Modelling Cadence Perception Via Musical Parameter Tuning to Perceptual Data

Maximos Kaliakatsos-Papakostas[(✉)], Asterios Zacharakis,
Costas Tsougras, and Emilios Cambouropoulos

Department of Music Studies, Aristotle University of Thessaloniki,
54124 Thessaloniki, Greece
{maxk,aszachar,tsougras,emilios}@mus.auth.gr

Abstract. Conceptual blending when used as a creative tool combines the features of two input spaces, generating new blended spaces that share the common structure of the inputs, as well as different combinations of their non-common parts. In the case of music, conceptual blending has been employed creatively, among others, in generating new cadences (pairs of chords that conclude musical phrases). Given a specific set of input cadences together with their blends, this paper addresses the following question: are some musical features of cadences more salient than others in defining perceived relations between input and blended cadences? To this end, behavioural data from a pairwise dissimilarity listening test using input and blended cadences as stimuli were collected, thus allowing the construction of a 'ground-truth' human-based perceptual space of cadences. Afterwards, the salience of each cadence feature was adjusted through the Differential Evolution (DE) algorithm, providing a system-perceived space of cadences that optimally matched the ground-truth space. Results in a specific example of cadence blending indicated that some features were distinguishably more salient than others. This pilot study was a first step towards building self-aware blending systems and revealed that the salience of features in conceptual blending is an essential part for producing perceptually relevant blends.

1 Introduction

The cognitive theory of conceptual blending by [9] has been extensively used in linguistics, music composition [21], music cognition [1,2] and other domains mainly as an analytical tool for explaining the cognitive processes that humans undergo when engaged in creative acts. In computational creativity, conceptual blending has been modelled by Goguen [10] as a generative mechanism, according to which two *input spaces* are blended to generate novel *blended spaces*, using tools of *category theory*. A computational framework that extends Goguen's approach has been developed in the context of the COncept INVENtion Theory[1]

---

[1] http://www.coinvent-project.eu.

(COINVENT) project [17] based on the notion of *amalgams* [8,16]. Following this framework, systems have been developed that blend features of two input musical cadences [7,19] (pair of ending chords) or chord transitions [13], producing novel blended ones that incorporate meaningful characteristics of the input ones.

The paper at hand presents a pilot study for defining the salient features of cadences in the context of a cadence blending system, based on the data collected from perceptual experiments. The system-produced blended cadences incorporate combinations of features from manually-made input cadences, while the importance of each feature differs according to its perceptual salience. The salience of features is obtained by applying the Differential Evolution (DE) algorithm for optimally matching (in terms of pairwise dissimilarity) the system-perceived cadence relations with the perceptual space extracted by the experiment with humans. This study is a first step towards increasing self-awareness in a creative system that produces cadences through conceptual blending.

2 A Formal Description of Cadences for Generative Conceptual Blending

In this paper a cadence is considered as a special case of a transition (a chord following another), but with the second chord is fixed. Therefore, when blending two input cadences, the characteristics of the penultimate chords of the inputs are combined to produce new penultimate blended chords that are paired with the fixed final chord to constitute the blended cadence. For instance, the case of blending, e.g., the perfect with the Phrygian cadences is described by the transitions I_1: G7 → Cm and I_2: B♭m → Cm respectively, while a blend of these inputs is the tritone substitution cadence, C♯7 → Cm [7,19]. The perceptual characteristics of the penultimate chord that are considered for describing a cadence are the following:

1. *fcRoot*: Root of the first chord.
2. *fcType*: Type of the first chord as presented by the GCT.
3. *fcPCs*: Pitch classes of the penultimate chord.
4. *rDiff*: Root difference for the transition.
5. *DIC0*: Existence of fixed pitch class.
6. *DIC1*: Existence of upward semitone movement in pitch classes.
7. *DIC-1*: Existence of downward semitone movement in pitch classes.
8. *DIC*: The compete DIC vector of the chord transition.
9. *asc*: Existence of ascending semitone to the tonic.
10. *desc*: Existence of descending semitone to the tonic.
11. *semi*: Existence of semitone movement towards the tonic.

