job View Error Details View Run Job Settings</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>Sequence</th> <th>Job Name</th> <th>Status</th> <th>Error</th> <th>UUID</th> <th>Page Name</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>Crop Bord...</td> <td>Has finished</td> <td></td> <td>9d65fc9012fe4...</td> <td>csg-0390_110.jpg</td> </tr> <tr> <td>2</td> <td>To Greysc...</td> <td>Has finished</td> <td></td> <td>30179f3e4e504...</td> <td>csg-0390_110.jpg</td> </tr> <tr> <td>3</td> <td>Bernsen T...</td> <td>Has finished</td> <td></td> <td>01ebd132cb88...</td> <td>csg-0390_110.jpg</td> </tr> <tr> <td>4</td> <td>Rdn Desp...</td> <td>Has finished</td> <td></td> <td>bf17ef9e78524...</td> <td>csg-0390_110.jpg</td> </tr> <tr> <td>6</td> <td>Pixel Seg...</td> <td>Has finished</td> <td></td> <td>478ba796135d...</td> <td>csg-0390_110.jpg</td> </tr> </tbody> </table>	Sequence	Job Name	Status	Error	UUID	Page Name	1	Crop Bord...	Has finished		9d65fc9012fe4...	csg-0390_110.jpg	2	To Greysc...	Has finished		30179f3e4e504...	csg-0390_110.jpg	3	Bernsen T...	Has finished		01ebd132cb88...	csg-0390_110.jpg	4	Rdn Desp...	Has finished		bf17ef9e78524...	csg-0390_110.jpg	6	Pixel Seg...	Has finished		478ba796135d...	csg-0390_110.jpg
Sequence	Job Name	Status		Error	UUID	Page Name																																	
1	Crop Bord...	Has finished		9d65fc9012fe4...	csg-0390_110.jpg																																		
2	To Greysc...	Has finished		30179f3e4e504...	csg-0390_110.jpg																																		
3	Bernsen T...	Has finished		01ebd132cb88...	csg-0390_110.jpg																																		
4	Rdn Desp...	Has finished		bf17ef9e78524...	csg-0390_110.jpg																																		
6	Pixel Seg...	Has finished		478ba796135d...	csg-0390_110.jpg																																		
try again: Crop, greyscale Niblack, de...		csg-0390_110.jpg																																					
Testing: crop, greyscale, BERNSEN, ...																																							

Results

<p style="text-align: center; margin-top: 5px;">View Result</p>	<p>Page</p> <p>Page name: csg-0390_110.jpg URL: http://rodan.simssa.ca/page/f7603e9fb397461db367f77849771e5b/</p>
<p style="text-align: center; margin-top: 5px;">View Result</p>	<p>Run Job name: Bernsen Threshold URL: /data/projects/c50c031df54e4e4a986852e715f06ad8/workflows/837f52c Created on: Mon Jun 16 2014 10:43:07 GMT-0400 (EDT) Updated on: Mon Jun 16 2014 10:50:01 GMT-0400 (EDT)</p> <p style="text-align: center; margin-top: 5px;">View Result</p>
<p style="text-align: center; margin-top: 5px;">View Result</p>	<p>Run Job name: Rdn Despeckle URL: /data/projects/c50c031df54e4e4a986852e715f06ad8/workflows/837f52c Created on: Mon Jun 16 2014 10:49:52 GMT-0400 (EDT) Updated on: Mon Jun 16 2014 10:50:08 GMT-0400 (EDT)</p> <p style="text-align: center; margin-top: 5px;">View Result</p>

2016: Breakthrough in OMR preprocessing!

- ❖ Pixel-level classification
 - ❖ Background
 - ❖ Text
 - ❖ Staff lines
 - ❖ Musical symbols
- ❖ Convolution Neural Network
- ❖ Jorge Calvo Zaragoza: “Calvo’s Method”



punctum

d d c dc cb

c dedcd dfd edc

C clef

- Greyscale
- Binarization
- Border Removal
- Lyric Removal
- Staff Removal
- Shape Classification
- Music Reconstruction
- Shape/Image Alignment



