Harmonizing in Real-Time with Neural Networks

Karin Höthker
hoethker@ira.uka.de

Dominik Hörnel
dominik@ira.uka.de

Institut für Logik, Komplexität und Deduktionssysteme
Universität Karlsruhe (TH)
Am Fasanengarten 5
D-76128 Karlsruhe, Germany

ABSTRACT
The task of real-time harmonization is characterized by the fact that only incomplete knowledge about the input melody is available. In particular, the global structure (phrase segmentation, overall harmonic relationships) and the longer-term melodic and tonal development cannot be taken into account. In this paper we introduce the method of dynamic tonal center forecasting to improve the real-time performance. The main idea of our approach consists in predicting changes of the underlying key, i.e. tonal centers, and in learning harmonic progressions relative to these tonal centers. In this way, both the short-term harmonic prediction performance and the longer-term tonal development can be improved. Our demonstration system HARMONET offers the possibility to interactively evaluate the proposed real-time harmonization system on arbitrary melodies entered on a MIDI keyboard. The results can be compared to a posteriori harmonizations of the same melodies based on a fundamental key which are carried out by the neural network-based system HARMONET described in [4]. Neural networks for both systems have been trained with four-part chorales of J. S. Bach and Max Reger.

1. INTRODUCTION
Many different computational paradigms have been applied to solving the task of inventing four-part harmonizations of a melody. For example in [3], an elaborated rule-based system is presented to describe the Bach chorale style, but is based on the Lerdahl-Jackendoff approach, [1] harmonizes melodies by completing them with a Boltzmann machine. The system HARMONET[4] uses feed-forward neural networks to model harmonic knowledge. In [9], four-part chorales are evolved by a genetic algorithm. The mentioned systems harmonize melodies a posteriori, such that they can take advantage of the overall melodic structure.

Real-time accompaniment systems must predict the harmonic development of a song when playing unknown melodies. Several systems have been developed for the learning of Jazz style chord sequences. In [12], a learning algorithm is used to perform real-time chord prediction in Jazz song contexts. The basic idea is to learn a probability estimation of a chord occurrence given its predecessors. In [2], a predictor combines prior knowledge (acquired by means of offline learning) with on-the-fly knowledge (acquired by means of online learning). The hybrid system is composed of a neural network which encompasses prior knowledge about typical chord sequences, and a rule-based sequence tracker which looks for recurrent chord sequences in real time.

Both systems take advantage of the fact that Jazz songs are organized in terms of well-defined chord blocks, called sections, which usually last 4 or 8 measures. These sections are repeated along the song, sometimes with slight differences. The situation is more complicated if the melody cannot be easily structured, e.g. for chorale melodies. For this reason, only a few researchers have dealt with real-time harmonization of classical harmonization styles. One approach addresses the task of inventing Bach-like chord progressions in real time learned by rule-based neural networks [10]. However, all melodies are assumed to be in a fixed C major key.

2. LEARNING HARMONIC PROGRESSIONS
The harmonization task consists in finding a suitable sequence \(H(0), H(1), \ldots\) of harmonies or chords accompanying a given melody. Each harmony \(H(t)\) is usually determined based on a local window which describes its harmonic and melodic context.

In HARMONET, neural networks are trained to predict the harmony at time \(t\), given information about previous harmonies and the melody from time \(t - 1\) to time \(t + 1\). The training patterns for the harmonic prediction (HP) network are obtained by extracting information from a local context at any time \(t\) as follows (see Figure 1): The target to be learned is the harmony \(H(t)\) which accompanies the melody quarter \(s(t)\). The input vector consists of the left harmonic \(H(t - 3), H(t - 2), H(t - 1)\) and the surrounding melodic context \(\{s(t - 1), s(t), s(t + 1)\}\). As further input features, we use the position of \(t\) relative to the begin-
First a harmonic analysis of the training examples is performed. From this analysis the frequencies of occurrence (prior probabilities \( P \)) of the analyzed harmonic functions are computed. These values are used to re-analyze a music example relative to all 24 major and minor keys. At time \( t \) a key is chosen as a tonal center \( TC(t) \) if it maximizes the sum of prior probabilities summed up over a fixed context of \( k \) harmonies:

\[
TC(t) = \arg \max_{k} \sum_{i=0}^{k} P_{k}(H(t+i) \mid H_{key})
\]

where \( H_{key} \) represents harmony \( H \) as a harmonic function relative to key. If one or more context harmonies cannot be represented relative to a tonal center, then this tonal center is not considered. We got the best results for \( k = 1 \) in our experiments. Larger context sizes lead to stranger sections for tonal centers, but local changes of tonal center are less reliably recognized.

In a second step isolated tonal centers \( TC(t) \neq TC(t-1) \) and \( TC(t) \neq TC(t+1) \) are eliminated and the extensions of the tonal center sections are determined. Figure 2 shows two analyses of the first phrase of a Bach chorale, one relative to the fundamental key and the other considering tonal centers. Notice that the tonal center analysis leads to a simplified musically coherent harmonic description because the second half of the phrase \( \text{S.SS-T.S or IV-bVII-I} \) in A major is identified as a simple cadence \( \text{T.S-D-T or IIV-V-I} \) in the new tonal center D major.

