Song Prompter: 歌詞とコードをスクロール表示するインタラクティブ演奏支援システム

Matthias Mauch[†]　藤原弘将[†]　後藤真孝[†]

本稿では，ベース音と MIDI ドラム音を伴奏として再生しながら，それに同期して，楽曲の歌詞，コード，ビート位置，小節区切り位置を横スクロール表示することで，ユーザーが曲に合わせて歌唱，演奏することを支援するインタフェース Song Prompter について述べる．歌詞・コードの時間情報と伴奏音は，元楽曲の音響信号から自動的に計算される．さらに，ユーザが楽曲の構造を把握し，演奏したい区間に簡単にアクセスできるように，楽曲の区間構成と各区間の名前を表示する．Song Prompter により，ユーザは楽曲の歌詞やコードを暗記したり，楽譜のページをめくったりすることなく，元楽曲のタイムラインと同期して歌や演奏を楽しむことができる．また，コードと伴奏音は異なる調に移調して表示，再生することができ，楽曲の再生スピードも変更することができる．

Song Prompter: an Interactive Performance Assistant with Scrolling Lyrics and Chord Display

Matthias Mauch,[†] Hiromasa Fujihara[†] and Masataka Goto[†]

We present *Song Prompter*, a software system that acts as a performance assistant by showing horizontally scrolling lyrics and chords in a graphical user interface, together with an audio accompaniment consisting of bass and MIDI drums. Both alignment and accompaniment are automatically calculated from the original song recording. A song outline displays the song structure, including names and positions of sections for easy overview and navigation. *Song Prompter* enables users to sing and play live along the timeline of an original song, without having to memorize lyrics and chords or turning pages. Chord labels, and bass and audio playback can be transposed to a different key, and the playback speed can be changed.

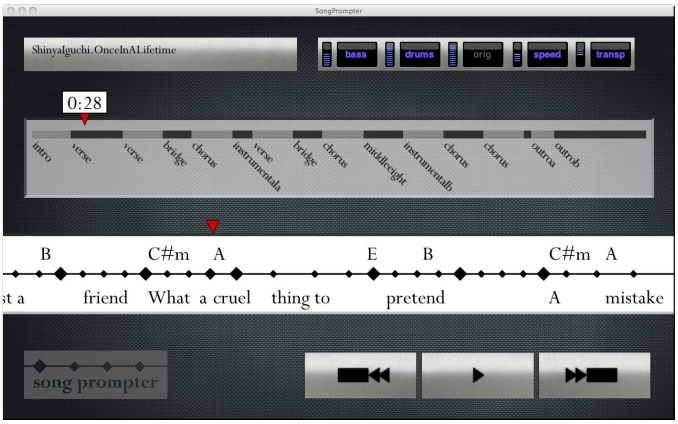
1. Introduction

Song Prompter is a software system that acts as a performance guide for musicians who play songs, from novice to professional. Once the user has opened a song, two aspects of the song are displayed in *Song Prompter*: a global song view (song outline), including song sections (chorus, verse, etc.), and a zoomed-in view of the lyrics, chords, beats, and bars at the current position in the song (song stream pane). The user can flexibly navigate to the desired song position and click to start performing: in the song stream pane the lyrics, chords, and beats are fluidly scrolling from right to left, synchronised with an accompaniment consisting of bass and drums, or the original song recording. These visual and auditory cues provide all the information necessary for a musician to play the guitar and sing along without having to memorize lyrics and chords or turning pages – even to songs that he or she has not practiced before.

Song Prompter is not a score following system (for a review of score following systems see[1)]), and does not currently use any microphone input. Instead, the user plays along to the fixed time line of the original song, like music games such as *SingStar*, *Guitar Hero*, and *Rock Band*. What makes *Song Prompter* stand out from these musical games is that it provides the navigation interface and the musical information (chords, beats, lyrics) which is needed by musicians to help them practice and perform, combining elements of games to support true musicianship. Furthermore, *Song Prompter* does not rely on commercially preproduced song information: it uses intelligent music processing methods to automatically calculate the lyrics-to-audio alignment, extract and re-synthesize the bass and drums from available commercial music recordings. Users need to provide only an audio recording of the song and the corresponding text and lyrics file in a format often used in song books and in the Internet.

