Conceptual Blending of High-Level Features and Data-Driven Salience Computation in Melodic Generation

Maximos Kaliakatsos-Papakostas and Emilios Cambouropoulos,
School of Music Studies,
Aristotle University of Thessaloniki,
Thessaloniki, Greece.
{maxk, emilios}@mus.auth.gr

Abstract

Conceptual Blending (CB) theory describes the cognitive mechanisms behind the way humans process the emergence of new conceptual spaces by blending two input spaces. CB theory has been primarily used as a method for interpreting creative artefacts, while recently it has been utilised in the context of computational creativity for algorithmic invention of new concepts. Examples in the domain of music include the employment of CB interpretatively as a tool to explain musical semantic structures based on lyrics of songs or on the relations between body gestures and music structures. Recent work on generative applications of CB has shown that proper low-level representation of the input spaces allows the generation of consistent and sometimes surprising blends. However, blending high-level features (as discussed in the interpretative studies) of music explicitly, is hardly feasible with mere low-level representation of objects. Additionally, selecting features that are more salient in the context of two input spaces and relevant background knowledge and should, thus, be preserved and integrated in new interesting blends has not yet been tackled in a cognitively pertinent manner. The paper at hand proposes a novel approach to generating new material that allows blending high-level features by combining low-level structures, based on statistically computed salience values for each high-level feature extracted from data. The proposed framework is applied to a basic but, at the same time, complicated field of music, namely melodic generation. The presented examples allow an insightful examination of what the proposed approach does, revealing new

possibilities and prospects.

Keywords: Conceptual Blending, Computational Creativity, Feature Salience, Genetic Algorithms, Melody Generation

1. Introduction

Computational creativity studies the processes and products of computational systems when generating something new, in combination with the counterparts of human creativity [1]. Creativity in general is about generating novel material either by exploring unexplored regions of a given conceptual space through exploratory creativity, by using transformational creativity to transform rules that describe concepts, or by associating and combining diverse conceptual spaces with combinational creativity [2]. According to Boden [3], combinational creativity is among the hardest forms of creativity to describe formally. The theory of Conceptual Blending (CB) is about combinational creativity; it was developed by Fauconnier and Turner [4] and it describes the cognitive mechanisms that allow creative combinations of elements from diverse conceptual spaces, leading to the emergence of new conceptual spaces that incorporate new interesting properties. In music, as well as in other fields, CB theory has been primarily used for interpreting relations between high-level musical concepts and extra-musical meaning in existing works, e.g. see [5, 6, 7]. For instance, Zbikowski [7] explains the relation of a concept given in lyrics of a work by Palestrina, i.e., "falling from heaven" with the utilised musical concept of a descending passages; this relation exists on a high-level descriptions of elements in a musical surface.

Conceptual Blending has also been studied as a method for generating new concepts, rather than merely interpreting existing ones. Interesting results have been presented in many fields, e.g. for the creation of mathematical concepts [8, 9]. Other notable examples are in music, specifically, where generative implementations based on conceptual blending have been presented for the invention of blended cadences [10, 11], melodies [12], or the generation of blended harmonic spaces through blending chord transitions [13, 14]. The concepts describing the cadences or the chord transitions in the latter studies, incorporated the analytical description of specific low-level elements, e.g. the root notes, chord types, existence of leading note to the tonic etc. Blending, therefore, involved the combination of such low-level structural elements and not *concepts* describing qualitative features (e.g. emotions, ideas, concepts as with Zbikowski's [7] high-level descriptions) of the blended spaces. Additionally, the determination of which features of each input should be included in the blend in some studies was assigned "by hand", in the sense that, even though the most salient features of the inputs were automatically selected by the blending algorithm, salience values were manually assigned to each feature.

In the examined domain of music, statistical learning plays a significant role in music cognition via the exposition of listeners to available musical stimuli [15]. This is evident by the fact that empirical experiments on the perception of music structures correlate with the statistical findings in musical corpora; the tonal centre and the mode are examples of such correlations [16, 17]. Additionally, listeners with different backgrounds are expected to interpret musical information differently, according to the *cultural distance* hypothesis, according to which "the degree to which the musics of any two cultures differ in the statistical patterns of pitch and rhythm will predict how well a person from one of the cultures can process the music of the other." [18] Several empirical studies have been conducted that examine the differences in music perception between listeners as a function of their cultures, e.g. see [19, 20] among others (some of which are discussed in [21]), which involve listening tests for inferring the musical schemata learned by listeners in diverse musical cultures.

