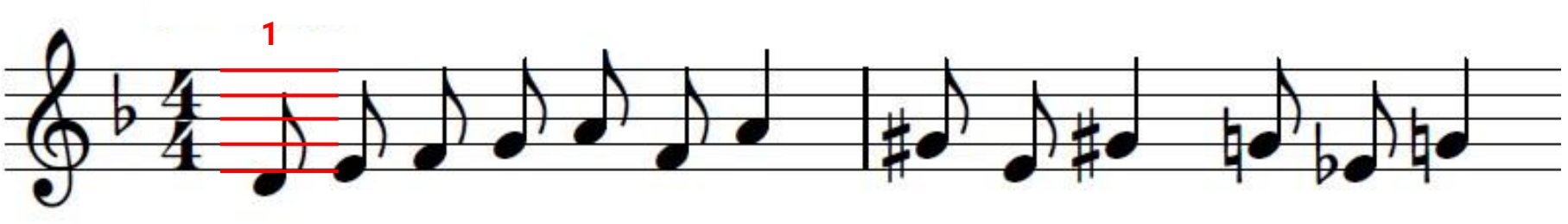
Generating Audio from Sheet Music

John Icke & Megan Shapiro

CS/EE 507: Intro to Computer Vision
Fall 2020

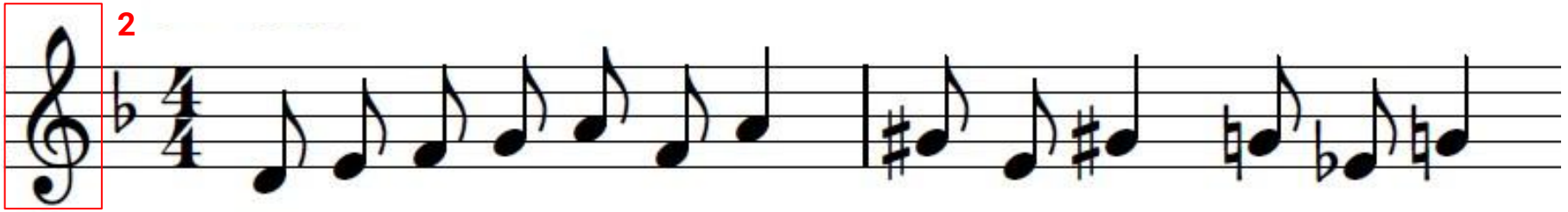
Goals



From an image of a line of sheet music

1. Locate staff lines
2. Detect clef (treble or bass)
3. Detect key (number of flats or sharps)
4. Find note locations in image
5. Find note lengths (⅛, ¼, ½, whole)
6. Find note pitch based on position w.r.t. staff
7. Find accidentals and adjust pitch
8. Play song

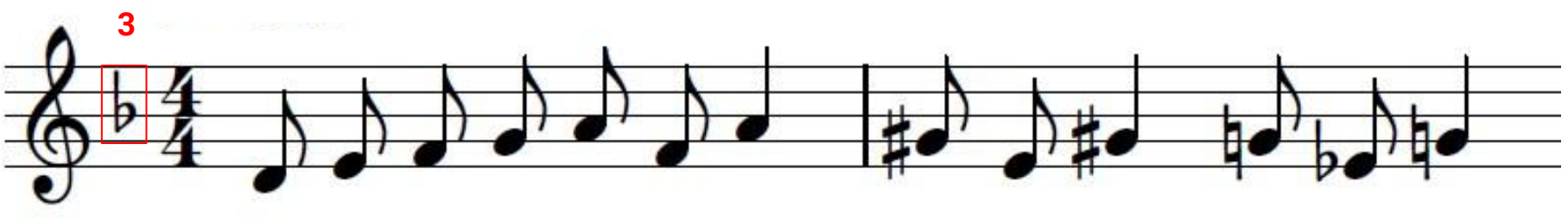
Goals



From an image of a line of sheet music

1. Locate staff lines
2. Detect clef (treble or bass)
3. Detect key (number of flats or sharps)
4. Find note locations in image
5. Find note lengths (⅛, ¼, ½, whole)
6. Find note pitch based on position w.r.t. staff
7. Find accidentals and adjust pitch
8. Play song