For computing the root and type in a consistent manner for all utilised chords, the General Chord Type (GCT) [5,14] has been employed, which allows the rearrangement of the notes of a harmonic simultaneity such that abstract types of chords along with their root may be derived. The GCT algorithm finds the

maximal subset that forms the base upon which the chord type is built, while the lowest note of the base is the root of the chord; the potentially remaining notes are assembling the extension of the GCT, which is the set of notes that would not be a part of the maximally consonant subset. For example, the GCT representation of the first degree (I) chord in a major scale is [0, [0 4 7]], where 0 indicates the root note in relation to the scale (0 is the scale as first degree) and [0 4 7] is the chord's type (4 indicates a major third and 7 a perfect fifth). Accordingly, a V7 chord is denoted by [7, [0 4 7], [10]], where 10 is the extension (minor seventh), which cannot be included in the base considering that the tritone and minor seventh intervals are dissonant.

The aforementioned properties 1–2 describe the first chord of the cadence and the first two properties (chord root and type) are extracted from the GCT algorithm, considering that all the examined cadences are in the key of C minor. Property 4, the difference between the chord roots is an integer between -5 and 6, indicating the pitch class difference between the roots of the first and the second chords of the cadence. Property 5 captures the existence of a common note between the two chords, while properties 6 and 7 indicate the existence of a semitone movement (upward and downward respectively) in any pitch class of the cadence transition. These properties actually indicate if there is a 0, 1 or -1 in the Directional Interval Class (DIC) [6], flagging whether there are small pitch class voice leading movements (repeating notes or semitone movements) in the cadence. Property 8 incorporates the entire DIC vector of the transition/cadence. Properties 9 to 11 are used to highlight the importance of whether there is a semitone movement (property 11) to the tonic from the first to the second chord of the cadence as well as whether this movement is ascending (property 9) or descending (property 10); these properties reflect the importance of the leading note (upwards or, even, downwards).

Table 1 illustrates a blending example, where the tritone substitution cadence is created from the perfect and the phrygian cadences. This blend incorporates properties from both input spaces with a good balance, i.e. many properties that are common in both input spaces, while new properties have also been added through completion. Specifically, this blend includes five properties of input 1, four properties of input 2, three common properties and four new properties that were not present in any input space. The properties of the blended space come from either input space, or are completed by logical deduction through axioms describing cadences, as indicated in the parentheses next to each respective property.

3   Approximating the Importance of Properties According to the Perceptual Pairwise Distances of Cadences

As discussed in the introduction, one desirable property for a creative system is the ability to self-evaluate its products [12]. In this respect, the cadence blending system should be able to make accurate predictions of how the blends are perceived in relation to the inputs. To this end, a vector containing differences

Table 1. Example of the tritone substitution cadence invention, by blending the perfect and the phrygian cadences.

Property name	Input 1 (Perfect)	Input 2 (Phrygian)	Possible blend
fcRoot	7	10	1
fcType	[0 4 7 10]	[0 3 7]	[0 4 7 10]
fcPCs	$\{7, 11, 2, 5\}$	$\{10, 1, 5\}$	$\{11, 1, 5, 8\}$
rDiff	5	2	1
DIC0	1	0	0
DIC1	1	0	1
DIC-1	1	1	1
DIC	$[1, 2, 0, 2, 0, 1, 2,\ 1, 0, 1, 2, 0]$	$[1, 0, 1, 1, 1, 0, 0,\ 3, 0, 0, 1, 1]$	$[2, 1, 0, 1, 2, 0, 1,\ 2, 0, 2, 0, 1]$
asc	1	0	1
desc	0	1	1
semi	1	1	1

on the utilised music properties, denoted by P_i^C. Since it is assumed that not all properties are of equal importance in deciding the distance between pairs of cadences, each property (P_i^C) is assumed to have a *weight* of importance, denoted by w_i. The overall distance between cadences can be then calculated by summing the weight values of the properties that are different in these cadences. Specifically, the distance between two cadences X and Y is calculated by:

$$D(X,Y) = \sum_{i=1}^{11} w_i \ f_i, \tag{1}$$

where f_i is a function related to how distance is measured for each property, as analysed in Eqs. 2, 3 and 4.

