

A CROWDSOURCED ENCODING APPROACH FOR HANDWRITTEN SHEET MUSIC

Florian Meier¹, Manuel Burghardt², Alexander Bazo², and Christian Wolff²
Information Science Group¹ and Media Informatics Group², University of Regensburg
name.surname@ur.de

ABSTRACT

We present an approach that uses *human computation* and *crowdsourcing* principles for encoding large amounts of monophonic, handwritten sheet music.

1. INTRODUCTION

The *Regensburger Volksmusik Portal* (RVP¹, Regensburg folk music portal) contains – among other resources – a collection of some 50.000 handwritten sheets of folk music from the 19th and 20th century. The songs are mostly monophonic transcriptions of traditional German folk songs, accompanied by the lyrics. Although the metadata of most sheets has already been cataloged in a database as part of previous DFG sub-project, the digitization of the lyrics and scores is still a desideratum. Our research objective is to find an efficient and reliable way to encode the handwritten scores in the *music encoding format* (MEI), to enable automatic analyses of the songs.

2. PROBLEM AND RESEARCH OBJECTIVES

Although a great number of *optical music recognition* (OMR) software already exists (e.g. *Notescan* in *Nightingale*, *Midiscan* in *Finale*, *Photoscore* in *Sibelius*, *Smartscore*, *Sharpeye* etc.), most of the tools struggle to produce a reliable transformation of music sheets into a machine-readable format [1]. Due to its heterogeneity and stylistic variability, handwritten scores are even more challenging for OMR tools, or as Müller (2007, p. 20) puts it: "So far, optical music recognition of handwritten scores is a major unresolved problem" [2]. To overcome the shortcomings of automatic OMR approaches we are planning to use human encoders to transcribe the handwritten scores into a machine-readable format. As manual encoding of over 50.000 scores is a time-consuming and laborious task, we suggest a crowdsourced approach² to get the encoding done. *Crowdsourcing*³ describes the

process of outsourcing tasks, which are traditionally performed by regular employees, by distributing them to a large group of voluntary workers on the web. Crowdsourcing platforms like *Amazon Mechanical Turk*⁴ allow the distribution and management of such tasks – the challenge however is: How to translate the task of encoding handwritten scores in a way it can be achieved by arbitrary workers from the crowd? We present a prototype for a web-based tool that allows manual encoding of the scores without having to know about notation or markup formats.

3. PROTOTYPE DESCRIPTION

At the outset we were thinking about a tool with a virtual keyboard that would allow the users to encode while actually playing along with the song. This approach was quickly dismissed as we discovered some severe problems during the conceptual design of such a tool:

- Users must know how to read music in order to be able to play the tune on the fingerboard, which reduces the pool of potential encoders from the crowd
- Using the keyboard or the mouse as a method for piano interaction is very error-prone
- The synchronization between user input/interactions and the actual position in the score is difficult to realize, especially via a web interface
- The users do not get immediate feedback about their input, it is only possible to generate audio or image feedback after the whole score has been played (difficulties for error correction)

All these drawbacks led us to think about another type of tool, which for now we have named *PlayIt* (cf. Figure 1). In the top area, the user is shown the original handwritten score. In the area below the user sees an empty line of staff. The encoders are asked to *copy* the original score by clicking on the respective position on the empty sheet of music. The user can choose from different music notation symbols, which are presented as a multi-level bubble menu. Our tool will recognize the type of symbol and its

¹ RVP Online: <<http://www.bibliothek.uni-regensburg.de/rvp/>>, all web links in this paper were last accessed July 1, 2013

² There are few projects that use crowdsourcing for the transcription of music scores, for example the Bodleian Libraries "What's the score" project at Oxford University, which is aiming to encode metadata about piano songs, <<http://www.whats-the-score.org/>>

³ The term was coined by Jeff Howe in his infamous WIRED article "The Rise of Crowdsourcing", <<http://www.wired.com/wired/archive/14.06/crowds.html>>

⁴ Amazon Mechanical Turk: <<https://www.mturk.com/mturk/>>

position in the score, and will be able to generate audio-output that can be played in the browser. It also generates a MEI-encoded song file that is saved in a database. We are planning to implement the prototype by only applying recent web standards: HTML5 and JavaScript are used to realize a web-based user interface with the necessary interactive elements. There is a huge number of existing JavaScript APIs and libraries, such as JQuery or Midi.js,

which facilitate the implementation of the prototype, and help to create an innovative user experience. In order to assure the quality of the produced transcriptions, we are planning to have every score encoded multiple times by different users. This allows to automatically identify differences in the encoding and to have these differences corrected by an expert.