It appears that listeners from different backgrounds develop "sensitivity" to different schematic/high-level attributes of music. This sensitivity to specific feature values can be measured by a "salience value": higher salience values in a feature of an object makes this feature a more decisive factor about recognising this object. Features are not only related with an object, but also with a category: e.g. humans recognise the category of zebras mostly by their colour feature. Inferring the salience (or prominence) of musical features has been studied in the computational methods as the "Unscramble" algorithm [22], which has been used for clustering objects [23], e.g. musical patterns in "Träumerei" [24].

Enabling computers to consider the conveyed high-level features and their accompanying salience values in musical excerpts and their categories, is a direction that becomes all and more feasible, but also necessary for developing next generations of tools for music information retrieval [25] and, subsequently, computational creativity in music. In the paper at hand it is maintained that the utilisation of generative CB for computational creativity can be enhanced by introducing blending of high-level concepts instead of merely blending structural elements.

Generating music by utilising high-level features as target-features is a field that has received increased attention the last decades with methodologies that pertain to evolutionary computation [26]. New ways for generating meaningful sets of desired target features have been researched ,e.g. for generating rhythm variations based on variation percentage [27, 28, 29, 30, 31]. Using CB to generate target combinations of the most salient high-level feature of input spaces appears to be a perfect fit for such algorithms. The paper at hand proposes an extension of creative CB to a data-driven workflow that blends higher-level features extracted from input spaces, based on the distinction of the most important high-level features through automatically assigned salience values. The blended features are then used as target features for evolutionary algorithms that combine low-level information from the inputs and generate output that reflects these features. The field of application of this methodology is melodies and specifically blending of Chinese and German traditional melodies. Previous work on blending drum rhythms [32] demonstrated a more basic approach of the methodology presented herein with a large amount of employed drum rhythm features (40) that did not allow a clear intuitive assessment of what the algorithm really does. The work presented here employs 6 simple melodic features that allow direct intuitive insights about what the proposed methodology does.

2. Motivation for New Approaches in Generative Conceptual Blending

This section defines the *representation* and the *salience* problems of the current framework for generative conceptual blending of low-level concepts/attributes. Through an intuitive discussion around basic cognitive principles, the proposed methodological step to tackle these problems are illustrated.

2.1. Addressing the Representation Problem: High-Level Features and Conceptual Blending

The theory of CB was first employed as a method for interpreting creative ideas rather than generating new ones. The interpretative (not generative) power of the Conceptual Blending theory has produced important results in providing explanations about how abstract concepts from different domains relate with each other, generating new conceptual spaces of abstract concepts. There is extensive scientific work where the CB theory was employed for interpreting creative outcomes of humans in diverse areas, with a common starting point in cognitive linguistics. Among many papers in many areas, some indicative studies present the employment of conceptual blending theory to explain the metaphorical expression of time [33] and the dialogues between beloved persons [34] in poetry, to analyse how abstract concepts have been blended in successful advertisements [35] and news headlines [36, 37].

Modelling musical concepts has proven useful for studying the cognitive grounding of structures in music theory [38], analyse findings in empirical tests on the construction of elementary musical concepts [39] and compact musical structures (e.g. motives, themes and chords), or even complete music parts [7, 40]. The common idea that these studies (among many others) build on, is that musical ideas are related with extra-musical meaning via schemas [41, 42], which are abstract concepts describing general attributes and relations in human perception and cognition. In music, the idea of schemas is generally different from the one related with studies on analogy and is mainly associated with tools that create abstractions from musical excerpts and facilitate the acquisition of mental knowledge structure [43]. Example of such abstractions, as studied in [43], are the concepts of tonal centre and mode [16], which humans unconsciously extract when exposed to musical stimuli. Those abstractions allow listers to relate and compare musical excerpts on more abstract levels, for example: two pieces are similar in terms of the emotion they elicit (e.g. both sound "happy" because they both utilise elements of a major scale similarly), but they are not in the same key (because their tonal centre differs). Similarly, those abstractions also allow the ordering of objects, since they are quantitative in some sense, for example: one can measure if a piece A adheres more to the major scale than piece B, or if piece C is closer to piece A than to B in terms of pitch class content.