Properties with indexes 1, 2, 4, 5, 6, 7, 9, 10, 11 have a binary f_i function similar to the Kronecker delta function:

$$f_i = \begin{cases} 1, & \text{if } P_i^X \neq P_i^Y, \\ 0, & \text{if } P_i^X = P_i^Y \end{cases}, \text{ for } i \in \{1, 2, 4, 5, 6, 7, 9, 10, 11\}. \tag{2}$$

Equation 2 indicates that these properties need to be equal in both cadences in order not to be penalised by the respective w_i values. The function for property 3 is computing the number of common over the number of total pitch classes in the first chords of two cadences. Specifically,

$$f_3 = \frac{N(\cup(P_3^X, P_3^Y)) - N(\cap(P_3^X, P_3^Y))}{N(\cup(P_3^X, P_3^Y))}, \tag{3}$$

where $N(\cap(P_3^X, P_3^Y))$ and $N(\cup(P_3^X, P_3^Y))$ is the number of elements in the intersection and union of the pitch class sets. Equation 3 indicates that there is a proportional penalty to w_i for pitch classes that are not common in the first chord of two cadences. Finally, DIC information (property 8) is measured according to the correlation of the DIC vectors of the cadences under examination. Weaker

correlations are penalised proportionally with regards to w_8, according to the following equation:

$$f_8 = (1 - \text{corr}(P_8^X, P_8^Y))/2. \tag{4}$$

Correlation between DIC vectors conveys harmonic meaning at some extend, as indicated by the genre categorisation results based on DIC correlation reported in [4].

By calculating the distances between all pairs of the examined cadences according to Eq. 1, a dissimilarity matrix that represents the pairwise differences among the nine cadences is constructed. This dissimilarity matrix is subsequently analysed through non-metric weighted MDS and results in a spatial configuration of the cadences that from now on will be called the 'algorithmic space'. Therefore, in order to define the contribution (i.e. weight value) of each parameter on deciding the overall distance, a differential evolution (DE) algorithm [18] was used to optimise the fit between pairwise distances in the perceptual space (used as ground truth) and the respective ones in the algorithmic space. An overview of the optimisation process is schematised in Fig. 1.
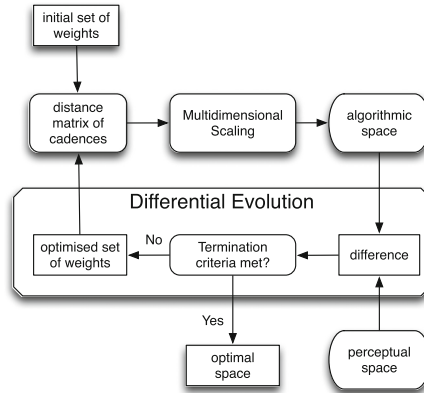

Fig. 1. The optimisation through the Differential Evolution.

The difference between the perceptual and algorithmic spaces is quantified through a fitness function that is estimated by taking the average of two similarity metrics, namely the m^2 statistic for Procrustes analysis [11] and Tucker's congruence coefficient [3]. For a detailed discussion on the application of these metrics to comparison between timbre spaces please see [20].

4   Perceptual Experiment

A pairwise dissimilarity listening test was deemed appropriate to act as a ground truth for modelling how a set of musical cadences is perceived by listeners, as

the dissimilarity matrices it produces allow Multidimensional Scaling (MDS) analysis to generate geometric configurations that represent the relationships between percepts.

Twenty listeners (age range = 18–44, mean age 24.9, 10 male) participated to the listening experiment. Participants were students in the Department of Music Studies at the Aristotle University of Thessaloniki. All of them reported normal hearing and long term music practice (16.5 years on average, ranging from 5 to 35).