Layout Analysis

Calvo's Method

Greyscale

Binarization

Border Removal

Lyric Removal

Staff Removal

Shape Classification

Music Reconstruction

Shape/Image Alignment

Three Different Outputs in One Step! Using Convolutional Neural Networks



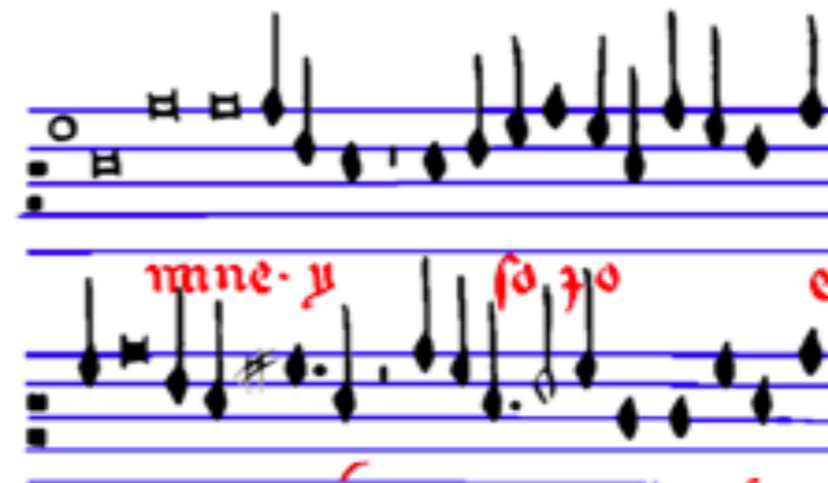
(a) Input image



(b) Binarization



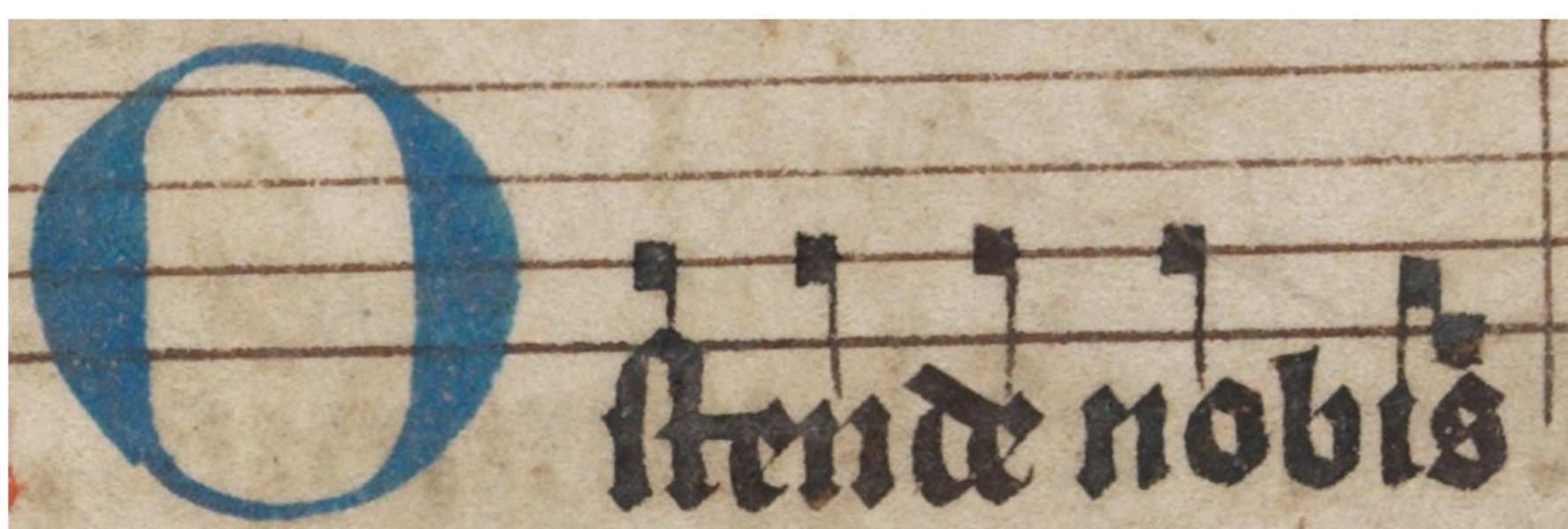
(c) Text detection



(d) Staff-line detection

Calvo's Method

Complete Layout Analysis



(a) Example of input piece of score



(b) Input score after layout analysis

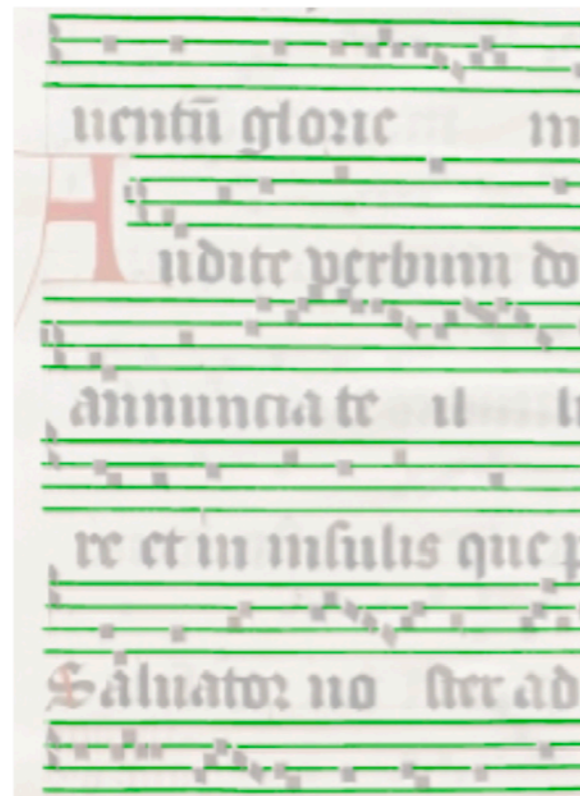
Separation of Staff, Notes, & Text

Jorge Calvo Zaragoza

Actual



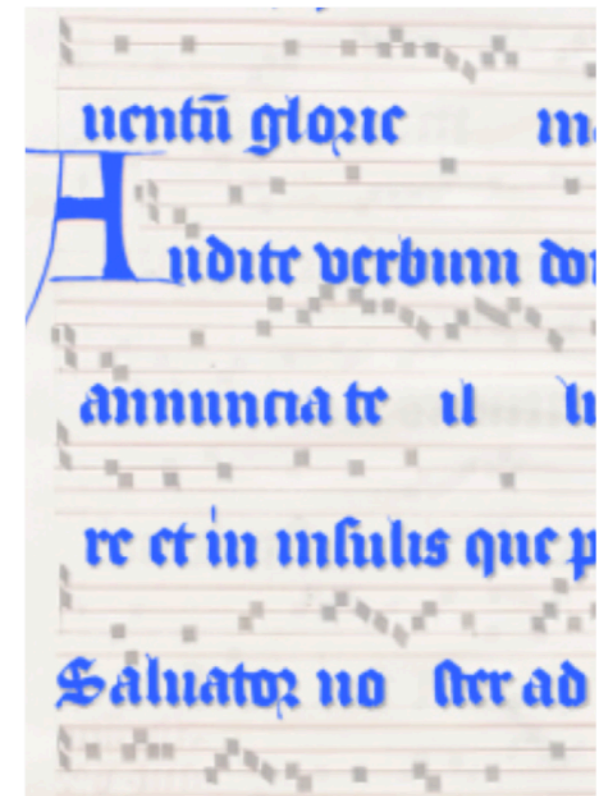
Staff



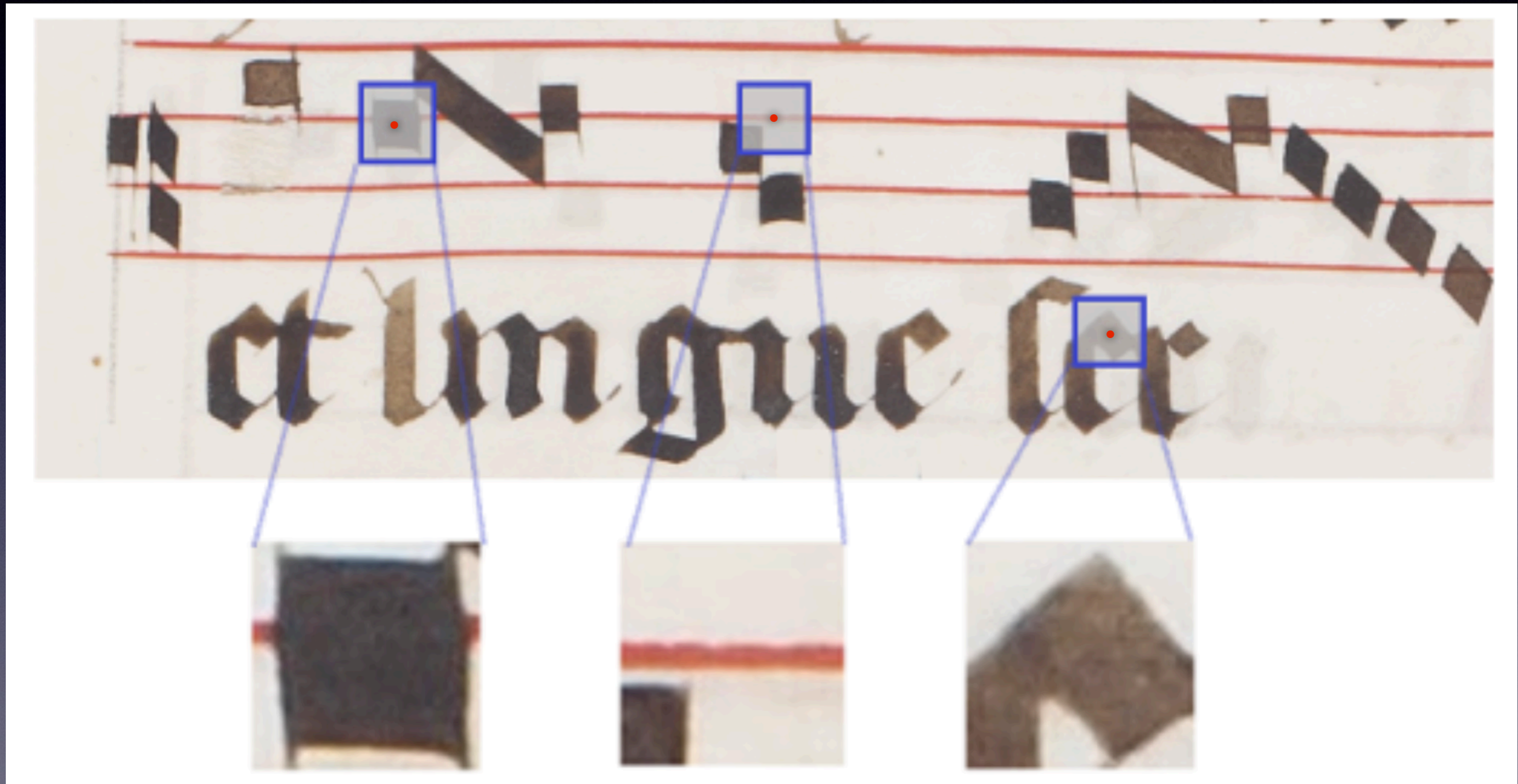
Note



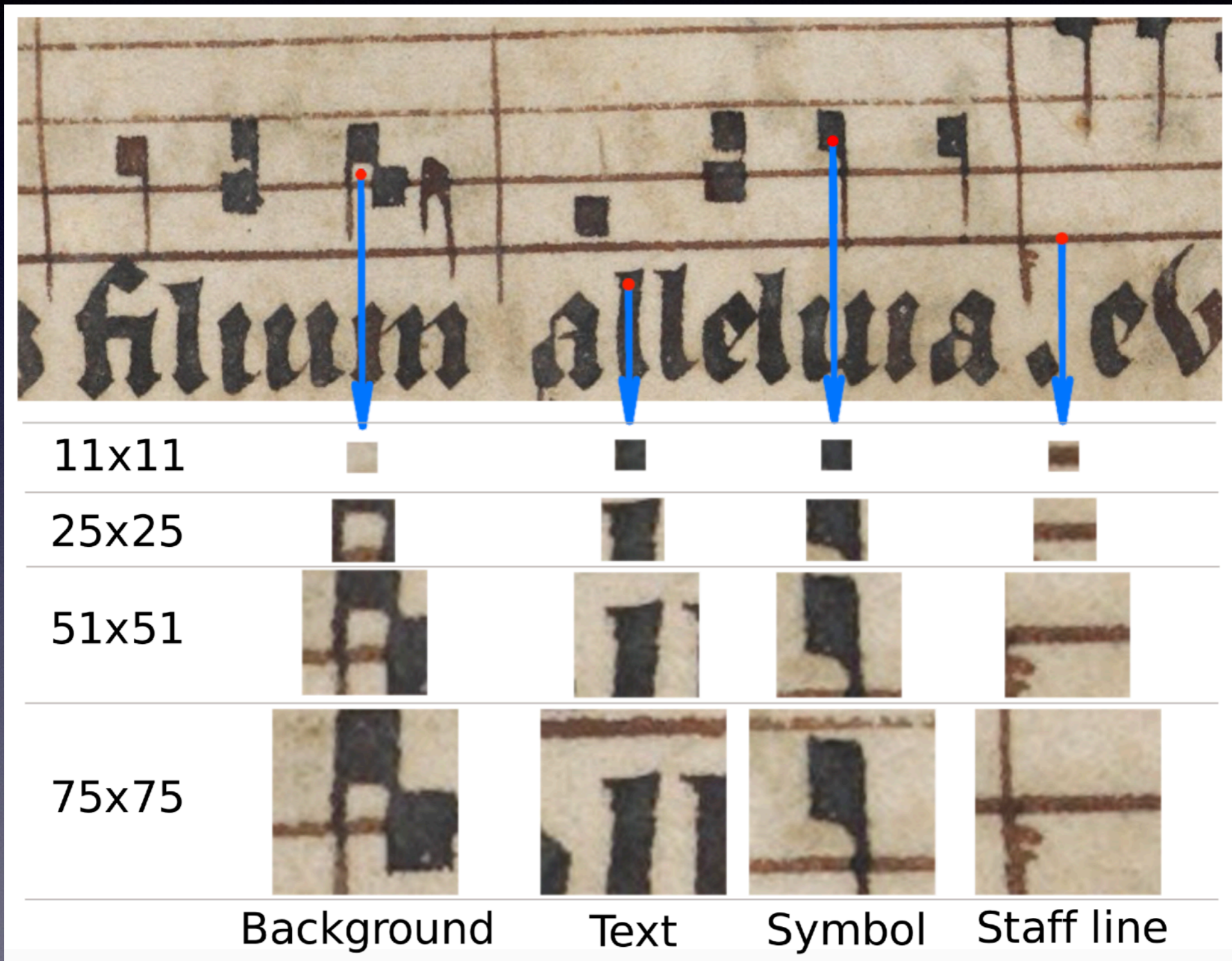
Text



