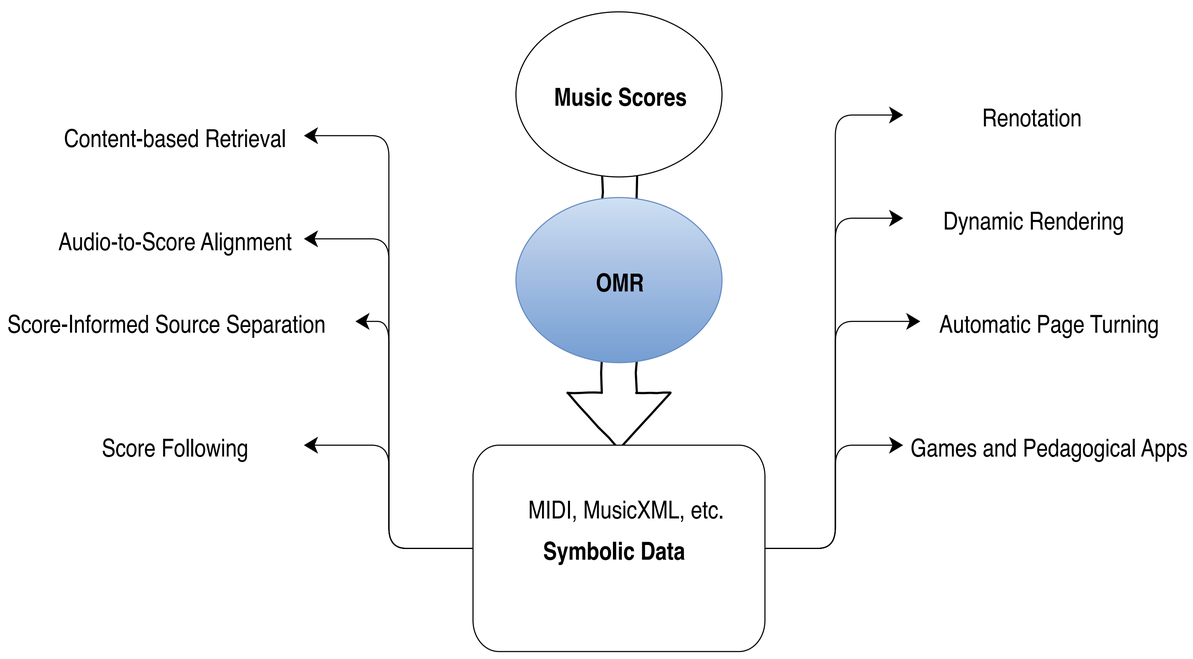
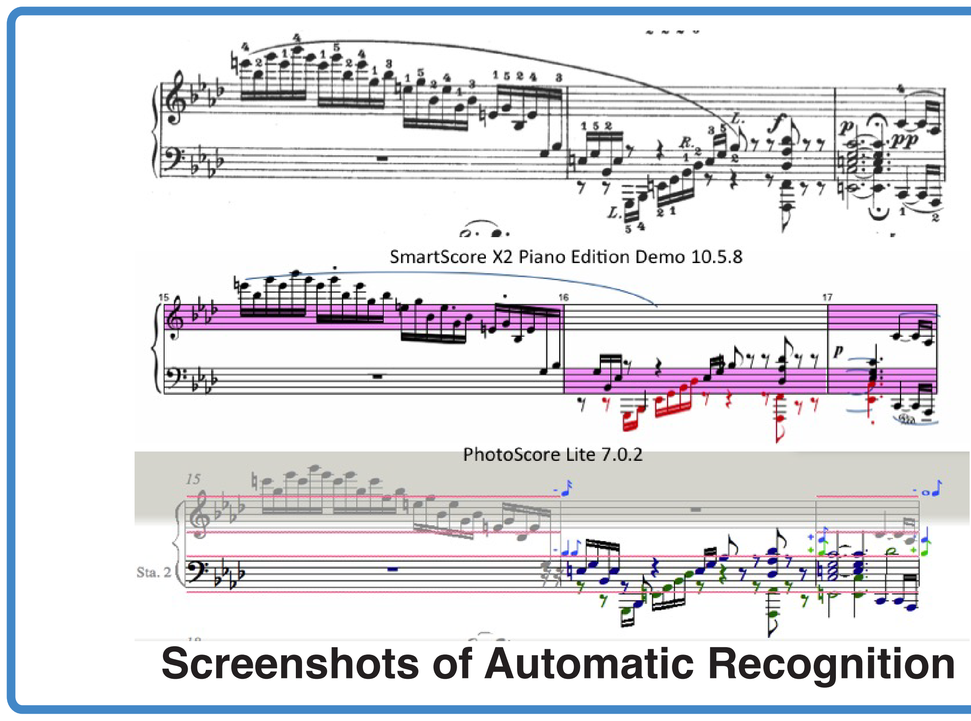
17th International Society for Music Information Retrieval Conference - August 7-11, 2016 NYC