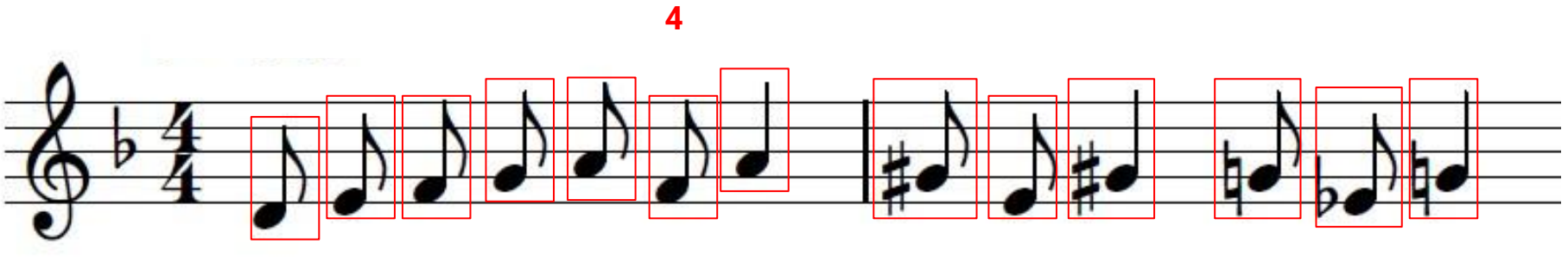
Goals



From an image of a line of sheet music

1. Locate staff lines
2. Detect clef (treble or bass)
3. Detect key (number of flats or sharps)
4. Find note locations in image
5. Find note lengths (⅛, ¼, ½, whole)
6. Find note pitch based on position w.r.t. staff
7. Find accidentals and adjust pitch
8. Play song

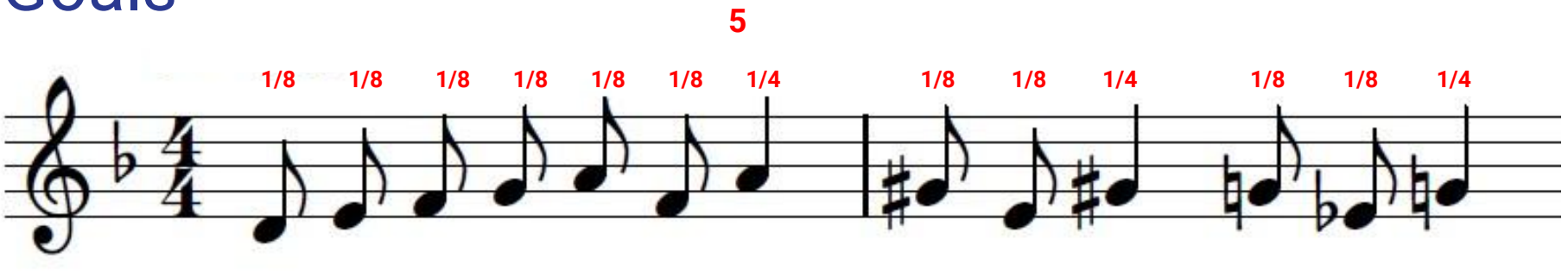
Goals



From an image of a line of sheet music

1. Locate staff lines
2. Detect clef (treble or bass)
3. Detect key (number of flats or sharps)
4. Find note locations in image
5. Find note lengths (⅛, ¼, ½, whole)
6. Find note pitch based on position w.r.t. staff
7. Find accidentals and adjust pitch
8. Play song

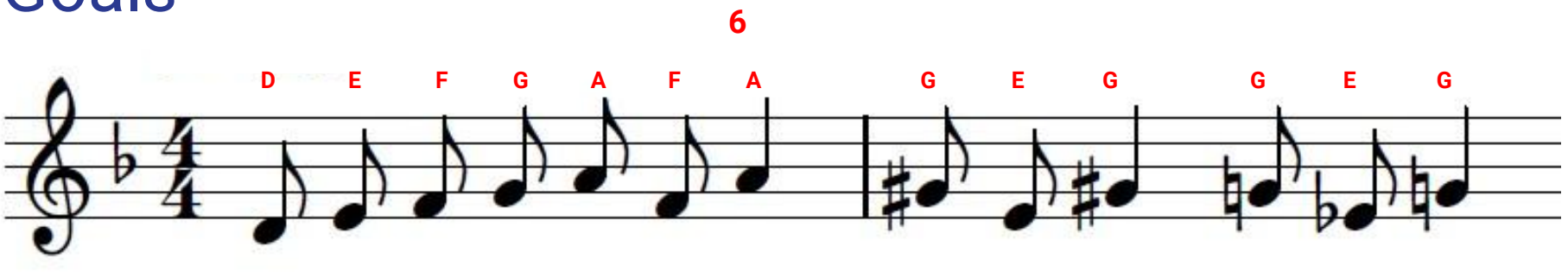
Goals



From an image of a line of sheet music

1. Locate staff lines
2. Detect clef (treble or bass)
3. Detect key (number of flats or sharps)
4. Find note locations in image
5. Find note lengths (⅛, ¼, ½, whole)
6. Find note pitch based on position w.r.t. staff
7. Find accidentals and adjust pitch
8. Play song