Creating the Ground Truth



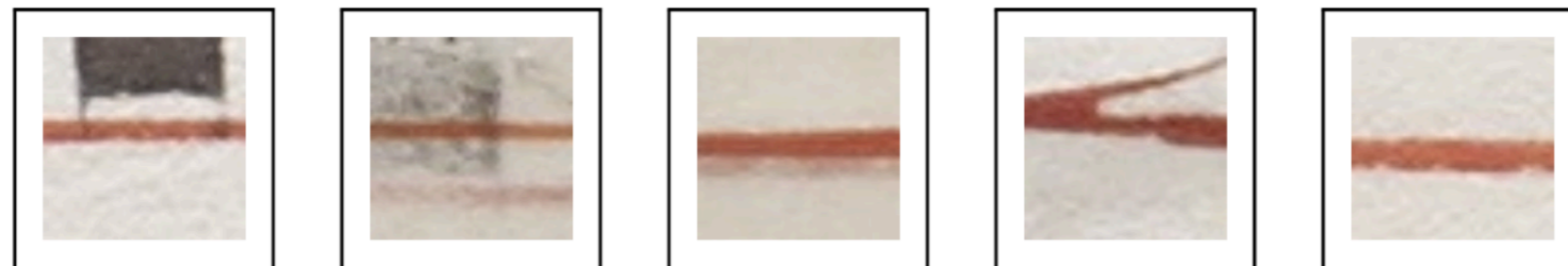
Examples of Different CNN Input Window Size



Samples of different classes



(a) Samples of *background* class



(b) Samples of *staff* class



(c) Samples of *text* class

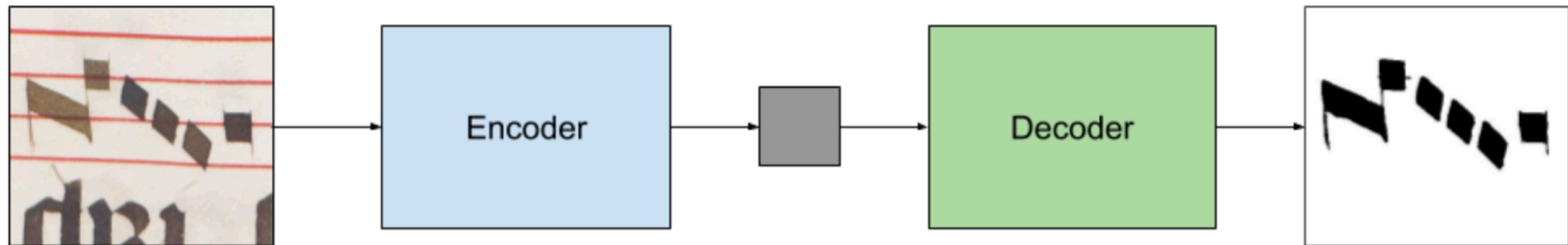
Separation of Staff, Notes, & Text



2018: A Different Neural Network Model

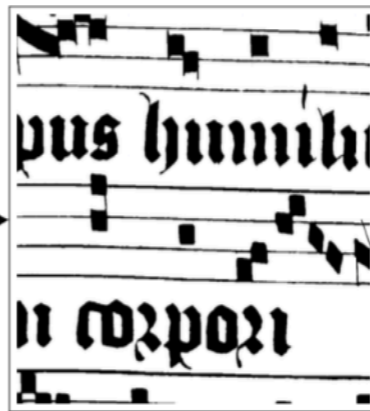
Selective Auto Encoders

Jorge Calvo Zaragoza

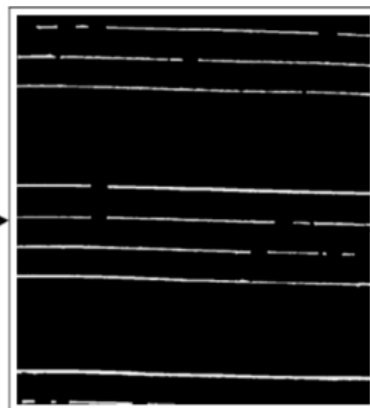
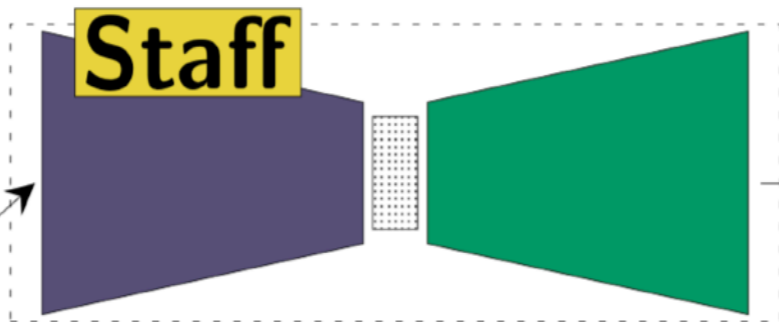


Encoding Decoding

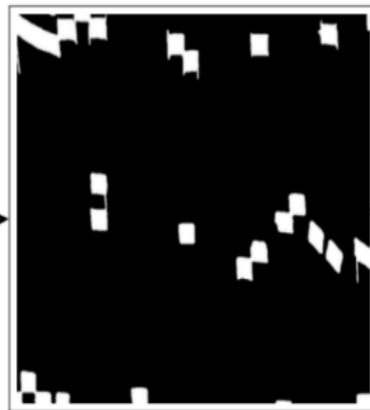
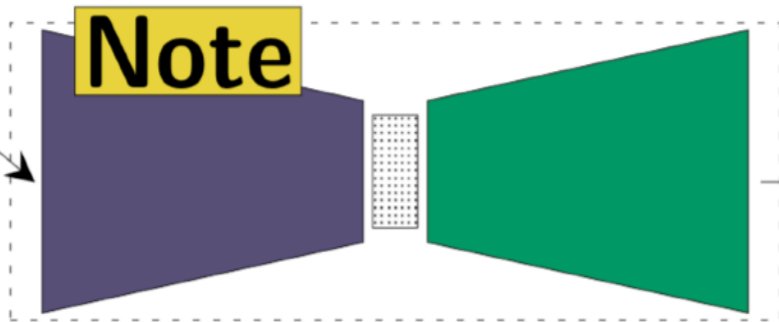
Background



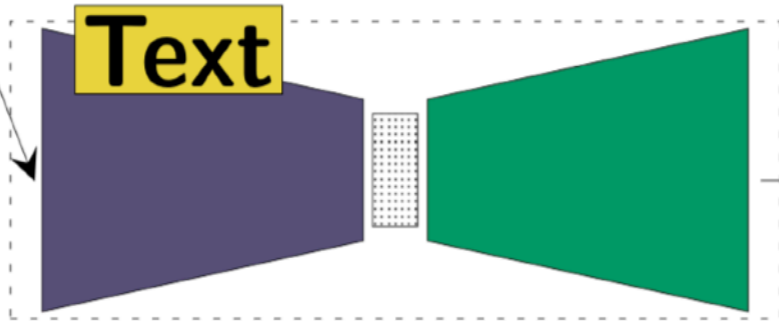
Staff



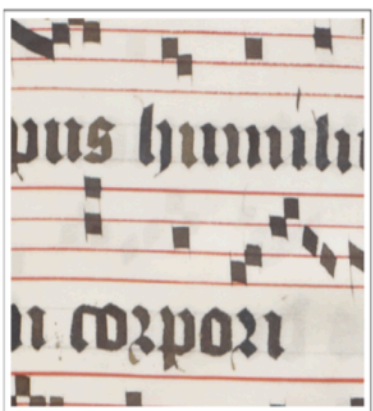
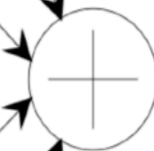
Note



