Optical Music Recognition with Human Labeled Constraints

Christopher Raphael, Liang Chen, Yucong Jiang, Rong Jin, Erik Stolterman. Indiana University Bloomington

What is Optical Music Recognition (OMR)

OMR system is aimed to convert digital music score images to symbolic music data (e.g. MusicXML or MIDI).

Existing Optical Music Systems:

Notescan, Midiscan, Photoscore, Smartscore, Sharpeye, Audiveris, etc.

Challenges:

1, A heavy tail of rare symbols and execptional configurations that downgrade the performance of automatic recognition.
2, Lack of paradigm to identify measure-level symbol structures.
3, Hard to incorporate higher-level music knowledge into recognition.
4, Image degradation due to noise, distortion, or illumination.

Current state of the art:

1, Unstable performance in terms of accuracy and usability.
2, Require a considerable amount of human labor for post-processing.

Model for Automatic Recognition:

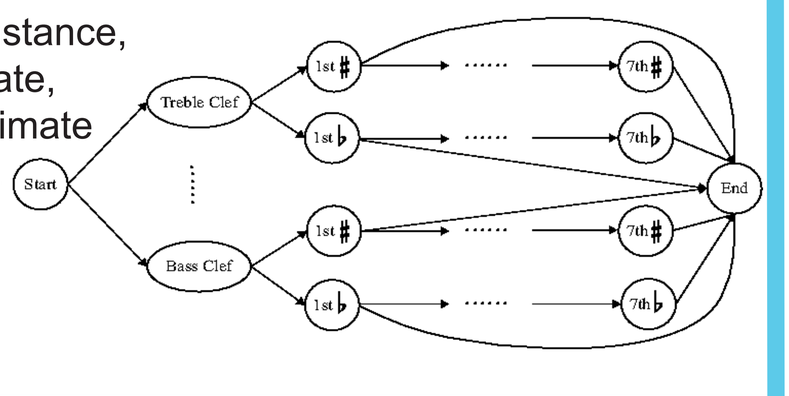
1, Template Matching for symbol primitive detection

Each primitive template is partitioned into several different states: inside (black), outside (white) , or unknown (background), etc.
We evaluate the goodness of a primitive model by measuring the likelihood normalized by background model.

$$H(B,W) = \sum_{x \in B} \log \frac{p_B(g(x))}{p_U(g(x))} + \sum_{x \in W} \log \frac{p_W(g(x))}{p_U(g(x))}$$

2, Represent composite symbols via graphical model

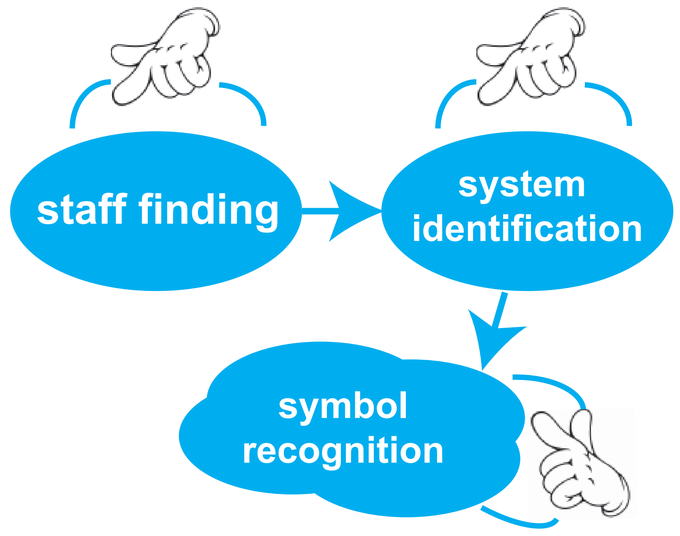
In the clef-key graph, for instance, each node represents a state, and the arrows define legitimate transitions between states.



