A probabilistic approach to determining bass voice leading in melodic harmonisation

Dimos Makris[a*], Maximos Kaliakatsos-Papakostas[b], and Emilios Cambouropoulos[b]

[a]Department of Informatics, Ionian University, Corfu, Greece
[b]School of Music Studies, Aristotle University of Thessaloniki, Thessaloniki, Greece
c12makr@ionio.gr
{maxk,emilios}@mus.auth.gr

Abstract. Melodic harmonisation deals with the assignment of harmony (chords) over a given melody. Probabilistic approaches to melodic harmonisation utilise statistical information derived from a training dataset to harmonise a melody. This paper proposes a probabilistic approach for the automatic generation of voice leading for the bass note on a set of given chords from different musical idioms; the chord sequences are assumed to be generated by another system. The proposed bass voice leading (BVL) probabilistic model is part of ongoing work, it is based on the hidden Markov model (HMM) and it determines the bass voice contour by observing the contour of the melodic line. The experimental results demonstrate that the proposed BVL method indeed efficiently captures (in a statistical sense) the characteristic BVL features of the examined musical idioms.

Keywords: voice leading, hidden Markov model, bass voice, conceptual blending

1 Introduction

Melodic harmonisation systems assign harmonic material to a given melody. Harmony is expressed as a sequence of chords, but the overall essence of harmony is not concerned solely with the selection of chords; an important part of harmony has to do with the relative placement of the notes that comprise successive chords, a problem known as *voice leading*. Voice leading places focus on the horizontal relation of notes between successive chords, roughly considering chord successions as a composition of several mutually dependent voices. Thereby, each note of each chord is considered to belong to a separate melodic stream called a *voice*, while the composition of all voices produces the chord sequence.

Regarding melodic harmonisation systems, there are certain sets of "rules" that need to be taken under consideration when evaluating voice leading. However, these "rules" are defined by musical systems, called *idioms*, with many

* Corresponding author.

differences. The work presented in this paper is a part of an ongoing research within the context of the COINVENT project [10], which examines the development of a computationally feasible model for conceptual blending. Therefore, the inclusion of many diverse musical idioms in this approach is required for achieving bold results that blend characteristics from different layers of harmony across idioms.

The aspect of harmony that this paper addresses is voice leading of the bass voice, which is an important element of harmony. Experimental evaluation of methodologies that utilise statistical machine learning techniques demonstrated that an efficient way to harmonise a melody is to add the bass line first[11]. To the best of our knowledge, no study exists that focuses only on generating voice leading contour of the bass line independently of the actual chord notes (i.e. the actual chord notes that belong to the bass line are determined at a later study).

2 Probabilistic bass voice leading

The proposed methodology aims to derive information from the melody voice in order to calculate the most probable movement for the bass voice, hereby referred to as the *bass voice leading* (BVL). This approach is intended to be harnessed to a larger modular probabilistic framework where the selection of chords (in GCT form [2]) is performed on an other probabilistic module [6]. Therefore, the development of the discussed BVL system is targeted towards providing indicative guidelines to the overall system about possible bass motion rather than defining specific notes for the bass voice.

The level of refinement for representing the bass and melody voice movement for the BVL system is also a matter of examination in the current paper. It is, however, a central hypothesis that both the bass and the melody voice steps are represented by abstract notions that describe pitch direction (up, down, steady, in steps or leaps etc.). Several scenarios are examined in Section 3 about the level of refinement required to have optimal results. Table 1 exhibits the utilised refinement scales in semitone differences for the bass and melody voice movement. For example, by considering a refinement level 2 for describing the melody voice, the following set of seven descriptors for contour change are considered: $\texttt{mel}_2 = \{\texttt{st\_v}, \texttt{s\_up}, \texttt{s\_down}, \texttt{sl\_up}, \texttt{sl\_down}, \texttt{bl\_up}, \texttt{bl\_down}, \}$ while an example of refinement level 0 for the bass voice has the following set of descriptors: $\texttt{bass}_0 = \{\texttt{st\_v}, \texttt{up}, \texttt{down}\}$. On the left side of the above equations, the subscript of the melody and the bass voice indicators denotes the level of refinement that is considered. Under this representation scheme, a given chord sequence in MIDI pitch numbers, such as: $[67, 63, 60, 48]$, $[67, 62, 65, 47]$, $[63, 60, 65, 48]$, $[65, 60, 60, 56]$ gives bass and melody (soprano) voice leading: $[-1, 0]$, $[+1, -4]$, $[+8, +2]$, which eventually becomes: [down, st_v], [up, bl_down], [up, sl_up].