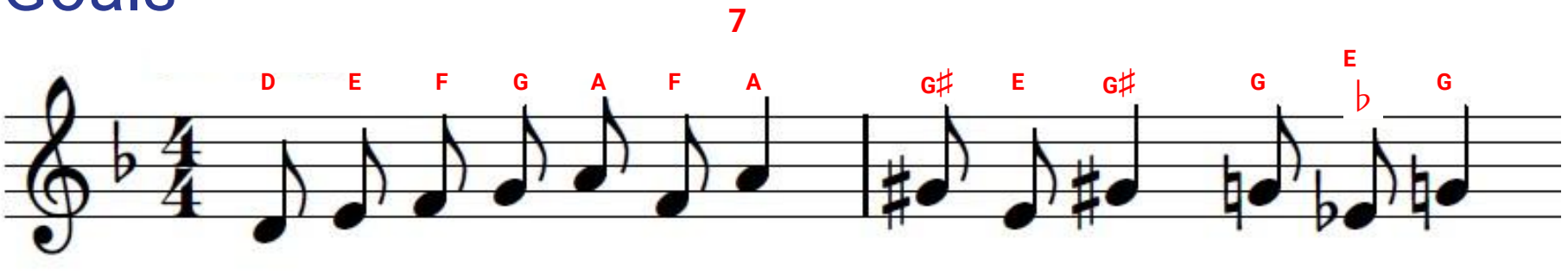
Goals



From an image of a line of sheet music

1. Locate staff lines
2. Detect clef (treble or bass)
3. Detect key (number of flats or sharps)
4. Find note locations in image
5. Find note lengths (⅛, ¼, ½, whole)
6. Find note pitch based on position w.r.t. staff
7. Find accidentals and adjust pitch
8. Play song

Goals



From an image of a line of sheet music

1. Locate staff lines
2. Detect clef (treble or bass)
3. Detect key (number of flats or sharps)
4. Find note locations in image
5. Find note lengths (⅛, ¼, ½, whole)
6. Find note pitch based on position w.r.t. staff
7. Find accidentals and adjust pitch
8. Play song

Previous Work

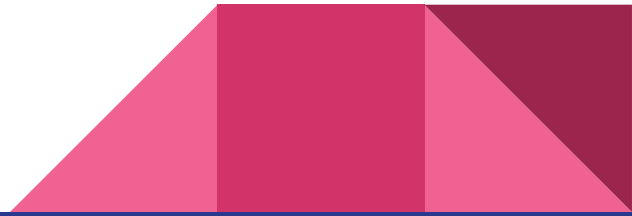
- Optical Music Recognition
  - Heavier focus on handwriting interpretation
    - Digitization, storage, duplication, preservation
  - Wide Range of techniques
    - Template Matching
    - Machine Learning
  - Typically done in three steps
    - Preprocessing
    - Symbol recognition
    - Reconstruction (Sounding the note)

Technical Approach

- Preprocessing
  - Thresholded using an inverse binary threshold
  - Staff lines detected and removed using HoughLinesP() function
- Symbol Detection
  - Clef detected using template matching (which match was stronger, treble or bass)
  - Key template matching (which was stronger, flat or sharp, and how many of them are present)
  - Note locations using connected components
  - Note type using blobs (is the note head circle filled?) and then templates (⅛ vs ¼ have filled circles, ½ vs whole have empty circles)
  - Accidentals using same templates as key signature

Technical Approach

- Reconstruction
  - Note position compared with staff line heights to determine note location
  - Note location converted to pitch using predefined arrays according to key signature
  - Pitches played using musicalbeeps library



Results

Live Demo

Going Further

- Tempo
- Time Signatures
- Measure Markers
- Rests
- Dynamics
- Articulation
- Multiple different Instruments