Other studies have examined *quantitative* descriptions of high-level concepts, in a sense that concepts are represented with a magnitude value describing a qualitative feature. For example, the high-level feature of "rhythm complexity" (related with syncopation) in one-bar musical rhythms is described with a numerical value that ranks rhythms according to a complexity scale [44]. Among other musical qualities, there are successful examples in the literature that relate feature extraction methods and perceived qualities of rhythm. For instance, the empirical studies presented in [45, 46] have shown that there are strong correlations between a proposed quantitative expression of syncopation and the sensation of groove in rhythms.

In contrast, the current *generative* frameworks proposed for CB incorporate analytic description of low-level properties of concepts in the input spaces based on structures related to formal logic. Even though some interesting results have been presented with currently studied low-level approaches, the scope of the results are focussed on small test-case examples. For instance, in [10] the tritone substitution cadence used in jazz music was generated by blending the Perfect and the Phrygian cadences, which are cadences used centuries before the tritone substitution cadence was introduced. The methodology for blending cadences was combined with statistical learning techniques in [47], offering a way to combine entire chord transition spaces by blending chord transitions (successive chord pairs). Even though these results are very promising, the strength of the conceptual blending is deteriorated in low-level structural information, disregarding high-level concepts that in most cases can be only approximated from low-level properties. For example, imagine that we have two melodies, one with high rhythm (`highR`) complexity and low harmonic complexity (`lowH`) and one with the opposite, low rhythm (`lowR`) complexity and high harmonic complexity (`highH`); those are high-level features. If we want to construct a new melody that has, e.g., high values for both (`highR` and `highH`), low-level blending does not explicitly allow for it. Blending low-level features may eventually lead to melodies that have the desired characteristics, but the "objective" of the blending process (i.e. having a result with `highR` and `highH`) cannot be explicitly stated. Explicit blending of high-level information is not available in current frameworks since the representation concerns explicit definition of low-level attributes; we call this problem the *representation problem*.

2.2. Addressing the Feature Salience Problem: Identifying Feature Importance through Data

The idea of creating novel concepts by "combining existing ideas and concepts in a manner useful for an intended purpose" [48] is well established in the cognitive science and artificial intelligence literature since many years. The determination of which out of many combinations of ideas and concepts serve an "intended purpose", however, is a task-dependent problem. Identifying the relevant or salient concepts from the non-relevant or non-salient concepts toward serving an intended purpose is a key-component according to Goel [49]; this segregation relates with the "frame problem" in AI [50]:

usually there are overwhelmingly more non-relevant elements, actions, concepts or combinations that need to be rejected. The aforementioned problem can be summarised in one question: how can we determine the *salience* of features that need to be combined from two input spaces in order to have *meaningful* blends? We refer to this problem as the *salience problem*.

Regarding the term *"meaningful"*, during the process of meaning construction humans tend to focus on specific aspects that generally help us compress information and obtain global insight of a field (see [4], page 312). When meaning is constructed through conceptual blending of spaces that possess meaning of their own, there are several theoretic criteria, referred to as *optimality principles* [4], that have been proposed as conditions that would make the generated space *meaningful*; the implementation of such criteria in specific applications depend on the "intended purpose" of each application (for examples where such criteria have been implemented see [51, 52]). Humans are good at compressing information and generating abstractions by creating categories of objects on many levels of detail, e.g., even though there are exceptions, birds have wings; fishes live in the sea; zebras have black and white stripe patterns. Those category-related features that allow us to separate objects into classes are related with the concept of the *salience* of those features in their class and between classes. For example, a zebra without the stripe patterns would most possibly be categorised as a horse, while even a dog with the zebra stripe pattern would directly evoke the image of a zebra; therefore the stripe pattern feature is salient for the zebra category. The salience of features does not only allow the compression of the (input) conceptual spaces, but also facilitates the acquisition of global insight in the blended space by allowing the inference of the involved input spaces.