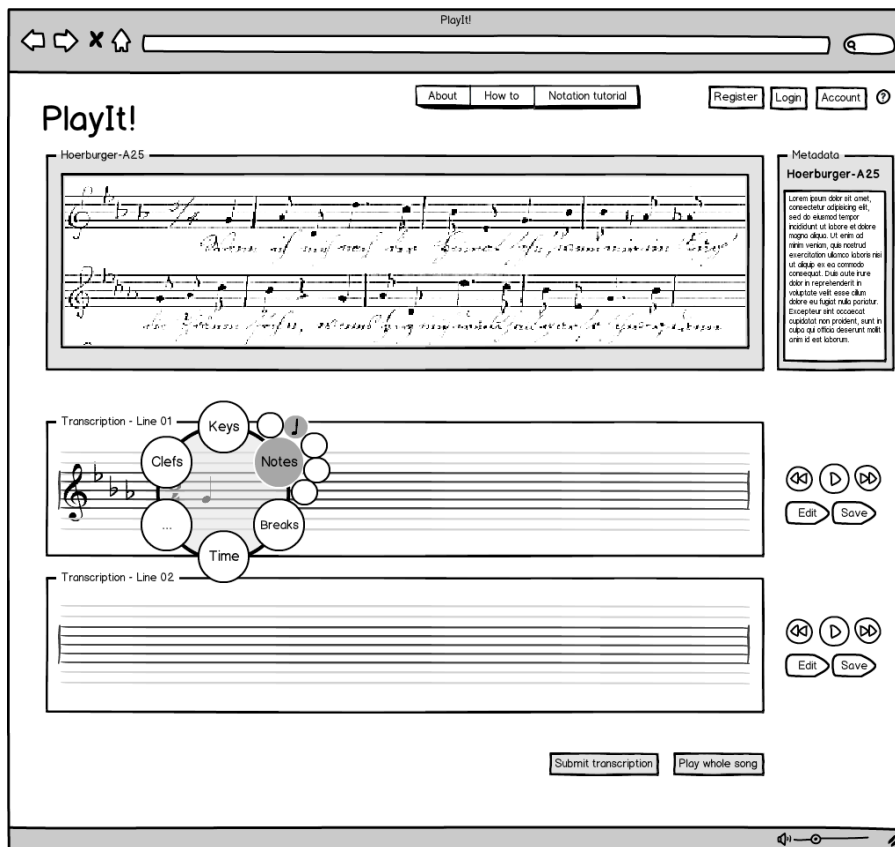


Figure 1. Mockup for the *PlayIt* encoding interface.

4. OUTLOOK

Our next steps will be to implement the prototype and conduct some tests with regard to the usability of the tool, which is considered a key feature to ensure that users are motivated to encode the songs. At this point, motivating the users is one of the main challenges: With several thousand scores that have to be encoded multiple times, we will not be able to pay high wages for the crowdworkers. We are considering *gamification* elements as possible motivation boosters [3][4]. These might include virtual rewards, like e.g. folk dance videos or additional background information on a song, and achievements for the total number of encoded songs or the overall quality of all encoded songs.

5. REFERENCES

- [1] Rebelo, A., Capela, G. & Cardoso, J. S.: “Optical recognition of music symbols: A comparative study”, *International Journal on Document Analysis and Recognition - IJDAR* , No. 13 (1), pp. 19-31, 2010.
- [2] Müller, M.: *Information Retrieval for Music and Motion*. Springer: Berlin, 2007.
- [3] von Ahn, L. & Dabbish, L.: “Labeling Images with a Computer Game”, *Proceedings of the 2004 Conference on Human Factors in Computing Systems, CHI 2004, Vienna, Austria*, pp. 319-326, 2004.
- [4] Ridge, M.: “Playing with Difficult Objects – Game Designs to Improve Museum Collections”. *Museums and the Web 2011: Proceedings*. Toronto, 2011.