Text



Combination



Accuracy & Training Time Comparison

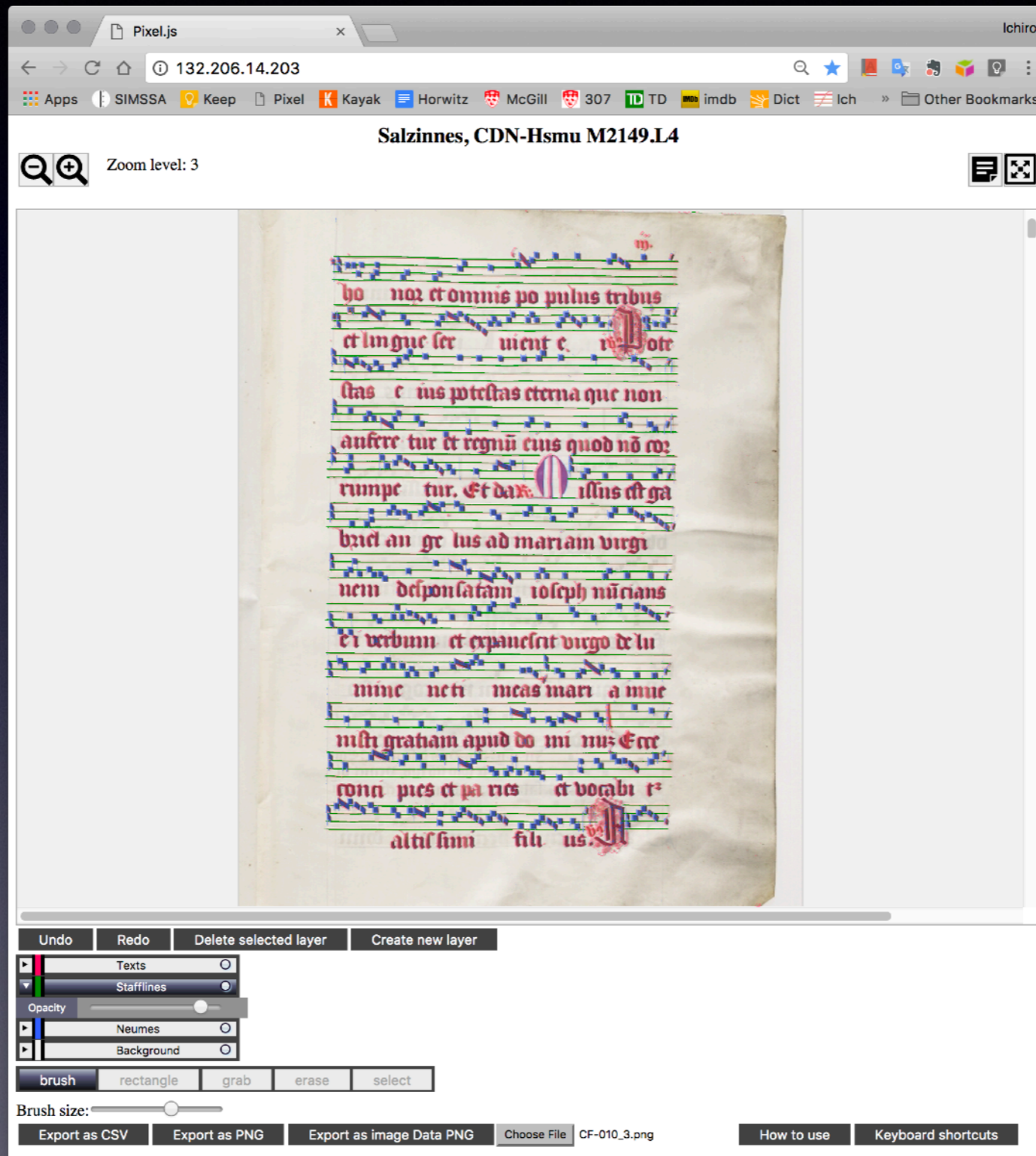
Selective Auto Encoders (SAE) vs Convolutional Neural Nets (CNN)

Two Medieval Manuscripts: Salzinnes & Einsiedeln

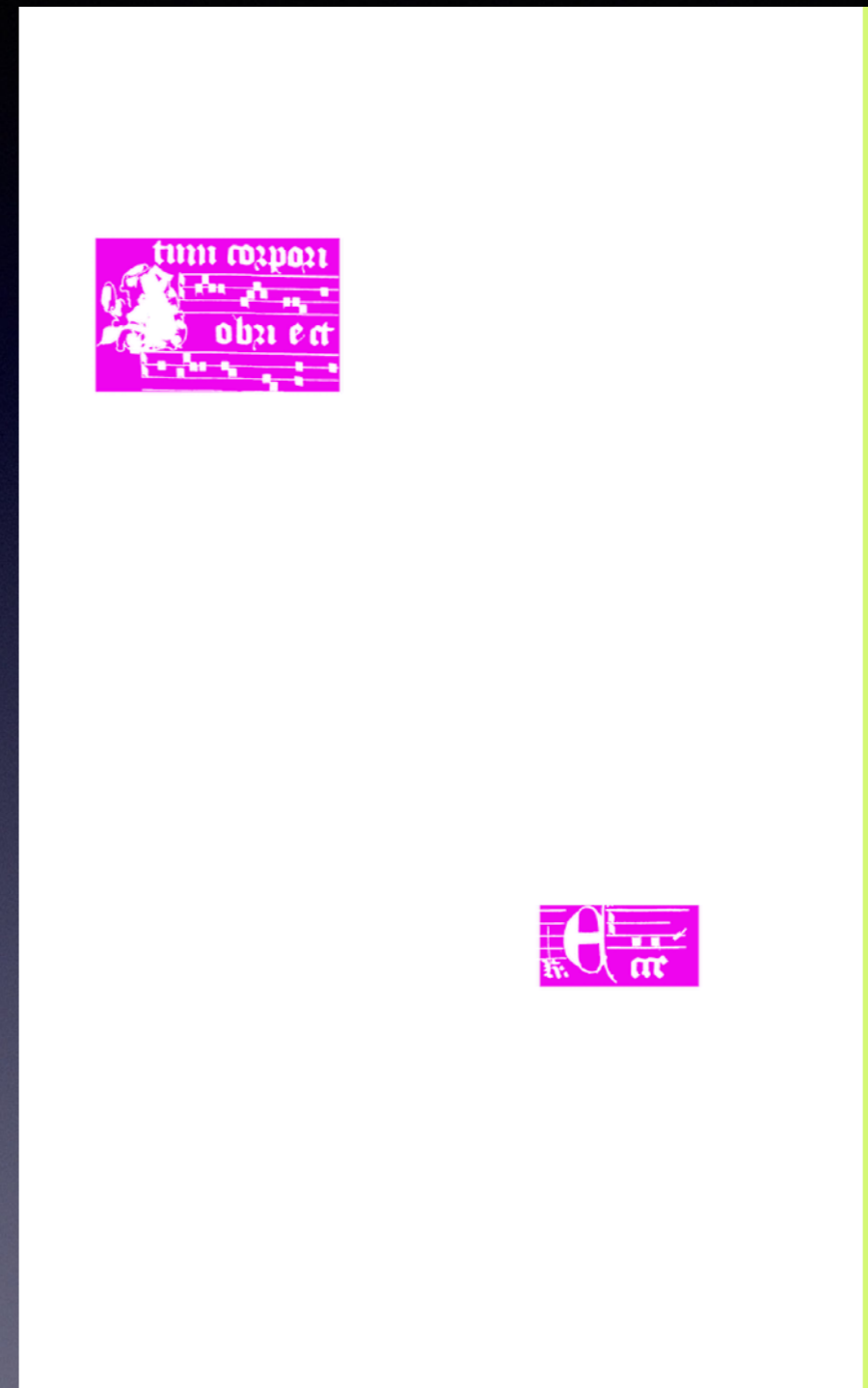
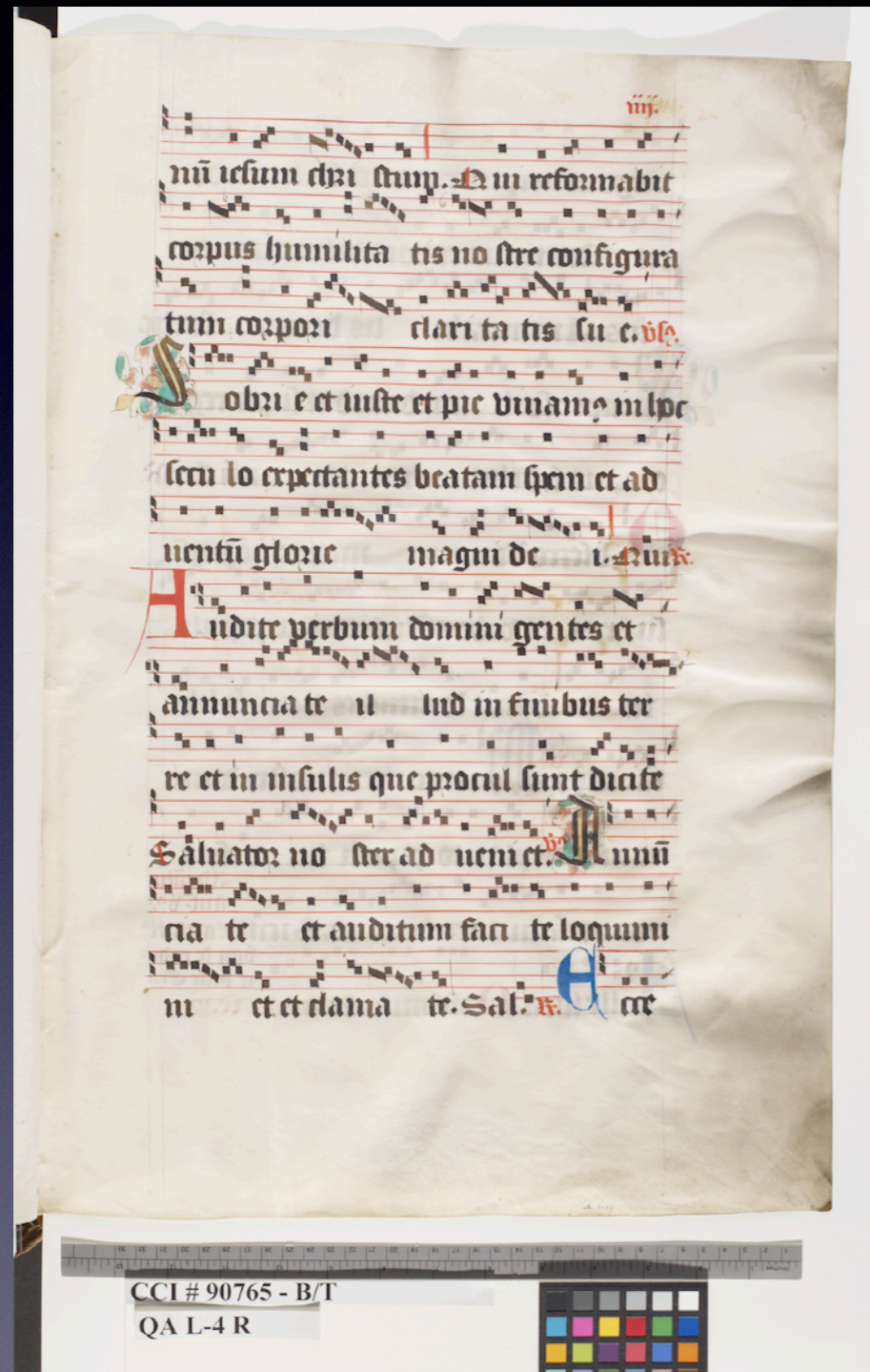
Strategy	Macro F_1		Time per page
	Salzinnes	Einsiedeln	
SAE	96.4	89.3	~ 1 minute
CNN	91.3	88.4	~ 6 hours

Pixel.js: Ground Truth Creator

Zeyad Saleh, Ké Zhang & Eric Liu



Partial Creation of Ground Truth



To classify over 30 million pixels: 3 days (24 hours)!