Participants were asked to compare all the pairs among 9 cadences using the free magnitude estimation method. Therefore, they rated the perceptual distances of 45 pairs (same pairs included) by freely typing in a number of their choice to represent dissimilarity of each pair (i.e., an unbounded scale) with 0 indicating a same pair. Each stimulus lasted around 4 s and interstimulus interval was set at 0.5 s. Listeners became familiar with the range of cadences under study during an initial presentation of the stimulus set (random order). This was followed by a brief training stage where listeners rated the distance between four selected pairs of cadences. For the main part of the experiment, participants were allowed to listen to each pair of cadences as many times as needed prior to submitting their dissimilarity rating. The pairs were presented in random order and participants were advised to retain a consistent rating strategy throughout the experiment. In total, the listening test sessions, including instructions and breaks, lasted around thirty minutes for most of the participants.

The stimulus set consisted of the two input cadences (the perfect and Phrygian) together with seven blended cadences. The selection of cadences was made manually after evaluating their blending elements so as to attain a theoretically valid, maximally diverse corpus. All cadences were assumed to be in C minor tonality/modality, consisted of two chords and the final chord was kept constant (C minor), thus variation between the stimuli resulted from altering the penultimate chords. The nine cadential pairs of chords are described from a music-theoretical perspective in the following list:

1. Perfect authentic cadence, featuring the full V7 dominant chord that resolves to the i tonic chord without 5th, in order to achieve correct voice leading.
2. Phrygian cadence, with the ♭vii chord in first inversion resolving to the i tonic chord.
3. Tritone substitution progression, with the ♭II7♭ chord (German-type augmented-6th chord) leading to the tonic.
4. Backdoor progression, with the ♭VII7 chord in first inversion, in order to achieve maximum voice-leading uniformity.
5. Contrapuntal-type tonal cadence, with the viio6 resolving to the minor tonic.
6. Plagal-type cadence, with the iio6/5 progressing to the tonic.
7. Minor-dominant to minor-tonic progression, utilising chords from the natural minor scale (Aeolian mode).
8. Altered dominant-7th chord to minor-tonic progression, with the dominant in second inversion and with its 5th lowered (French-type augmented 6th chord).
9. Half-diminished 'dominant'-7th chord to minor-tonic progression.

5   Results

Before proceeding to the main body of the analysis for the dissimilarity data
we examined the internal consistency of the dissimilarity ratings. Cronbach's
alpha was .94 indicating high inter-participant reliability. In the main body of
the analysis, the dissimilarity ratings within each linguistic group were analysed
through non-metric (ordinal) MDS with dimension weighting (INDSCAL within
SPSS PROXSCAL algorithm)   [15]. A two-dimensional solution was deemed
optimal for data representation as the improvement of measures-of-fit when
adding a third dimension was minimal. Figure 2a and shows the configuration of
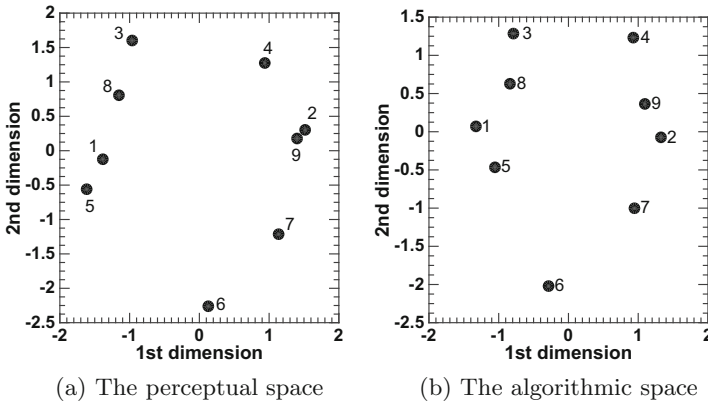the cadences within this 2-D space.

(a) The perceptual space   (b) The algorithmic space

Fig. 2. (a) The 2 dimensional perceptual space of the nine cadences. The perfect and
the Phrygian cadences (No. 1 & 2) are positioned far away from each other on the 1st
dimension. (b) The optimised algorithmic cadence space that resulted from modelling
the perceptual space through optimal weighting of the musical parameters.