For illustrating the way that the notion of feature salience can potentially offer global insight in the process of "meaning" construction through CB, let us use some intuitive and simple examples from the toy-domain of animal blending [53]. Let us imagine the example of blending a zebra with a shark. Let us also keep in mind that there exists a fish called the "zebra shark", which scientists, creatively indeed, named this way because of its colour and shape. In this example we are going to follow the reverse process of "constructing" such a fish as if it did not exist. What would a good blend between a zebra and a shark look like? A good blend would creatively combine the elements of both inputs (zebra and shark), allowing the observer to distinguish as clearly as possible that a zebra and a shark are involved in this blend. Many blends can be constructed with those inputs and some of those blends could be characterised as successful, depending on the purpose that the blends are expected to serve. For instance, the blend of a zebra with a shark fin on its back could serve the purpose of a four-leg peaceful earthly animal. Similarly, a shark with zebra stripes would be a good blend of a dangerous predator that lives in the sea and, when flocking with other individuals of its species, can confuse other predators because of their combined black-white colour patterns. However, the blend of a zebra with grey colour (colour of a shark) would not be successful, since this blend would look as a simple, non-blended grey horse – at least to someone who had no information about the fact that this is the output of a blending process with these specific inputs. Similarly, the blend of a shark with four legs (without the zebra stripes) would not be successful, since many other animals with four legs could have produced this blend with a shark.

The distinctive characteristics of zebras are in general their black and white stripe patterns. If this element is not present in any blend that includes a zebra as input, then the blend would possibly fail to convey the existence of a zebra, since almost all the remaining characteristics of zebras are similar to characteristics of other animals, especially horses. Similarly with sharks: if we use, e.g., only the grey colour property, then the resulting blend would not necessarily indicate the involvement of a shark since many animals have grey colour. Therefore, in the aforementioned example successful blends need to incorporate features that are distinctive or *salient* for the involved categories – the "`Zebra`" and the "`Shark`" categories in the running example. Regarding the actual zebra shark, it appears that biologists have given this name to this type of fish by decomposing the most salient features of a shark and a zebra, which are naturally evoked to the observer of such a fish. Data play an important role in defining the salient characteristics of a class, since the salience of the feature is related with the "commonality" of this feature in the class. Even though complex cognitive processes may play significant role regarding how humans perceive the salience of a feature, this paper proposes an approach that is based on statistical learning; the formalisation of this data-related approach is given in Section 3.3

3. Methodological implementation and application of high-level data-driven Conceptual Blending

The aim of the proposed methodology is to incorporate high-level features in generative conceptual blending (addressing the *representation problem*),

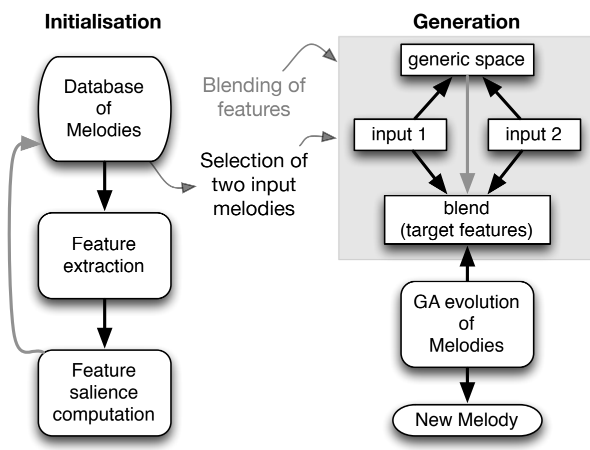


Figure 1: Overview of the proposed methodology. During the initialisation phase (left), features are extracted from the melodies in the database and the salience of each feature in every melody is computed. After two input melodies are selected, their features are blended (top-right), producing the target *blended* features for the Genetic Algorithm which constructs the output blended melody (bottom-right).

while selecting the high-level features of inputs to be included in the blend according to their salience values that are computed from data (addressing the *salience problem*); the output will be new objects (blends) that reflect the blended input high-level features and incorporate low-level information inherited from the input objects. This methodology is applied to melodic generation in Section 4, using as background knowledge sets of melodies extracted from the Essen corpus [54] belonging to two styles (Chinese and German) with different characteristics. For each melody 6 quantitative high-level features are extracted, which are descriptive of some basic rhythmic and melodic cognitively-based qualities. Figure 1 illustrates an overview of the proposed methodology applied to melodies; the methodology includes an initialisation stage and a generation stage. In the initialisation stage (left side) features are extracted from each melody in the database of melodies and afterwards the salience values of each feature in every melody are computed (with a process described in Section 3.3). After the initialisation step the database includes melodies that are accompanied by a vector of features and a vector of the corresponding saliences for each feature (one salience value for each feature in each melody).