Ground Truth

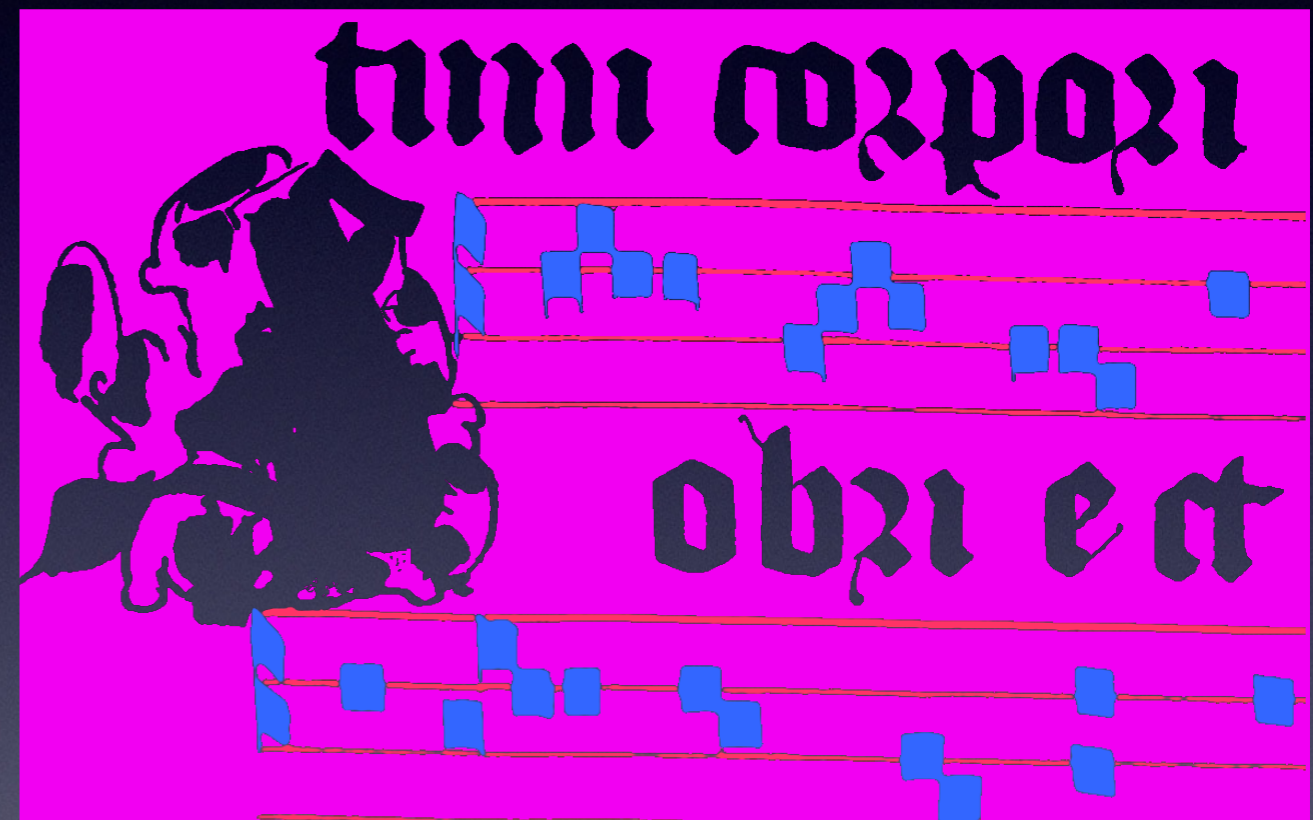
tumi cozpozi

obzi e ot

Original Image & Ground Truth



Original Image



Ground Truth

Classification of a Page: Notes

vi.

flere quia cito veniet salus tu a. Dis.

In laud. Anti. **A**lleluia. ps Misereere mei. *Ve.*

Ueni ad liberandum nos. Domine

deus virtutum. *v.* **E**t ostende faciem

tuam et salvi erimus. Domine. *ymus.*

Splendor paterne. vers. Emitte agnū domine dominatorem terre. De petra deserti ad montem filie syon. *Ad benedictus Antiphona.*

Spiritus sanctus in te descendet ma

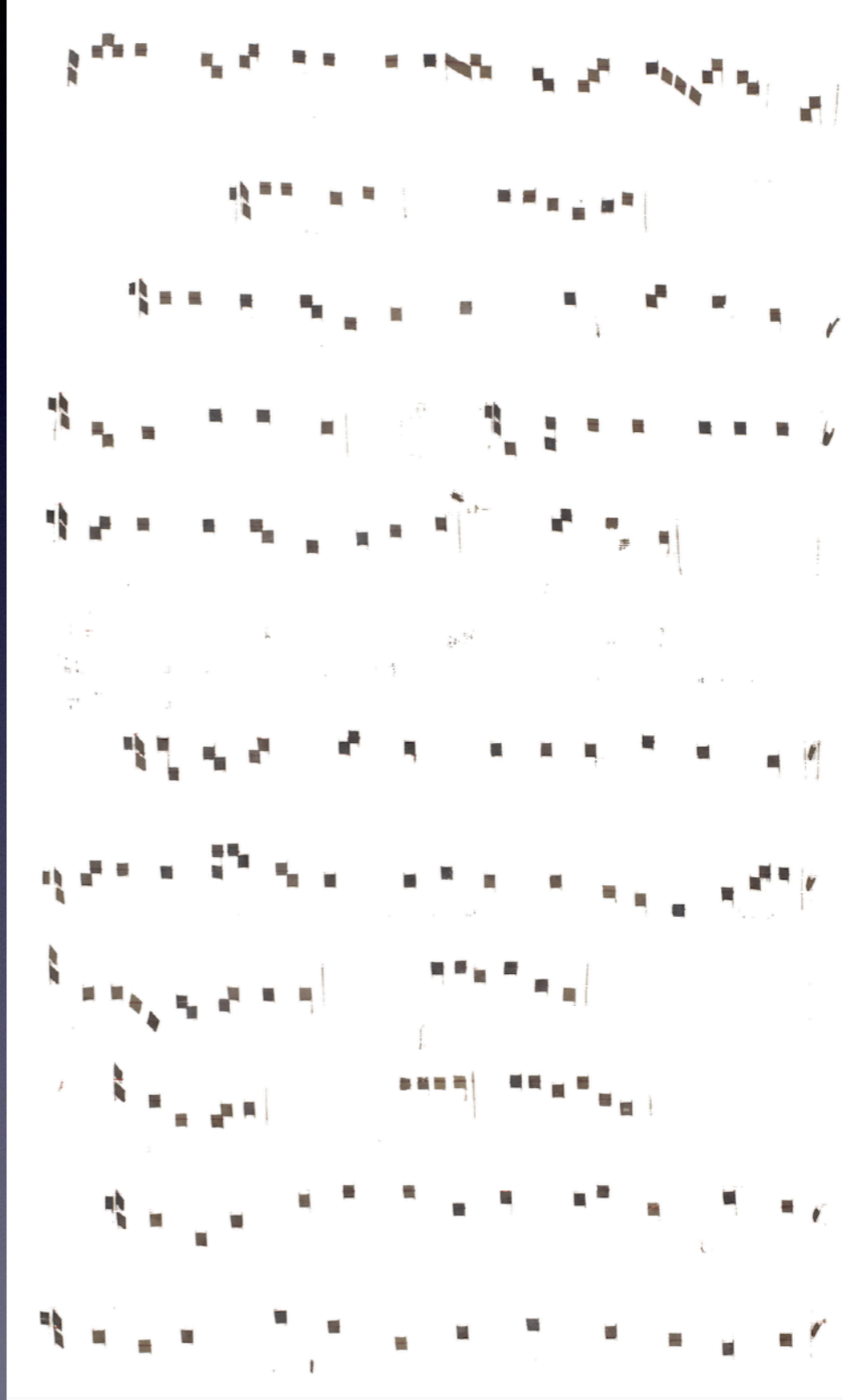
ria ne ti meas habebis in utero filius

dei alleluia. *ca.* Benedictus. *Ad ymnā. au*

Ad m. **A**lleluia. ps. Beati immaculati. *ymus.*

Conditor alme siderum eterna lux cre

dentium christe redemptor omnium ex



Classification of a Page: with Staves

vi.

flere quia cito veniet salus tu a. Dis.