Human-Interactive Optical Music Recognition

Liang Chen, Erik Stolterman, Christopher Raphael, Indiana University Bloomington





Screenshots of Automatic Recognition

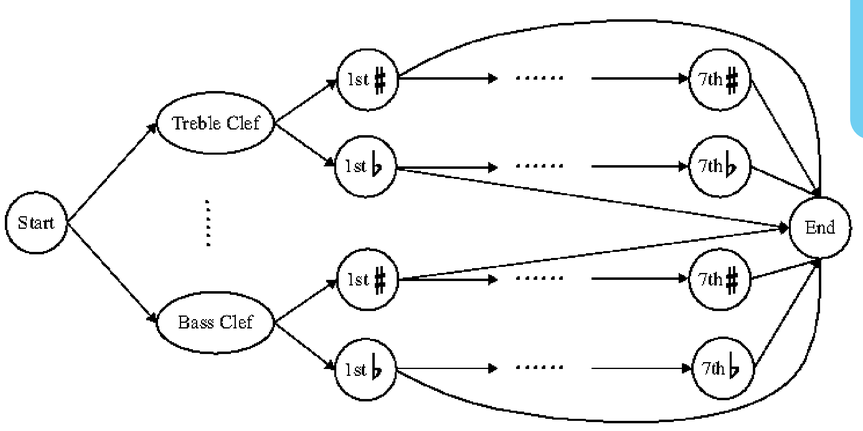
Automatic Staff/ System/ Symbol Recognition

Primitive Level: template matching using Maximum Likelihood Estimation

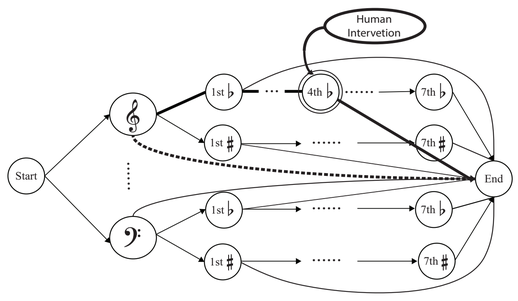
$$p(g) = \prod_{x \in B} p_B(g(x)) \prod_{x \in W} p_W(g(x)) \prod_{x \in U} p_U(g(x))$$

Symbol Level: graphical models incorporating grammatical constraints

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



Human-in-the-loop computation



Label Constraints: using human-imposed pixel labels

$$H^* = arg \max_H Q_H$$

$$Q_H = S_H + T_H$$

$$T_H = \sum_x t(x, L_H(x))$$

$$t(x, L_H(x)) = \begin{cases} C & \text{Pixel label correctly matched} \\ -C & \text{Pixel label not correctly matched} \\ 0 & \text{Otherwise} \end{cases}$$

The Problem

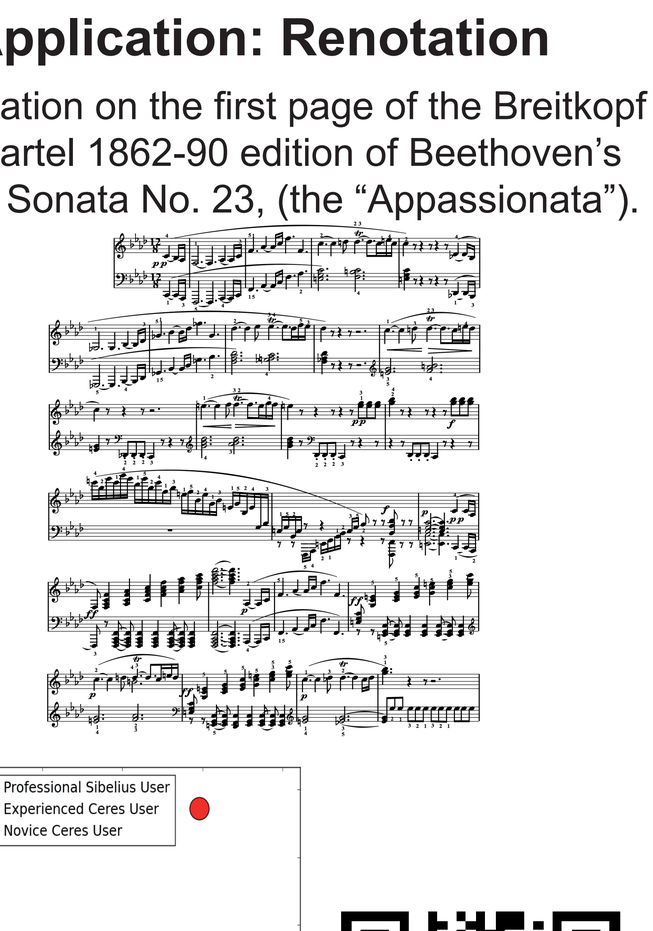
Current systems **separate** score recognition and proofreading as two independent steps. The post-proofreading process was tedious and time-consuming, sometimes taking even longer time than creating the score from scratch.