The main assumption for developing the presented BVL methodology is that bass voice is not only a melody itself, but it also depends on the piece's melody. Therefore, the selection of the next bass voice note is dependent both on its previous note(s), as well as on the current interval between the current and

refinement description	short name	semitone difference range	refinement level
steady voice	`st_v`	0	0, 1, 2
up	`up`	above 0	0
down	`down`	below 0	0
step up	`s_up`	between 1 and 2	1, 2
step down	`s_down`	between −2 and −1	1, 2
leap up	`l_up`	above 2	1
leap down	`l_down`	below −2	1
small leap up	`sl_up`	between 3 and 5	2
small leap down	`sl_down`	between −3 and −5	2
big leap up	`bl_up`	above 5	2
big leap down	`bl_down`	below −5	2

Table 1. The pitch direction refinement scales considered for the development of the proposed BVL system, according to the considered level of refinement.

the previous notes of the melody. This assumption, based on the fact that a probabilistic framework is required for the harmonisation system, motivates the utilisation of the *hidden Markov model* (HMM) methodology. According to the HMM methodology, a sequence of observed elements is given and a sequence of (hidden) states is produced as output. The training process of an HMM incorporates the extraction of statistics about the probabilities that a certain state (bass direction descriptor) follows an other state, given the current observation element (melody direction descriptor). These statistics are extracted from a training dataset, while the state sequence that is generated by an HMM system, is produced according to the maximum probability described by the training data statistics – considering a given sequence of observation elements.

3 Experimental results

Aim of the experimental process is to evaluate whether the presented approach composes bass voice leading sequences that capture the intended statistical features regarding BVL from different music idioms. Additionally, it is examined whether there is an optimal level of detail for grouping successive bass note differences in semitones (according to Table 1), regarding BVL generation. To this end, a collection of five datasets has been utilised for training and testing the capabilities of the proposed BVL-HMM, namely: 1) a set of Bach Chorales, 2) several chorales from the 19th and 20th centuries, 3) polyphonic songs from Epirus, 4) a set of medieval pieces and 5) a set of modal chorales. These pieces are included in a dataset composed by music pieces (over 400) from many diverse music idioms (seven idioms with sub-categories). The Bach Chorales have been extensively employed in automatic probabilistic melodic harmonisation [1, 3, 9, 8], while the polyphonic songs of Epirus [7, 5] constitute a dataset that has hardly been studied. Several refinement level scenarios have been examined for the melody and the bass voices that are demonstrated in Table 2.

Each idiom's dataset is divided in two subsets, a *training* and a *testing* subset, with a proportion of 90% to 10% of the entire idiom's dataset. The training subset is utilised to train a BVL-HMM according to the selected refinement

scenario	bass refinement	melody refinement	states × observations
1	1	1	5 × 5
2	1	2	5 × 7
3	0	2	3 × 7
4	0	1	3 × 5
5	0	0	3 × 3

Table 2. The examined scenarios concerning bass and melody voice refinement levels. According to Table 1, each refinement level is described by a number of states (bass voice steps) and observations (melody voice steps).

scenario. A model trained with the sequences (bass movement transitions and melody movement observations) of a specific idiom, X, will hereby be symbolised as M_X while the testing pieces denoted as D_X. The evaluation of whether a model M_X predicts a subset D_X better than a subset D_Y is achieved through the cross-entropy measure. The measure of cross-entropy is utilised to provide an entropy value for a sequence from a dataset, $\{S_i, i \in \{1, 2, \ldots, n\}\} \in D_X$, according to the context of each sequence element, S_i, denoted as C_i, as evaluated by a model M_Y. The value of cross-entropy under this formalisation is given by $-\frac{1}{n}\sum_1^n \log\ P_{M_Y}(S_i, C_{i,M_Y})$, where $P_{M_Y}(S_i, C_{i,M_Y})$ is the probability value assigned for the respective sequence element and its context from the discussed model.