In laud. Anti. **A**lleluia. ps Misere mei. *Ve.*

Ueni ad liberandum nos. Domine

deus virtutum. *v.* **E**t ostende faciem

tuam et salvi erimus. Domine. *Ymnus.*

Splendor paterne. vers. Emitte agnū domine dominatorem terre. De petra deserti ad montem filie syon. *Ad benedictus Antiphona.*

Spiritus sanctus in te descendet ma

ria ne ti meas habebis in utero filius

dei alleluia. *ca.* Benedictus. *Ad ymnus. au*

Alleluia. ps. Beati immaculati. *Ymnus.*

Conditor alme siderum eterna lux cre

dentium christe redemptor omnium ex

Classification of a Page

vi.

flore quia cito veniet salus tu a. Dis.

In laud. Anti. **A**lleluia. ps Misereere mei. *Ve.*

Ueni ad liberandum nos. Domine
 deus virtutum. *v.* **E**t ostende faciem
 tuam et salui erimus. Domine. *ymus.*
Splendor paterne. vers. Emitte agnū domine
 dominatorem terre. De petra deserti ad montem
 filie syon. *Ad benedictus Antiphona.*

Spiritus sanctus in te descendet ma
 ria ne ti meas habebis in utero filius
 dei alleluia. *ca.* Benedictus. *Ad p̄mā. aū*
Ad m.
p totū aduct

Alleluia. ps. Beati immaculati. *ym.*

Qonditor alme siderum eterna lux cre
 dentium christe redemptor omnium ex

vi.

flore quia cito veniet salus tu a. Dis.

In laud. Anti. **A**lleluia. ps Misereere mei. *Ve.*

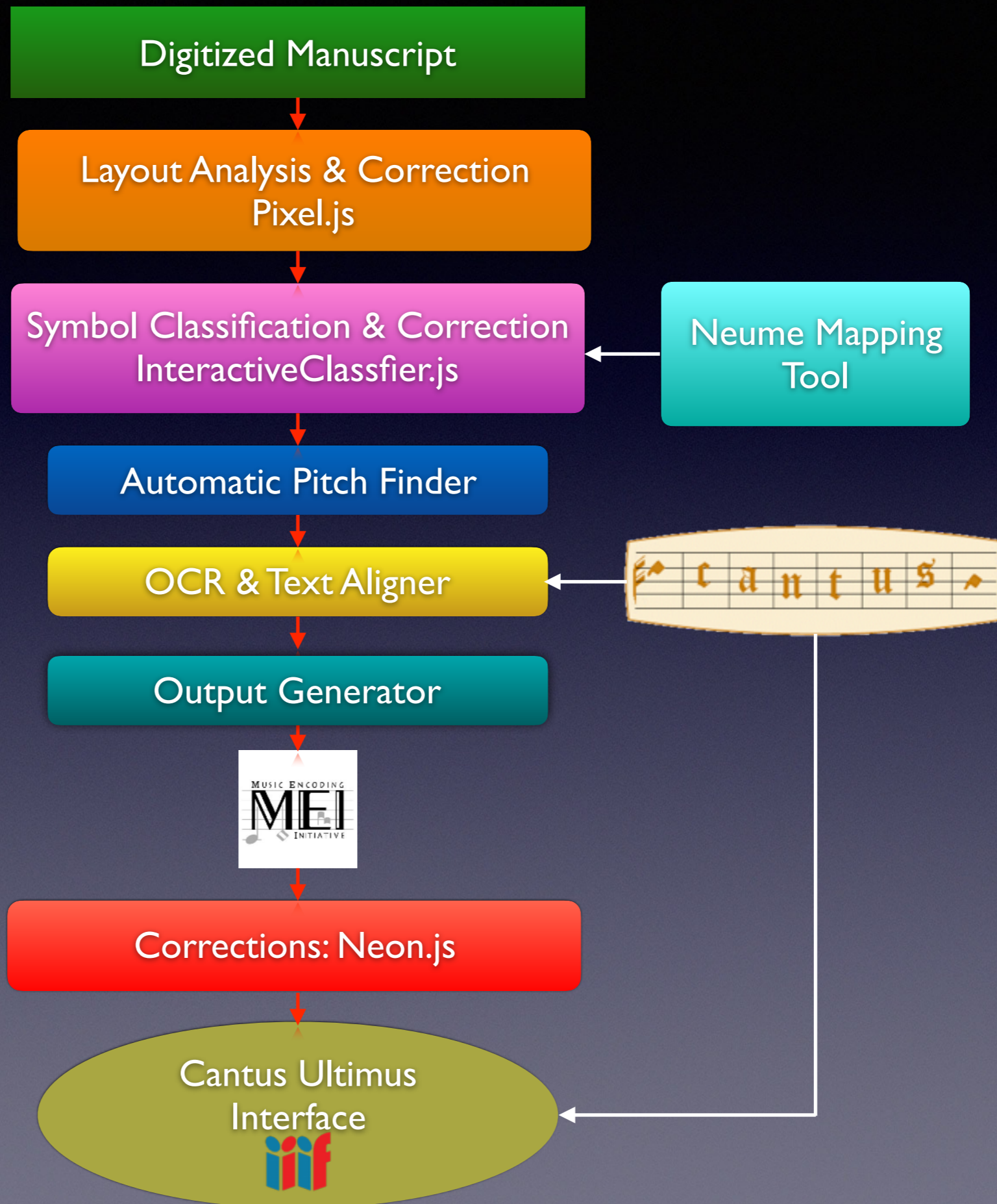
Ueni ad liberandum nos. Domine
 deus virtutum. *v.* **E**t ostende faciem
 tuam et salui erimus. Domine. *ymus.*
Splendor paterne. vers. Emitte agnū domine
 dominatorem terre. De petra deserti ad montem
 filie syon. *Ad benedictus Antiphona.*

Spiritus sanctus in te descendet ma
 ria ne ti meas habebis in utero filius
 dei alleluia. *ca.* Benedictus. *Ad p̄mā. aū*
Ad m.
p totū aduct

Alleluia. ps. Beati immaculati. *ym.*

Qonditor alme siderum eterna lux cre
 dentium christe redemptor omnium ex

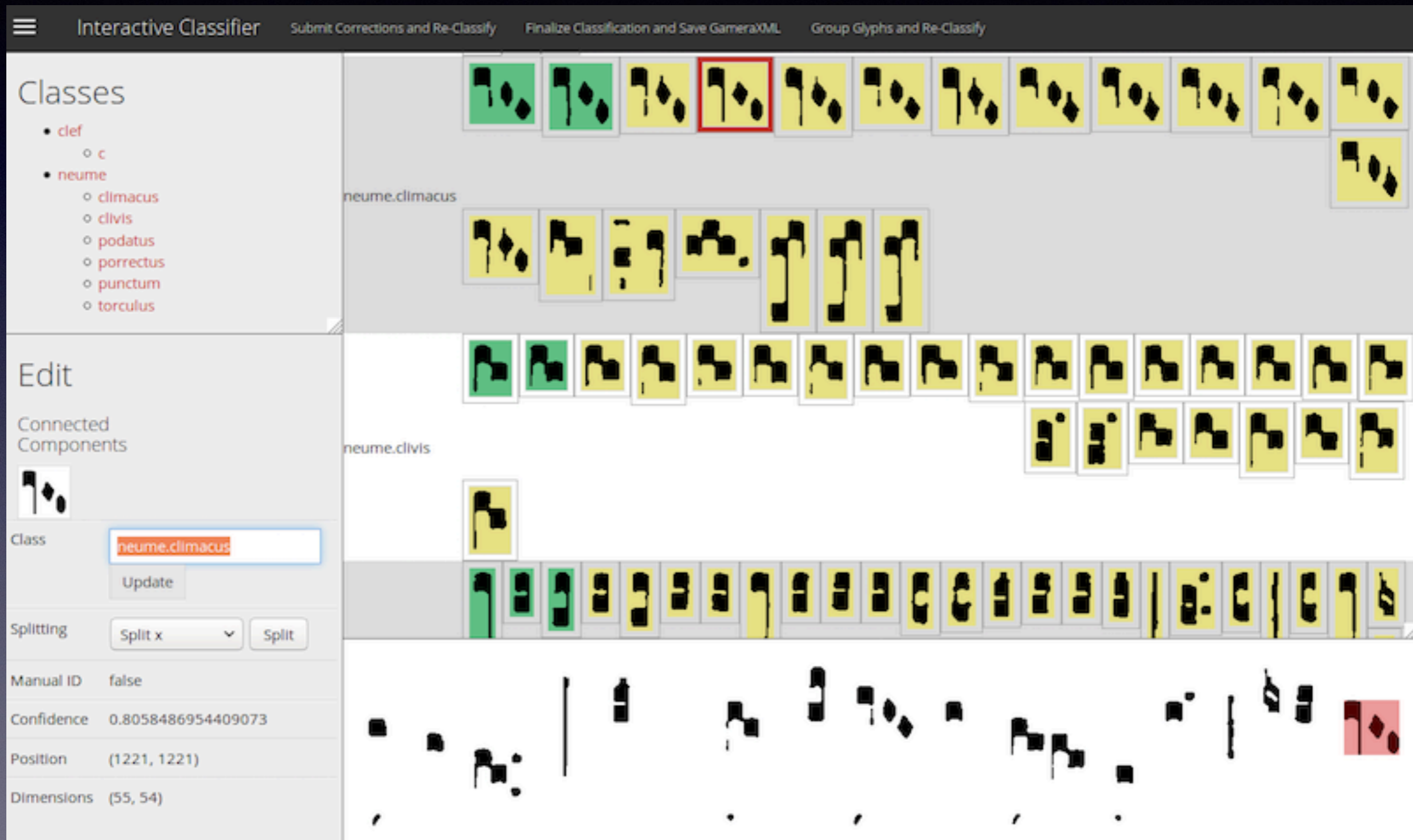
SIMMSA Workflow for Neume Notation



Workflow Management System

InteractiveClassifier.js

Minh Anh Nguyen



The screenshot shows the 'Interactive Classifier' web application interface. At the top, there are navigation buttons: 'Submit Corrections and Re-Classify', 'Finalize Classification and Save GameraXML', and 'Group Glyphs and Re-Classify'. The main area is divided into several sections:




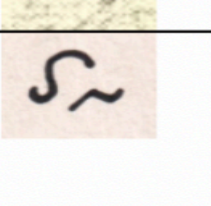
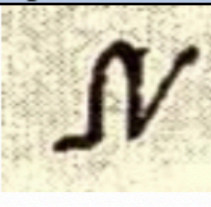
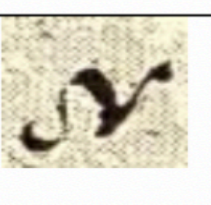
- Classes:** A tree view on the left showing a hierarchy: `def` (expanded) with sub-items `c` and `neume` (expanded). Under `neume`, there are sub-items: `climacus`, `clivis`, `podatus`, `porrectus`, `punctum`, and `torculus`.
- Edit:** A section on the left with a 'Connected Components' icon and a 'Class' dropdown menu currently set to `neume.climacus`. Below the dropdown is an 'Update' button. There are also 'Splitting' controls with a 'Split x' dropdown and a 'Split' button.
- Manual ID:** A field set to 'false'.
- Confidence:** A field showing the value '0.8058486954409073'.
- Position:** A field showing the coordinates '(1221, 1221)'.
- Dimensions:** A field showing the dimensions '(55, 54)'.

The main workspace displays a grid of musical glyphs. The top row shows a sequence of glyphs, with the fourth one highlighted by a red border. Below this, there are sections labeled 'neume.climacus' and 'neume.clivis', each containing a row of glyphs. At the bottom, a larger view shows a sequence of glyphs with a red highlight on the final one.



- ❖ The Music Encoding Initiative (MEI) is a community-driven effort to define a system for encoding musical documents in a machine-readable file format (XML).
- ❖ In development since 1999.
- ❖ MEI is based on Text Encoding Initiative (TEI).
- ❖ MEI is an alternative to MusicXML.

Neume Mapping Table to MEI

		Torculus 3		
3 pitches: n-h-l				
	21,5 (10r) plenitudine m	Torculus	neume.torculus	<pre><neume> <nc tilt= "e" /> <nc angled= "true" tilt= "n" intm= "u" /> <nc tilt= "se" intm= "d" /> </neume></pre>
	21,5 (10r) tua	Torculus	neume.torculus	<pre><neume> <nc curve= "a" /> <nc tilt= "nw" intm= "u" /> <nc curve= "c" intm= "d" /> </neume></pre>
	21,5 (10r) et	Torculus melodic	neume.torculus	<pre><neume> <nc tilt= "e" /> <nc angled= "true" tilt= "n" intm= "u" /> <nc tilt= "se" intm= "d" rel_len= "1" /> </neume></pre>
		Torculus	neume.torculus	<pre><neume> <nc curve= "a" /> <nc tilt= "n" intm= "u" curve= "c" /> <nc con= "g" tilt= "ne" angled= true" intm= "d" /> </neume></pre>
		Torculus resupinus 4		
4 pitches: n-h-l-h				
	21,12 (10r) nostrum	Torculus resupinus	neume.torculus_resupinus	<pre><neume> <nc tilt= "e" /> <nc tilt= "n" angled= "true" intm= "u" /> <nc tilt= "s" intm= "d" /> <nc angled= "true" tilt= "ne" intm= "u" /> </neume></pre>
	22,13 (10v) tribuisti	Torculus resupinus	neume.torculus_resupinus	<pre><neume> <nc curve= "a" /> <nc tilt= "n" intm= "u" /> <nc tilt= "se" intm= "d" /> <nc angled= "true" tilt= "ne" intm= "u" /> </neume></pre>

Neume Mapping Tool

Imane Chafi

[Home](#)
[New neume](#)
[About us](#)
[Contact](#)
[Help](#)
[Log-out](#)
[User1](#)

Punctum

Images:



Name:

Punctum

Folio:

1/r9

Description:

Punctum

Classification:

neume.punctum

Update

Delete

Mei Snippet:

```

1 //mei snippet for neume.punctum
2 <neume>
3 <nc>
4 </neume>
    
```

Pes

Images:



Name:

Pes

Folio:

1r/3, 1r/5

Description:

Classification:

neume.pes.b.3

Update

Delete

Mei Snippet:

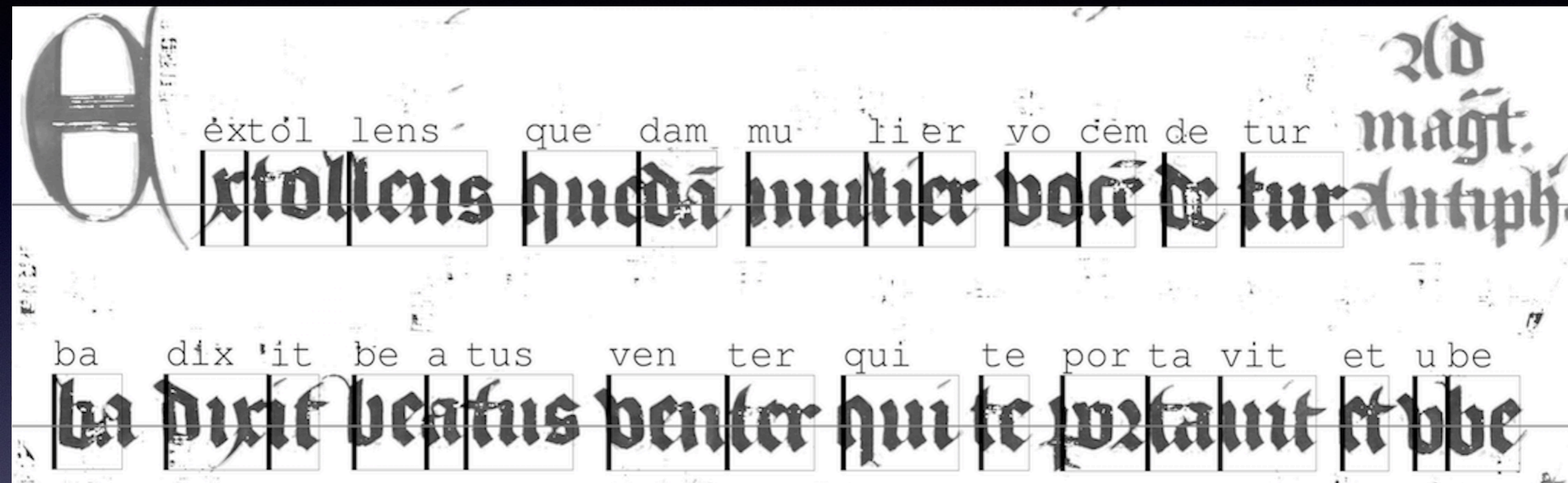
```

1 //mei snippet for neume.pes.b.3
2 <neume>
3 <nc/>
4 <nc intm = "25"/>
5 </neume> <|
6
    
```

element parse error: Error: invalid tagName:

OCR & Text Aligner

Timothy de Reuse



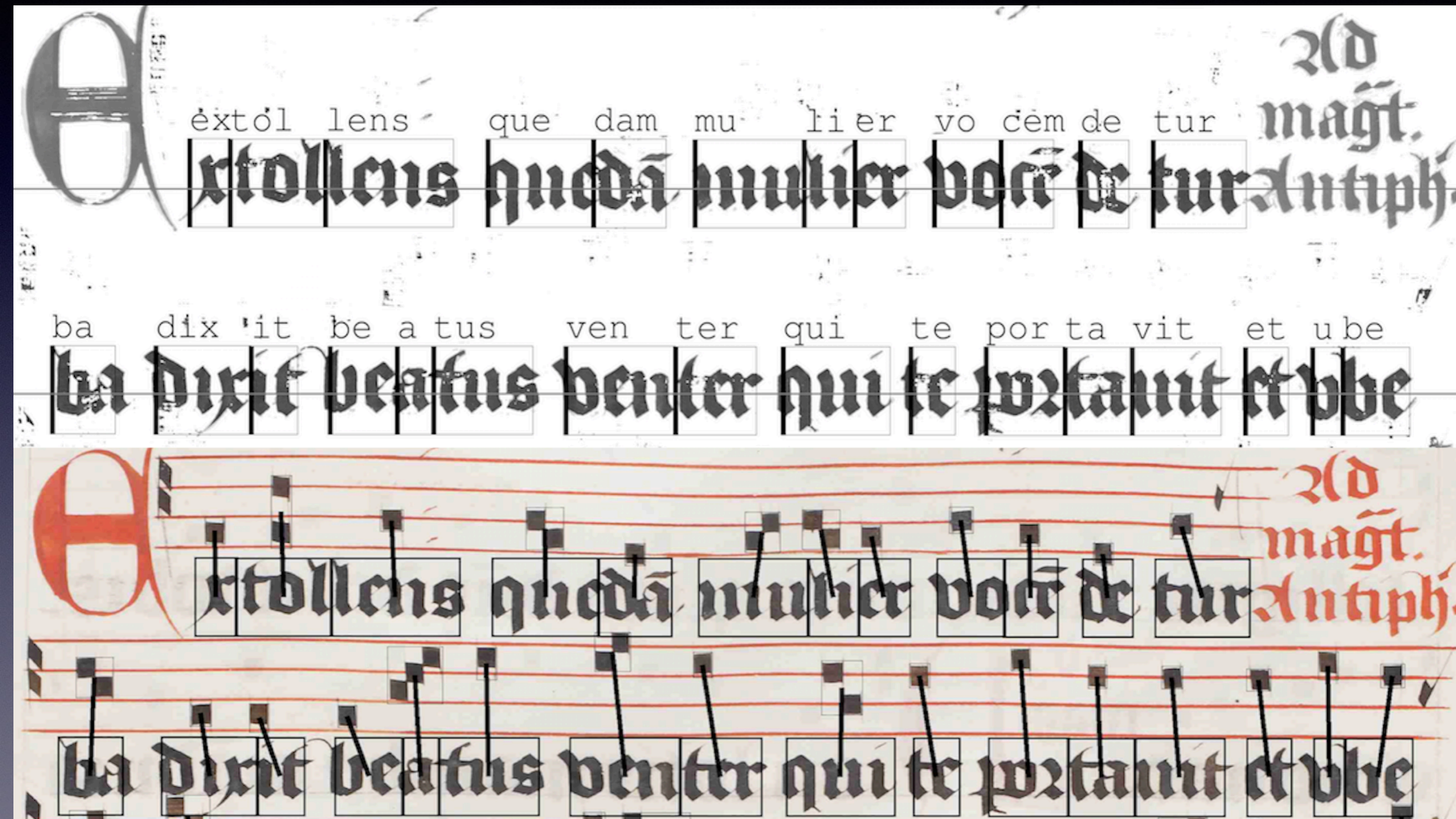
From Cantus Database

Extollens quaedam mulier vocem de tur-
ba dixit beatus venter qui te portavit et ube...

- ❖ OCR used: OCROPUS (recurrent neural network: LSTM)
- ❖ Sequence alignment: Needleman-Wunsch algorithm

OCR & Text Aligner

Timothy de Reuse



The image displays a manuscript page with Latin text and musical notation. The text is presented in two rows, with the first row starting with a large initial 'A'. The text is: **A** extol lens que dam mu li er vo cem de tur magt. Antiph. ba dix it be a tus ven ter qui te por ta vit et ube. The text is overlaid with a grid of boxes, indicating the alignment of the OCR output with the original manuscript text. The musical notation is visible in the background, with red lines and black notes.

Neume Editor ONline: Neon.js

Juliette Regimbal

The screenshot displays the Neon.js Neume Editor interface. The main window shows a manuscript page titled "Salzannes, CDN-Hsmu M2149.L4" with folio "011r". The manuscript text is in Gothic script with neumes above it. The text includes: "omnis homo quia et", "a et. **L**t prepara bi tur in mi", "solum e us et sedebit super", "cans in equita te. **T**uc.", and "gre **A**...". The neumes are represented by colored shapes (green, yellow, blue, orange, pink) on a four-line staff. The interface includes a menu bar with "Neon" and "File", a search bar, a zoom level of 5, and a "Go" button. On the right, a control panel has sections for "Display" (Glyph and Image Opacity sliders), "Display Text" (Highlight - Syllable), "MEI Status: VALID", "Insert" (Neume, Grouping, Clef, System), "Edit" (Select By: Syllable, Neume, Neume Component, Staff), "Delete", "Undo", and "Redo".

Neon.js: Version 3

Juliette Regimbal

What's new

- ❖ Background images displayed with diva.js
 - ❖ IIIF (International Image Interoperability Framework) compliant!
- ❖ Editing via *Verovio* (an online music engraver)
 - ❖ The first version of *Verovio* that is editable!

Neon.js: Text Editing

Caitlin Hutnyk



Display ▼

Zoom 100

Glyph Opacity 100

Image Opacity 100

Display Text: BBoxes: Info: Highlight - Syllable ▼

MEI Status: VALID

Insert ▼

Neume Grouping Clef System

Edit ▼

Select Syllable Neume Neume Component

By: Staff

Egre di tur do mi nus de sa ma ri a ad por tam qu res pi cit le em ad o ri en ten et ve ni et in beth le em am bu lans su per a quas re dem ptio nis u de tunc sal ce vus e rit Et nis ho mo qui a ec ce ve ni et Et pre pa ra bi tur in mi se ri cor dia so li um e ius et se de bit su per il

Cantus Ultimus Interface

The screenshot displays the Cantus Ultimus web interface. At the top, the browser address bar shows the URL `cantus.simssa.ca/manuscript/133/?folio=002r&chant=1`. The page title is "Cantus Ultimus > Salzinnes, CDN-Hsmu M2149.L4". Navigation links for "About", "Activities", "Team", "Manuscripts", and "Search" are visible. The interface includes a zoom control set to "Zoom level: 2" and a "Folio 002r (3 of 479)" indicator. A "Search Manuscript" button is present. The main content area is split into two panels. The left panel shows a high-resolution image of a manuscript page (Folio 002r) featuring a large, ornate initial 'S' and a miniature of a king and queen. The right panel displays the digital transcription of the chant, including the title "Aspiciens a longe ecce video", the Cantus ID (006129), sequence (1), feast (Dom. 1 Adventus), office (Matins), genre (Responsory), and mode (7). The full text of the chant is provided, along with a musical score in G-clef with square neumes on a four-line staff. The lyrics are: "As-pi-ci-ens a lon-ge ec-ce vi-de-o de-i po-tenti-am ve-ni-entem et ne-bu-lam to-tam ter-ram te-gen-tem i-te ob-vi-am ei et di-ci-".

Summary

“A Retrospective on Optical Music Recognition Research”

- ❖ Early developments
- ❖ OMR Thesis
- ❖ Gamera
- ❖ SIMSSA (Single Interface for Music Score Searching and Analysis)
- ❖ Cantus Ultimus

SIMSSA Team @McGill: Summer 2019



Rian Adamian

Caitlin Hutnyk

Cory McKay

Peter Schubert

Imane Chafi

Alessandra Ignesti

Zoé McLennan

Martha Thomae

Julie Cumming

Yaolong Ju

Néstor Nápoles

Andrew Tran

Alex Daigle

Sam Howes

Minh Anh Nguyen

Vi-An Tran

Tim de Reuse

Andrew Kam

Gustavo Pedro

Gabriel Vigliensoni

Glen Ethier

Ian Lorenz

Juliette Regimbal

Emily Hopkins

Sylvain Margot

Evan Savage

Acknowledgements

Christopher Antila
Claire Arthur
William Bain
Ryan Bannon
Laurier Baribeau
Noah Baxter
Laura Beauchamp
Justin Bell
Ruth Berkow
Marina Borsodi-Benson
John Ashley Burgoyne
Greg Burlet
Jorge Calvo-Zaragoza
Remi Chiu
Morgane Ciot
Nat Condit-Schultz
Julie Cumming
Alex Daigle
Marie DeYoung
Tim de Reuse
Natasha Dillabough
Daniel Donnelly
Neda Eshraghi
Meredith Evans
Wei Gao
David Garfinkel

Mahtab Ghamsari
Arielle Goldman
Ryan Groves
Jamie Klassen
Peter Henderson
Jason Hockman
Emily Hopkinson
Andrew Horwitz
Yaolong Ju
Andrew Kam
Anton Khelou
Jamie Klassen
Reiner Krämer
Véronique Lagacé
Saining Li
Eric Liu
Wendy Liu
Evan Magoni
Zoé McLennan
Nicky Mirfallah
Lillio Mok
Alexander Morgan
Catherine Motuz
Maria Murphy
Néstor Nápoles
Clare Neil

Minh Anh Nguyen
Chris Niven
Rory O'Connor
Laura Osterlund
Phyllis Ouyang
Jérôme Parent-Lévesque
Alexandre Parmentier
Gustavo Pedro
Sacha Perry-Fagant
Alastair Porter
Juliette Regimbal
Deepanjan Roy
Zeyad Saleh
Harry Simmonds
Brian Stern
Tristano Tenaglia
Martha Thomae
Andrew Tran
Vi-An Tran
Gabriel Vigliensoni
Tim Wilfong
Mike Winters
Ling-Xiao Yang
Ké Zhang

Cantus Ultimus: cantus.simssa.ca/manuscripts

Project page: simssa.ca

Github sources: github.com/DDMAL

Funded by



Social Sciences and Humanities
Research Council of Canada

Conseil de recherches en
sciences humaines du Canada

Canada



McGill



Schulich School of Music
École de musique Schulich

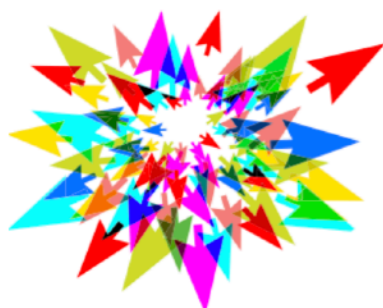


Centre for Interdisciplinary Research
in Music Media and Technology



Fonds de recherche
Société et culture

Québec



compute | calcul
canada | canada