The optimisation process presented in Sect. 3 produced combinations of
weights for each cadential feature that minimise the differences between the per-
ceptual and the algorithmic spaces, providing an optimised modelling of cadence
perception. The ideal combination should offer the highest possible fit (quantified
by the similarity metrics) with the perceptual space.

Several optimisation simulations were run with different setups for the DE
algorithm (concerning population members, number of generations etc.). Table 2
shows a property weight configuration that provided a satisfactory modelling of
the perceptual space (population members = 50, number of iterations = 30).
This configuration featured an $m^2 = .027$, a congruence coefficient = .994 and
an RV modified coefficient = .966 indicating an excellent fit between the per-
ceptual and algorithmic configurations. Had no optimisation taken place (i.e.,
all parameters assigned importance equal to 1), the similarity metrics between
the two configurations would become: $m^2 = .361$, congruence coefficient = .942

and an RV modified coefficient = .523, representing a serious divergence of the algorithmic space in relation to the perceptual. Figure 2b shows the optimised 2-dimensional configuration of the algorithmic space. As expected, based on the similarity measures reported above, the two spaces (perceptual and algorithmic) are very closely related since the obtained algorithmic space maintains the majority of the perceptual relationships between cadences.

Table 2. Optimal weights of the musical properties for modelling cadence relationships. A combination of four prominent parameters (in bold) and two weaker ones (in italics) achieved an excellent model of the perceptual space.

Weights (w_i values)										
fcRoot	fcType	fcPCs	rDiff	DIC0	DIC1	DIC-1	DIC	asc	desc	semi
.020	**.790**	**.751**	.000	.311	.014	.270	**.760**	**.731**	.001	.003

Based on the above, it can be concluded that the penultimate chord types, their pitch classes, the information provided by the DIC vector and the presence or absence of a leading note account for the way listeners perceived the relationships of cadences within this particular set. It should be also noted that according to the DE simulations there are two additional properties of 'moderate' importance: those that examine whether there is at least one fixed (*DIC0*) or one descending by one semitone (*DIC-1*) pitch class.

6   Conclusions

This paper presents a first pilot study towards increasing the self-evaluation ability of a creative cadence blending algorithm, by utilising data from perceptual experiments. The listening experiment incorporated two cadences (the perfect and the phrygian) as a starting point along with seven system-produced blends. The blending algorithm combines characteristics of the input spaces, generating several blends that include different combinations of characteristics from the input cadences. Aim of this paper is to identify whether any of the cadence characteristics are perceptually more salient in defining pairwise cadence similarities. The Differential Evolution (DE) algorithm was employed in order to fine-tune the salience weight of every cadence property, so that the relative placement of cadences obtained with system-based metrics optimally matches the user-obtained perceptual space.

The dissimilarity rating experiment revealed a categorical perception of cadences reflected by positioning on the 1st MDS dimension and clearly dictated by the existence of an upward semitone movement to the tonic (leading note) in the left-hand cadences in comparison to the lack of a leading note in the right-hand cadences. This fact is also evident by the high weight of the *asc* value shown in Table 2. Two major clusters of cadences were formed based on

this differentiation together with one outlier (the plagal cadence) that featured neither and upward semitone nor an upward tone to the tonic but a duplication of the tonic. It is also shown that both the intra and inter-cluster relations could be adequately modelled mainly through four salient musical properties, namely the penultimate chord type, its pitch classes, the DIC vector of the cadence and the existence of the leading note.

This being a pilot study, the generalisation of these findings for a wider range of cadences as well as a detailed mapping of musical properties to perceptual dimensions is a necessary step that is left for future work. For instance, initial findings indicate that the differentiation of cadences along the 2nd dimension could be explained by the inherent dissonance of the penultimate chords (as expressed by the MIR Toolbox roughness calculation) together with their distances from the final chord in Lerdahl's Tonal Pitch Space. Identification of such complementary measures could help towards increasing self-awareness of a cadence blending system, according to various diverse aspects of its creative products.