The melody generation phase shown on the right side of Figure 1, results in the generation of a new melody (blend) that encompasses the blended features of two input melodies from a database. The generation phase includes two steps: the feature blending and the melodic composition step. In the feature blending step (top right), the features of the selected melodies are blended, producing a set of *blended target features* that the blended melody should incorporate. The blended target features are selected through a process described in Section 3.3, which includes a balanced (among the two inputs) selection of the most salient features of selected input melodies. The blend of features constructed in the top-right side of Figure 1 are then employed as fitness targets for the Genetic Algorithm (GA) (bottom-right), generating a melody that best matches the target features.

Melody representation:. Melodies are represented both by low and high level information. Low level information for every note in the melody is given as pairs of pitch and onset time values. Additionally, during blending the Markov probability matrix of pitch transitions is also considered, as described in Section 3.4. Regarding high-level representation, a basic set of six features is used that are related with cognitively-relevant rhythmic and pitch attributes. A greater number of such features could have been used (e.g. see [55]) but only a representative number of two rhythmic and four pitch-related features was sustained for focussing more clearly on the effects of high-level feature blending. These features are forming a vector that describes each melody in the database, $\vec{m}_i = \{f_1^{(i)},\ f_2^{(i)},\ \dots,\ f_6^{(i)}\}$, where i is the index of the melody. A list of these features along with a short description is given as follows:

1. *Rhythm inhomogeneity:* ($\mathrm{r}_{\mathrm{hom}}$) Rhythm inhomogeneity can be viewed as an aspect of rhythm complexity in terms of length distribution of successive onset intervals, where more similar intervals (fewer rhythmic deviations) elicit the sensation of a less complex rhythm and vice versa [56, 27]. Methods for measuring rhythm complexity rhythm complexity span from simpler [57] to more complex [58, 59]; for this feature, rhythm inhomogeneity is approached by the value of the standard deviation of all inter-onset intervals over the value of the mean inter-onset interval in a melodic sequence.
2. *Pitch-class set complexity:* ($\mathrm{n}_{\mathrm{pcp}}$) This feature utilises the information entropy [60] of the Pitch Class Profile (PCP) set distribution of a melody for determining the relative quantities of each pitch class (pc)

used in the melody; the closer a pc distributions is to the uniform distribution, the more complex the PCP content of the melody.

3. *Small pitch intervals:* (n_{int}) Even though there appear to be global cognitive mechanisms that favour small pitch intervals (semitone or tone) [15], different percentages of small intervals are used in different styles. This feature measures the percentage of small successive pitch intervals among all successive pitch intervals.
4. *Pitch range:* (n_{range}) The difference between the smallest and the greatest midi pitch values in the melody over the ration of two octaves (24) – all involved melodies are within two octaves.
5. *Note repetitions:* (n_{rep}) The ratio of constant note intervals over the total number of intervals in a melody.
6. *Rhythm repetitions:* (r_{cns}) The ratio of consecutive constant rhythm intervals over the total number of rhythm intervals.

Since the music-related part of the methodology is meant to be kept simple (for focussing on the CB aspects), the information of tonality variability is "neutralised" – even though in future work "advanced" musical concepts related to notions of tonality variability could be used in the feature representation of melodies. To this end, all melodies are transposed to the keys of C major or A minor (depending on their initial tonality).

3.1. Conceptual Blending of melody features

Features blending in the proposed methodology employs a basic computational framework of the theory of CB proposed by Fauconnier and Turner [4] and formalised by Goguen [51]. The employed framework is based on the framework developed during the Concept Invention Theory (COINVENT[1]) project [61]; in the context of this first application, however, some simplifications are applied for making the blending process more straightforward. According to this theory, formalisation and implementation, two *input spaces* are described as sets of properties and relations. The *generic space* of these inputs is computed, which is the conceptual space that keeps the common structure of the input spaces, guaranteeing that this structure also exists in the blended space, and generalises or abstracts over the parts of the inputs that are distinct. Afterwards, an *amalgamation* process [62, 63] generalises

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

the input concepts until the generic space is found and "combines" generalised versions of the input spaces to create blends that are "*consistent*" or satisfy certain properties that relate to the knowledge domain. Regarding blends, the term "consistent" refers to whether all logical relations in the blend and the background knowledge are satisfied, i.e. there are no mutually canceling contradictions.