What we are doing

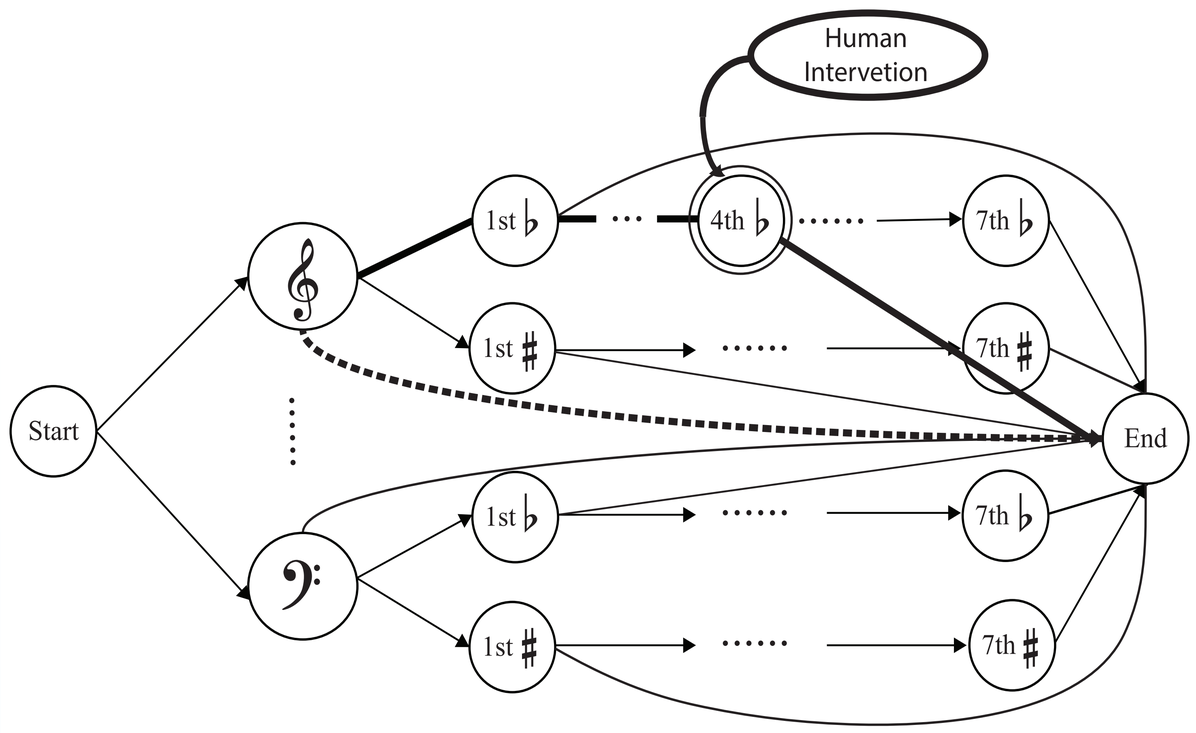
Instead of pinning our hope on fully automatic system, we design an interactive one which combines the power of both human and machine intelligence.

We construct a computational loop, in which machine takes human feedback as extra contraints for re-recognition. This loop continues until it converges to the correct recognition result.