![Figure 2: Harmonic analysis of the beginning phrase of the Bach chorale "Ach bleib bei uns, Herr Jesu Christ" (a) relative to fundamental key (top) [b] considering tonal centers (bottom)](image)

### 3. TONAL CENTER ANALYSIS

The analysis and recognition of tonal centers within a musical piece presents a challenging task which usually does not have an unambiguous solution. Rule-based approaches to the automatic computation of tonal centers can be found in [8] and [11]. The basic idea in [11] is to search for harmonic transitions that confirm a tonal center. Thereafter potential tonal centers and corresponding harmonies are determined. Using a tonal center model, it is tested for each chord whether it may be represented relative to the tonal center. Unexplained sections of the harmonic sequence are treated with a specific modulation model.

Since the algorithm is not very flexible and difficult to implement, we have chosen a statistical approach to tonal center recognition based on a training set which gives similar results in most cases. It uses a kind of bootstrapping strategy:

![Figure 1: Input features and target value at quarter-beat \( t \). The question mark indicates the harmony to be learned.](image)
description, the classification performance of the networks can be significantly improved. To show this effect, we have chosen two sets of Bach chorales and a set of Reger chorales. For every set, we trained HP and RTHP feed-forward neural networks with and without tonal center analysis. The networks were tested on independent test sets of the same style. Table 1 shows that the percentage of correctly classified harmonies increases by 9.1 to 19.6 percent if tonal centers are considered. Note that the input features of a network are identical whether the data is analyzed with respect to tonal centers or relative to a fixed fundamental key. This implies that the increase of the classification rate within the rows of Table 1 is due entirely to the tonal center analysis, and not to the selected input features.

When comparing the HP and the RTHP networks for each setting, the classification performance of the HP network is superior to the performance of the corresponding RTHP network, due to the additional features like the further melodic development in the HP input patterns.

The classification rate for the Reger style is generally lower, because the networks have to learn harmonizations of higher complexity from fewer training examples than for the Bach style.

An essential component of the system is the computation of the tonal centers. We use a tonal center change (TCC) network which decides whether a tonal center change takes place or not, and a tonal center prediction (TCP) network which computes a new tonal center if the TCC network has decided to change the tonal center. The TCC network takes its decisions according to the melody notes $s(t)$, $s(t-1)$ and $s(t-2)$, the previous harmony $H(t-1)$ and the previous tonal center $TC(t-1)$. The TCP network selects a new tonal center $TC(t)$ according to the current melody note $s(t)$, the previous tonal center $TC(t-1)$ and the previous harmony $H(t-1)$. Figure 3 shows a melody harmonized with the proposed approach.

5. THE HARMOTHEATER SYSTEM

Figure 4 gives an overview of the HARMOTHEATER system. The Real-Time Module offers the possibility to interactively evaluate the proposed real-time harmonization model on arbitrary melodies entered on a MIDI keyboard. The results can be compared to a posteriori harmonizations without tonal centers of the same melodies carried out by the neural network-based system described in [4] (Offline Module). Neural networks for both modules have been trained with four-part chorales of J. S. Bach and Max Reger. Additionally, a “random style” with untrained neural networks and original examples give an idea of the space of potential harmonizations.

The user interface is kept very simple so that it can be used by people with little musical background knowledge. It is built like a stage – hence the name HARMOTHEATER – on which the “virtual composer” Reger, Bach and Randon are acting (see Figure 5). The Real-Time Module performs the harmonization in the following way: When a note is played by the user, the TCC network decides whether a tonal center change takes place or not. If necessary, the TCP network predicts the next tonal center (Tonal Center Prediction). The RTHP network computes the harmony relative to the current tonal center. Then, the pitches for the alto, tenor and bass voice are chosen according to the harmonic function and additional rules from musicological text-books (Real-Time Harmonization). The resulting harmony is immediately sent to the output port.

![Figure 3: Harmonization by HARMOTHEATER modulating from C major through c minor to G Major (Bach style)](image)

![Figure 4: Overview of the HARMOTHEATER system](image)
6. CONCLUSION

We have presented a system that harmonizes arbitrary melodies in real time. The system allows the comparison of the results to a posteriori harmonizations of the same melodies carried out by a neural network-based system. A significant increase in performance could be achieved by dynamically forecasting tonal centers based on the development of the melody.

The proposed method is suitable for melodies improvised on a keyboard in real time, but can also be used to harmonize melodies a posteriori. For non-improvised melodies, however, the results are sometimes poorer with tonal center prediction than without, if the current tonal center is not correctly predicted, thus interrupting the flow of the overall harmonic development. One could therefore try to concurrently evolve harmonies and tonal centers within a genetic framework assessed by the above-mentioned networks as proposed in [5] for the completion of melodies. Initial experiments suggest that harmonizations are created which exhibit more elaborated tonal structure, maintaining the quality of local harmonic predictions.

References