The example of the zebra shark discussed above is illustrated in Figure 2 under the context of the COINVENT CB methodological framework. The two inputs considered are T_1 : "*Shark: a grey fish with fin*" and T_2 : "*Zebra: a striped horse-shaped animal*". There are many possible blends and many possible generalisation alternatives – under the constraint imposed by the generic space that all blends should correspond to the generic category: "*An organism with colour and shape*". Possible generalisations of T_1 are "*A grey organism*" or "*A fish with fin*". Accordingly there are many possible blends between T_1 and T_2 arising from those generalisation alternatives, e.g. "*A grey horse-shaped animal*" or "*A striped fish with fin*" (the actual zebra shark). There might be also inconsistent blends, e.g. "*A grey horse-shaped fish*" (this can be considered inconsistent since there cannot be horse-shaped fish), and, therefore, consistency check is necessary after a blend has been constructed. The value of each blend is assessed through blending *optimality principles* [4, 51, 52]. Even though there are extensive theoretical descriptions on optimality principles, e.g. see [4], the algorithmic implementations of such principles depend on the specific domain of application.

The first approach to melodic feature blending proposed in this paper, uses a basic and simplified version of the amalgamation process where no relations between features are considered, i.e., the value of one feature is not related with the value of the other (even though this should not necessarily be the case). Since relations between features are not considered, the amalgamation process plays the simple role of combining highly salient features of the input melodies, considering, however, the generic space restrictions. Figure 3 abstractly illustrates the feature blending process that generates the "target features" that are subsequently fed into the genetic algorithm (discussed in Section 3.4). The algorithm for this process (Algorithm 1) is described in Section 3.3.

3.2. Generic space

The generic space is the space of common features in the two input spaces; identifying common features is important in the conceptual blending theory

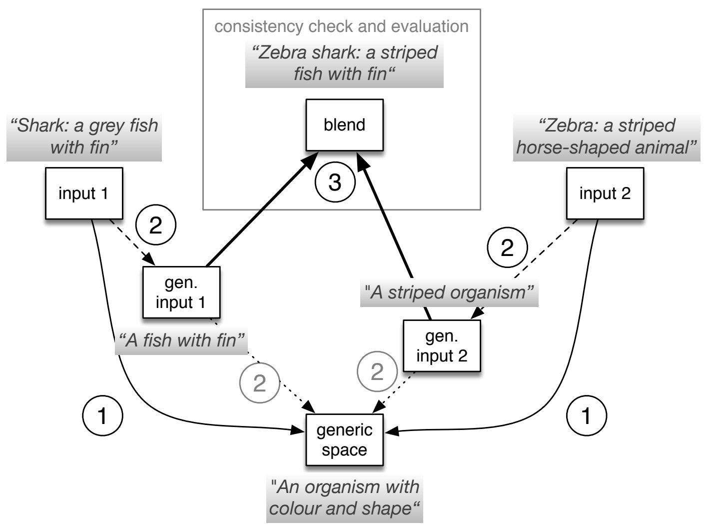


Figure 2: Illustration of the zebra shark example through a generative approach to conceptual blending based on amalgamation. The generic space is computed (1) and the input spaces are successively generalised (2), creating new potential blends (3). Some blends might be inconsistent or poorly evaluated according to domain specific criteria.

since those features should be identifiable in blend as well. The role of the generic space is to guarantee that the blend satisfies this property. In the theory of CB the generic space includes common aspects of the inputs, expressed as abstract concepts that have been approached by the utilisation of image schemas [42]; earlier, in Section 2, the notion of schemas and how it relates with high-level features was discussed.