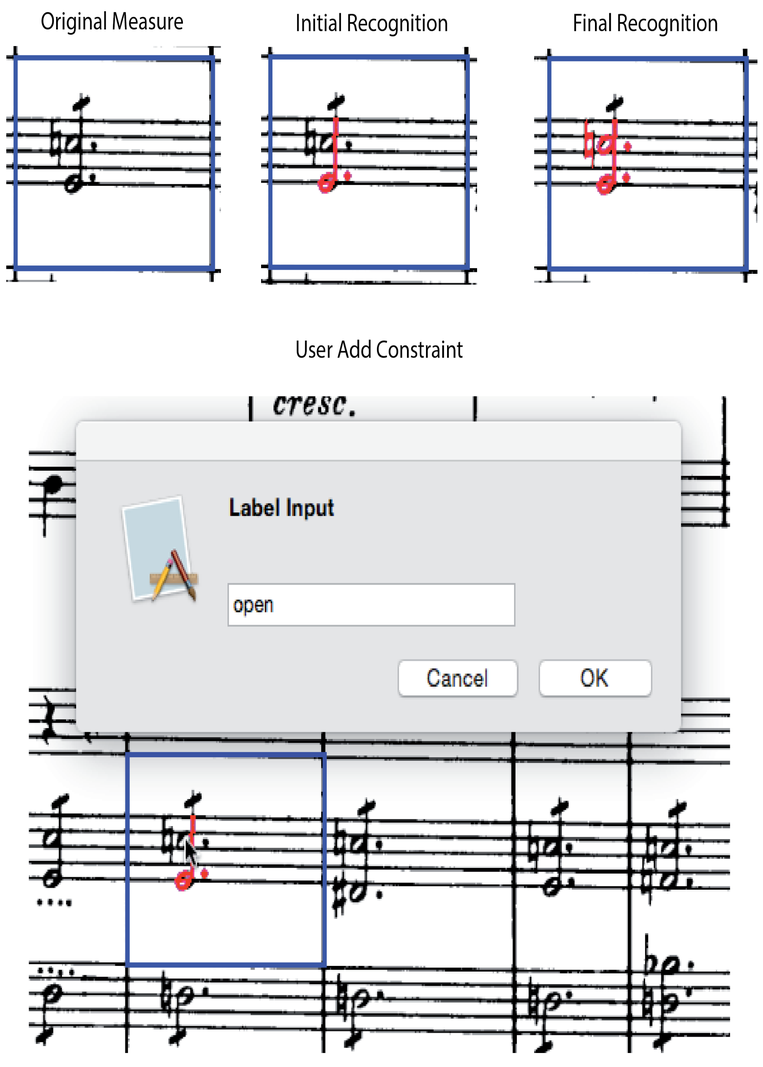


Idea of human-in-the-loop recognition

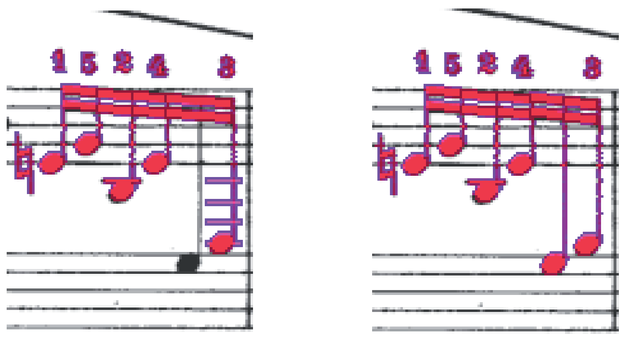
1, Label Constraints: user imposes pixel labels on the image, and machine re-recognize symbols subject to these constraints.
2, Model Constraints: user toggles switches on the interface making recognition engine change its underlying symbol graph.



Example of Label Constraint



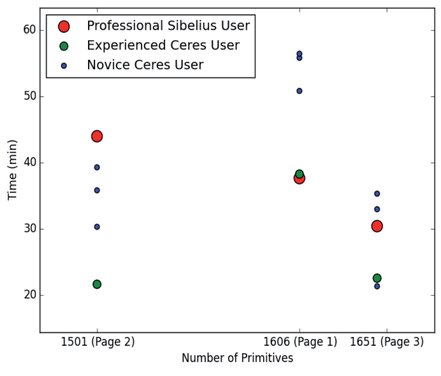
Example of Model Constraint

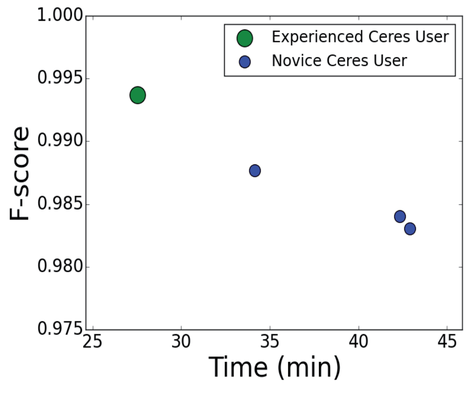


Incorrect graph

Correct graph
Beam using grand staff

User Study





Application: Renotation

Renotation on the first page of the Breitkopf and Hartel 1862-90 edition of Beethoven's Piano Sonata No. 23, (the "Appassionata").