The magnitude of the cross entropy value for a sequence S taken from a testing set D_X does not reveal much about how well a model M_Y predicts this sequence – or how good is this model for generating sequences that are similar to S. However, by comparing the cross-entropy values of a sequence X as predicted by two models, D_X and D_Y, we can assume which model predicts S better: the model that produces the *smaller* cross entropy value [4]. Smaller cross entropy values indicate that the elements of the sequence S "move on a path" with greater probability values. The effectiveness of the proposed model is indicated by the fact that most of the minimum values per row are on the main diagonal of the matrices, i.e. where model M_X predicts D_X better than any other D_Y.

Results indicated that scenarios 3 and 4 constitute more accurate refinement combinations for the melody and bass voices. Table 3 exhibits the cross-entropy values produced by the BVL-HMM under the refinement scenario 3, which is among the best refinement scenarios, where the systems are trained on each available training datasets for each test set's sequences. The presented values are averages across 100 repetitions of the experimental process, with different random divisions in training and testing subsets (preserving a ratio of 90%-10% respectively for all repetitions).

An example application of the proposed BVL system is exhibited in Figure 1, where GCT chords were produced by the cHMM [6] system. The chordal content of the harmonisation is functionally correct and compatible with Bach's style. The proposed bass line exhibits only two stylistic inconsistencies, namely the two $^6_4$ chords in the first bar. The overall voice leading is correct, except for the parallel octaves (first two chords) - note that the inner voices have been added by a very simple nearest position technique and that no other voice leading rules are

	M_{Bach}	$M_{\text{19th-20th}}$	M_{Epirus}	M_{Medieval}	M_{Modal}
D_{Bach}	**2.4779**	2.5881	31.0763	16.0368	5.3056
$D_{\text{19th-20th}}$	13.8988	**5.0687**	70.1652	31.6096	15.9747
D_{Epirus}	3.3127	3.1592	**2.8067**	2.9990	3.0378
D_{Medieval}	3.0988	3.0619	3.1845	**2.7684**	2.8539
D_{Modal}	3.0037	2.9028	3.3761	2.9611	**2.7629**

Table 3. Mean values of cross-entropies for all pairs of datasets, according to the refinement scenario 3.

accounted for. The presented musical example, among other examples, strongly suggests that further (statistical) information about the voicing layout of chords is required for generating harmonic results that capture an idioms style.

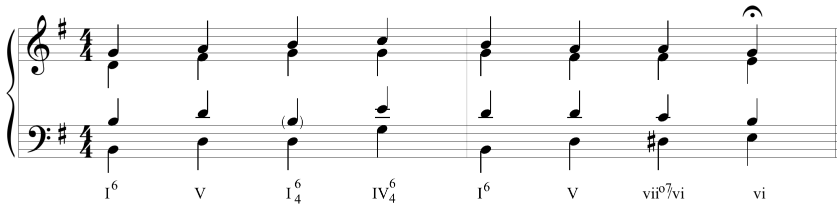

Fig. 1. Bach chorale melodic phrase automatically harmonised, with BVL generated by the proposed system (roman numeral harmonic analysis done manually).