† 産業技術総合研究所
National Institute of Advanced Industrial Science and Technology (AIST)

1. status block
2. song outline
3. song stream pane
4. transport buttons



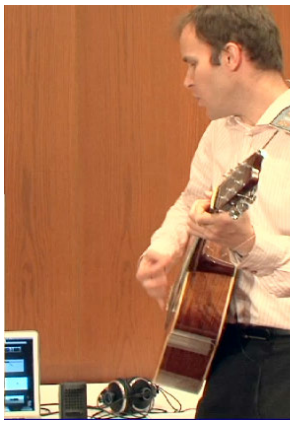


Fig. 1 *Song Prompter* interface screenshot and usage scenario of a single guitarist singing and playing to an accompaniment of bass and drums.

Many usage scenarios are conceivable. For example, in a cover band rehearsal situation, a band member can propose a new song by his favourite artist, and the band can immediately start playing the song based on the lyrics, beats, and chord progression displayed in *Song Prompter*. When the band performs live, *Song Prompter* can literally act as an automatic prompter. In a more informal setting, it can be used when a party of friends want to sing together, and someone has brought along a guitar: *Song Prompter* is more convenient than song books because no pages need to be turned, and it prevents people from loosing track of the current song position (as often happens with static formats). Furthermore, since the text is scrolling, only a small portion of the lyrics need to be seen, and they can be displayed in a larger font than is possible in a static book format.

In Section 2 we will describe the *Song Prompter* interface in more details, and Section 3 explains the underlying music processing technology. We conclude with some initial experiences with the system and a look to the future in Section 4.

A demonstration of *Song Prompter* can be viewed online☆.

2. Interface

As can be seen in the screenshot (Figure 1) the user interface is divided into four parts, from top to bottom: the *status block* with the song title and detailed performance controls, the *song outline* providing global song information and navigation, the *song stream pane* containing the main display function of *Song Prompter*, and finally a set of easy access *transport control* buttons.

2.1 Status Block

The status block (Figure 1) consists of two parts. On the left, the current song title is displayed, and on the right a panel allows to view and modify the playback parameters: bass on/off and bass volume, drums on/off and drum volume, playback of original recording on/off and volume, playback speed control, and key transpose control.

2.2 Song Outline

The song outline enables a user to see the structure of the whole song including section titles, but also provides an efficient way of navigating this structure. The user can click/tap on a section to access it directly and playback position will be set to the beginning of the selected section. The time display has two functions: it displays the current elapsed song time both in digital form and in analog form by its position on the song outline, but also functions as a navigation device: the user can simply drag the time display to the desired position.

2.3 Song Stream Pane

The song stream pane is the central display of *Song Prompter*, offering scrolling lyrics, chords, and beats synchronised to the timeline of the original recording of the song (see Figure 2). The display is designed so that lyrics and chords can easily be viewed at the same time, which is important for musicians both playing and singing. Diamond-shaped markers provide cues to where the beats are, larger markers mark the beginning of a bar (measure). The current song position is highlighted by pulsating beat marks, i.e. the beat marks enlarge at the current playback position. The playback position is additionally marked by a red triangle above the song stream pane. By dragging this triangle the display position can be adjusted to the left or right to suit the taste of the musician. Precise navigation in the song can be done by dragging the song stream pane directly.

To accommodate fast successions of lyrics without displaying overlapped words, the time line is stretched accordingly, similar to the variable lengths of bars (measures) in traditional sheet mu-