References

1. Antovic, M.: Musical metaphors in serbian and romani children: an empirical study. Metaphor Symbol **24**(3), 184–202 (2009)
2. Antovic, M.: Musical metaphor revisited: Primitives, universals and conceptual blending. Universals and Conceptual Blending (2011)
3. Borg, I., Groenen, P.J.F.: Modern Multidimensional Scaling: Theory and Applications, 2nd edn. Springer, New York (2005)
4. Cambouropoulos, E.: A directional interval class representation of chord transitions. In: Proceedings of the Joint Conference ICMPC-ESCOM (12th International Conference for Music Perception and Cognition, & 8th Conference of the European Society for the Cognitive Sciences of Music) (2012)
5. 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) (2014)
6. Cambouropoulos, E., Katsiavalos, A., Tsougras, C.: Idiom-independent harmonic pattern recognition based on a novel chord transition representation. In: Proceedings of the 3rd International Workshop on Folk Music Analysis (FMA) (2013)
7. Eppe, M., Confalonier, R., Maclean, E., Kaliakatsos-Papakostas, M., Cambouropoulos, E., Schorlemmer, M., Codescu, M., Kühnberger, K.U.: Computational invention of cadences and chord progressions by conceptual chord-blending. In: International Joint Conference on Artificial Intelligence (IJCAI) (2015)

8. Eppe, M., Maclean, E., Confalonieri, R., Kutz, O., Schorlemmer, M., Plaza, E.: ASP, amalgamation, and the conceptual blending workflow. In: Calimeri, F., Ianni, G., Truszczynski, M. (eds.) LPNMR 2015. LNCS, vol. 9345, pp. 309–316. Springer, Heidelberg (2015)

9. Fauconnier, G., Turner, M.: The Way We Think: Conceptual Blending and The Mind's Hidden Complexities. Basic Books, New York (2003)

10. Goguen, J.: Mathematical models of cognitive space and time. In: Andler, D., Ogawa, Y., Okada, M., Watanabe, S. (eds.) Reasoning and Cognition, vol. 2. Keio University Press (2006)

11. Gower, J.C., Dijksterhuis, G.B.: Procrustes Problems. Oxford University Press, Oxford (2004)

12. Jordanous, A.K.: Evaluating computational creativity: a standardised procedure for evaluating creative systems and its application. Ph.D. thesis, University of Sussex (2013)

13. Kaliakatsos-Papakostas, M., Confalonier, R., Corneli, J., Zacharakis, A., Cambouropoulos, E.: An argument-based creative assistant for harmonic blending. In: Proceedings of the 7th International Conference on Computational Creativity (ICCC) (2016)

14. Kaliakatsos-Papakostas, M., Zacharakis, A., Tsougras, C., Cambouropoulos, E.: Evaluating the General Chord Type representation in tonal music and organising GCT chord labels in functional chord categories. In: ISMIR. Malaga, Spain (2015)

15. Meulman, J.J., Heiser, W.J.: PASW Categories 18. SPSS Inc., Chicago (2008). Chapter 3

16. Ontañón, S., Plaza, E.: Amalgams: a formal approach for combining multiple case solutions. In: Bichindaritz, I., Montani, S. (eds.) ICCBR 2010. LNCS, vol. 6176, pp. 257–271. Springer, Heidelberg (2010)

17. 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)

18. Storn, R., Price, K.: Differential evolution - a simple and efficient adaptive scheme for global optimization over continuous spaces. J. Global Optim. **11**, 341–359 (1997)

19. Zacharakis, A., Kaliakatsos-Papakostas, M., Cambouropoulos, E.: Conceptual blending in music cadences: a formal model and subjective evaluation. In: ISMIR. Malaga (2015)

20. Zacharakis, A., Pastiadis, K., Reiss, J.D.: An interlanguage unification of musical timbre: bridging semantic, perceptual and acoustic dimensions. Music Percept. **32**(4), 394–412 (2015)

21. Zbikowski, L.M.: Conceptualizing Music: Cognitive Structure, Theory, and Analysis. Oxford University Press, New York (2002)