4 Conclusions

This paper presented a methodology for determining the bass voice leading (BVL) given a melody voice. Voice leading concerns the horizontal relations between notes of the harmonising chords. The proposed bass voice leading (BVL) probabilistic model utilises a hidden Markov model (HMM) to determine the most probable movement for the bass voice (hidden states), by observing the soprano movement (set of observations). Many variations regarding the representation of bass and soprano voice movement have been examined, discussing different levels of representation refinement expressed as different combinations for the number of visible and hidden states. Five diverse music idioms were trained creating the relevant BVLs, while parts of these idioms were used for testing every system separately. The results indicated low values in term of cross entropy for each trained BVL system with the corresponding testing dataset and high values for examples from different music idioms. Thereby, it is assumed that the proposed methodology is efficient, since some characteristics of voice leading are captured for each idiom.

For future work, a thorougher musicological examination of the pieces included in the dataset will be pursued, since great difference were observed for the voice leading of pieces included in some idioms (e.g. $M_{\text{19th-20th}}$ set). Additionally, our aim is the development of the overall harmonisation probabilistic system that employs additional voicing layout statistical information, while chord selection (based on a separate HMM module) will be also biased by the adequacy of each chord to fulfil the voice leading scenario provided by the voice leading probabilistic module – part of which is presented in this work.

Acknowledgements: This work is founded by the COINVENT project. The project COINVENT acknowledges the financial support of the Future and Emerging Technologies (FET) programme within the Seventh Framework Programme for Research of the European Commission, under FET-Open grant number: 611553.

References

1. Allan, M., Williams, C.K.I.: Harmonising chorales by probabilistic inference. In: Advances in Neural Information Processing Systems 17. pp. 25–32. MIT Press (2004)
2. Cambouropoulos, E., Kaliakatsos-Papakostas, M., Tsougras, C.: An idiom-independent representation of chords for computational music analysis and generation. In: Proceeding of the joint 11th Sound and Music Computing Conference (SMC) and 40th International Computer Music Conference (ICMC). ICMC–SMC 2014 (2014)
3. Jordan, M.I., Ghahramani, Z., Saul, L.K.: Hidden markov decision trees. In: Mozer, M., Jordan, M.I., Petsche, T. (eds.) NIPS. pp. 501–507. MIT Press (1996)
4. Jurafsky, D., Martin, J.H.: Speech and language processing. Prentice Hall, New Jersey, USA (2000)
5. Kaliakatsos-Papakostas, M., Katsiavalos, A., Tsougras, C., Cambouropoulos, E.: Harmony in the polyphonic songs of epirus: Representation, statistical analysis and generation. In: 4th International Workshop on Folk Music Analysis (FMA) 2014 (June 2011)
6. Kaliakatsos-Papakostas, M., Cambouropoulos, E.: Probabilistic harmonisation with fixed intermediate chord constraints. In: Proceeding of the joint 11th Sound and Music Computing Conference (SMC) and 40th International Computer Music Conference (ICMC). ICMC–SMC 2014 (2014)
7. Liolis, K.: To Epirótiko Polyphonikó Tragoúdi (Epirus Polyphonic Song). Ioannina (2006)
8. Manzara, L.C., Witten, I.H., James, M.: On the entropy of music: An experiment with bach chorale melodies. Leonardo Music Journal 2(1), 81–88 (Jan 1992)
9. Paiement, J.F., Eck, D., Bengio, S.: Probabilistic melodic harmonization. In: Proceedings of the 19th International Conference on Advances in Artificial Intelligence: Canadian Society for Computational Studies of Intelligence. pp. 218–229. AI'06, Springer-Verlag, Berlin, Heidelberg (2006)
10. Schorlemmer, M., Smaill, A., Kühnberger, K.U., Kutz, O., Colton, S., Cambouropoulos, E., Pease, A.: Coinvent: Towards a computational concept invention theory. In: 5th International Conference on Computational Creativity (ICCC) 2014 (June 2014)
11. Whorley, R.P., Wiggins, G.A., Rhodes, C., Pearce, M.T.: Multiple viewpoint systems: Time complexity and the construction of domains for complex musical viewpoints in the harmonization problem. Journal of New Music Research 42(3), 237–266 (Sep 2013)