☆ http://www.youtube.com/user/SongPrompter

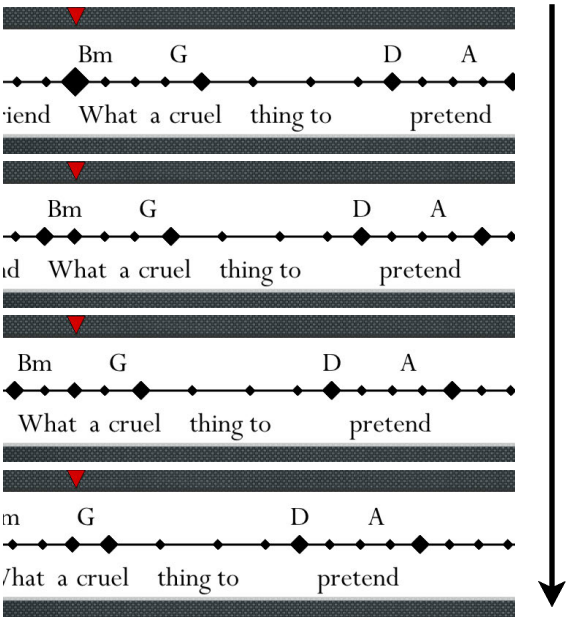


Fig. 2 Horizontally scrolling lyrics, chords, and beats. The figures provide four snapshots of the song stream pane moving from right to left. The current song position is marked by a red triangle and an enlarged beat marker.

sic. When the user chooses to transpose the piece to a different key, the chord labels are automatically updated (Figure 3) to match the transposed accompaniment, i.e. bass and original audio.

2.4 Transport Control

The large transport control buttons are located at the bottom of the display in order not to obstruct the view of the song stream pane when handling *Song Prompter* on a touch screen display. The buttons allow instant navigation to the previous or following song part as well as the traditional start/pause function.

3. Underlying Technology

Song Prompter is an interface to state-of-the-art music processing technology: lyrics-to-audio alignment and the automatic detection and re-synthesis of music from the original recording.

3.1 Lyrics-to-Audio Alignment

We use the automatic lyrics-to-audio alignment method proposed in[7], which makes use of additional chord information. This is necessary when only a textual description of the lyrics and chords, and the audio waveform of the original recording are given (see Figure 4). Hence, no musical score input is needed. The method combines automatic parsing of the lyrics and chords, calculation of and chroma features[5] and mel frequency cepstral coefficients (MFCCs) based on the segregated vocal signal[4]. The alignment method is an extension of

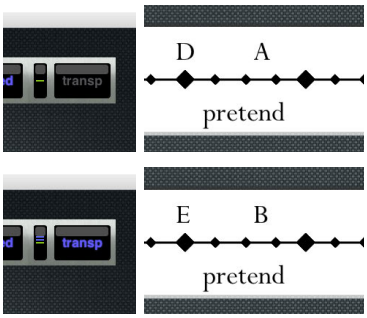


Fig. 3 Transpose to a different key: The user can choose to transpose the piece. Top: original (not transposed), the transpose status button is off, and the original chord is shown; bottom: transposed by two semitones up, the status button shows is on and indicates a transposition of two semitones, the chord display is updated accordingly. The audio playback is also transposed by the same amount.

3) and uses hidden Markov models[8], which is commonly used in speech recognition, but additionally models chords[6]. Once the alignment to the original audio has been performed, we know the position of every word and chord, but also every section of the song. This data is used in both the song outline and the song stream pane. The beat marks and larger bar marks complete the white song stream pane (see Figure). Their positions are obtained using either an established method[2] or a novel, yet unpublished method.

3.2 Automatic Accompaniment Generation

Like in the case of lyrics-to-audio alignment, the accompaniment provided by *Song Prompter* is also automatically derived directly from an original audio recording. The drum accompaniment is realised as MIDI, based on the automatically extracted beat and bar times. The fundamental frequency and the amplitude of the partials in the bass line are estimated using PreFEst[4], and are then re-synthesised during playback. The bass line can hence be transposed, sped up or slowed down. To achieve the same for the audio playback of the original recording, we use the implementation of granular synthesis provided in the Beads project☆.