Our Approach

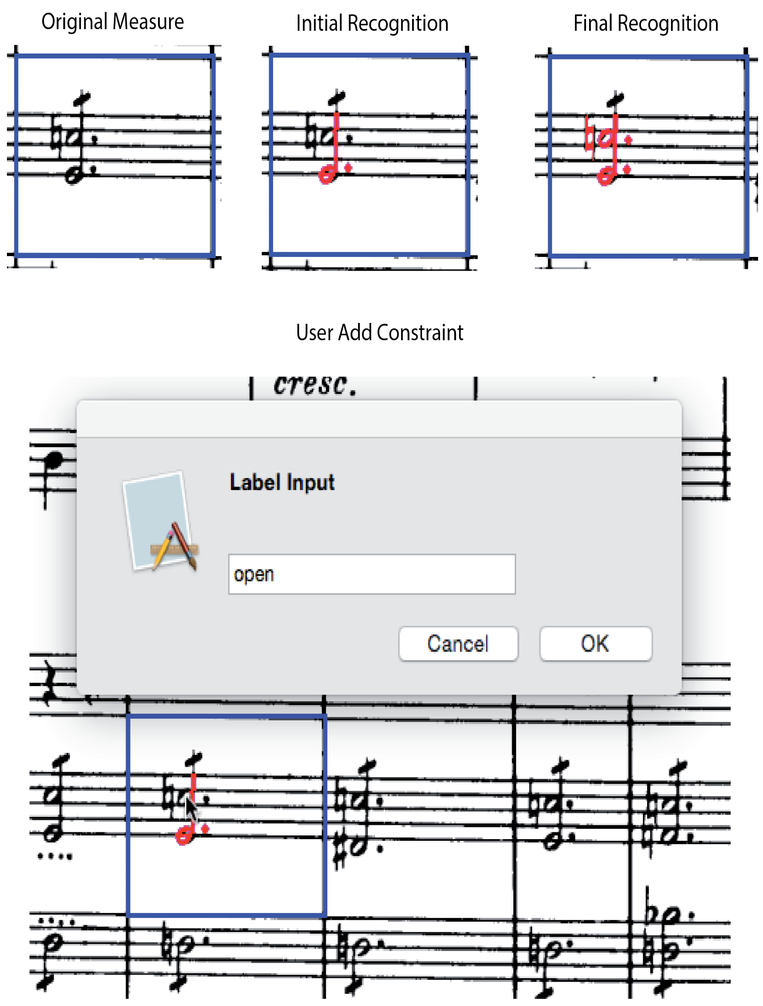
We hope to save the overall time by **combining** recognition and proofreading process via human-in-the-loop computation.

Inference:
The best configuration is parsed over the graph for different symbols using **Dynamic Programming.**

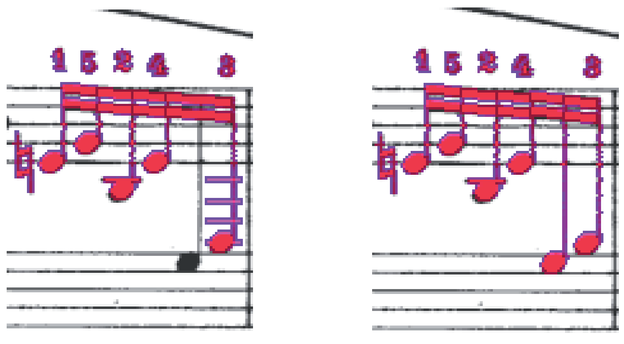
Model Constraints: changing underlying graph with toggle switches



Example of Label Constraint



Example of Model Constraint



Incorrect graph

Correct graph
Beam using grand staff

Application: Renotation

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

User Study (using the first 3 pages of "Appassionata", compared with Sibelius user)