4. Discussion

Song Prompter is features on a unique combination of music processing methods and provides an interface to make them available to musicians. In our experience, the quality of the alignment is usually good enough to easily follow the song, and the flow in the song stream pane makes it easy for a musician to follow the song both in terms of lyrics

☆ http://www.beadsproject.net/

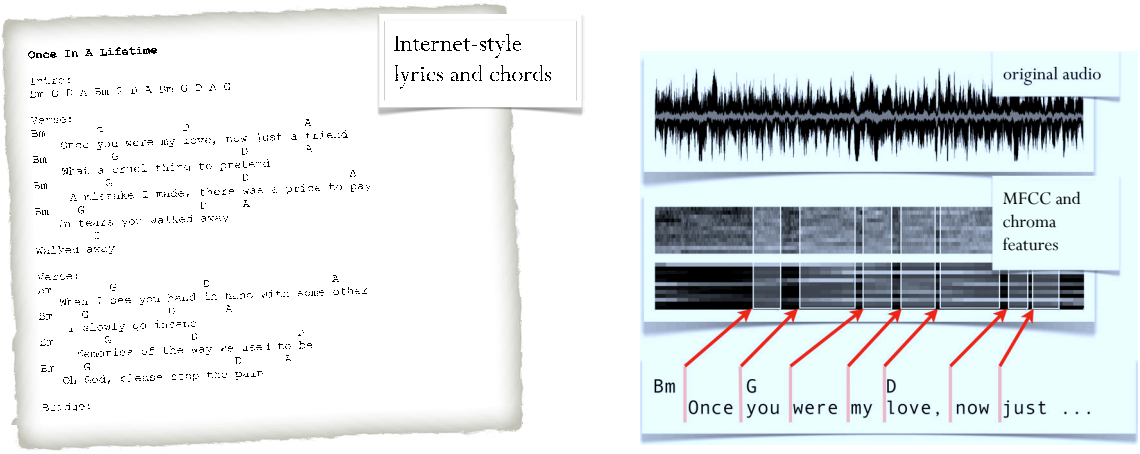


Fig. 4 The information needed to perform lyrics-to-audio alignment including chords. Left hand side: a typical lyrics and chords sheet in the format often found in song books and in the Internet. Right hand side: schematic view of the audio waveform, the audio features (MFCCs and chroma). Red arrows illustrate the alignment between lyrics, chords, and the audio features.

and chords. For popular music, an additional synchronisation between a performer and the software does not seem strictly necessary because the tempo is usually meant to be stable, and it's very easy to play along. However, in the future, we will investigate ways of robustly synchronising playback to the rhythm or harmony provided by the performer. Another possibility of improvement is to make the display more flexible: when vocals and the accompanying instrument (e.g. a guitar) are performed by two separate musicians, the vocalist usually does not require the chord flow in the song stream pane, and in the future we will add the option of separate lyrics and chord displays.

5. Acknowledgements

This work was funded by the CrestMuse Project of the Japan Science and Technology Agency (JST).

References

1) Roger B. Dannenberg and Christopher Raphael. Music score alignment and computer accompaniment. *Communications of the ACM*, 49(8):38–43, 2006.
2) Matthew E. P. Davies, Mark D. Plumbley, and Douglas Eck. Towards a musical beat emphasis function. In *Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA 2009)*, 2009.
3) H. Fujihara, M. Goto, J. Ogata, K. Komatani, T. Ogata, and H.G. Okuno. Automatic synchronization between lyrics and music CD recordings based on Viterbi alignment of segregated vocal signals. In *8th IEEE International Symposium on Multimedia (ISM'06)*, pages 257–264, 2006.
4) Masataka Goto. A real-time music scene description system: Predominant-F0 estimation for detecting melody and bass lines in real-world audio signals. *Speech Communication (ISCA Journal)*, 43(4):311–329, 2004.
5) Matthias Mauch and Simon Dixon. Approximate note transcription for the improved identification of difficult chords. In *Proceedings of the 11th International Society for Music Information Retrieval Conference (ISMIR 2010)*, pages 135–140, 2010.
6) Matthias Mauch and Simon Dixon. Simultaneous estimation of chords and musical context from audio. *to appear in IEEE Transactions on Audio, Speech, and Language Processing*, 18(6):1280–1289, 2010.
7) Matthias Mauch, Hiromasa Fujihara, and Masataka Goto. Lyrics-to-audio alignment and phrase-level segmentation using incomplete internet-style chord annotations. In *Proceedings of the 7th Sound and Music Computing Conference (SMC 2010)*, pages 9–16, 2010.
8) Lawrence R. Rabiner. A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):257–286, 1989.