In the generative approaches of conceptual blending proposed in the literature [52, 10, 13, 53, 62], the common structures and properties in the input spaces are forming the generic space. The generic space under the perspective of feature blending, on the other hand, can be used to provide numerical limits to which feature values can be considered the same, according to a predefined degree of granularity. For instance, if the two input melodies incorporate high levels of rhythm inhomogeneity, then the generic space should include the requirement that the blend should be a melody that also has a high value for this feature.

The notion of the generic space for spaces represented by continuous numerical features needs to be defined, since it differs from the logical-related formulation of "discrete" feature terms [64] that has been hitherto utilised in the literature [52, 10, 13, 53, 62]. For instance, if both inputs in the car blending example involve a red car, then the colour feature is included in the generic space; in case the input cars have different colours, the colour property would remain empty in the generic space, allowing cars of whatever

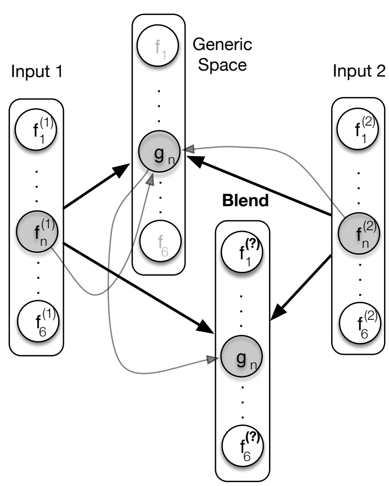


Figure 3: Simplified blending of melody features.

colour to be generated. However, in the proposed real-valued feature representation, the notion of whether two inputs have the "identical" feature is replaced by the notion of whether they have "similar" features.

Two melodies are considered to have a "similar" value in one of their features if their numeric difference is smaller than a given value. By convention, the values of a feature t in two melodies $\vec{m_{i_1}}$ and $\vec{m_{i_2}}$, denoted as $f_t^{(i_1)}$ and $f_t^{(i_2)}$ for $\vec{m_{i_1}}$ and $\vec{m_{i_2}}$ respectively, are considered "similar" if their distance is smaller than (the arbitrarily selected) 1/10-th of the range of all values of this feature in the dataset. In case $f_t^{(i_1)}$ and $f_t^{(i_2)}$ are similar, the generic space value for feature t, denoted as g_j in Figure 3, is the mean value of these features, which is directly passed to the blended set of features. In the abstract example of Figure 3, $f_n^{(1)}$ and $f_n^{(2)}$ in the two input melodies are considered similar and thus this feature is represented by a fixed value in the generic space (g_n) with a value equal to the mean of $f_n^{(1)}$ and $f_n^{(2)}$; subsequently this value (g_n) is also used in the blend.

3.3. Creating the "optimal" blend

The typical amalgamation process leads to the generation of many blends that correspond to different combinations of feature values inherited from the input spaces (keeping the "reserved" feature values from the generic space). Through this process the number of possible resulting blends is usually large and the selection of the "best" blend(s) is based on domain-specific blending *optimality principles* [4, 51, 52]. Two basic qualities are utilised in the current study, namely the *balance* of features inherited from the inputs in the blend and the *salience* value of each feature discussed in Section 2.2. With the presented blend generation process only the single best blend is retrieved, while extensions are possible that allow the preservation of an arbitrary number of highly ranked blends.

Balance:. One important aspect of meaningful blends is that they reflect characteristic of both inputs; to ensure that both inputs are represented in the blend, the constructed set of blended features should balanced mixture of the features in the two inputs. The modification of the amalgamation process used in the presented approach produces only the blend with the "optimal balance". Therefore, the resulting optimal blend of features includes an equal number of features (plus/minus one) from the two input spaces, keeping in mind that some feature values are reserved by the generic space.

Salience:. Selection which features will compose the "balanced" set is also crucial for generating blends that incorporate the most important or "salient" features of the inputs. Under the statistical perspective given in Section 2.2, the salience value of a feature value, $s(f_t^{(i)})$, in a melody $\vec{m}_i$ is computed from available data. Supposing that the considered "universe" of objects in the dataset (background of a listener) comprises N categories (e.g. melodic genres), where category c_I includes the objects with indexes i_I, the *concentration* value of feature t (the centre of the area where the feature values of most objects are) in each category is computed as:

$$C_{c_I}(t) = \frac{a_r + a_{r+1}}{2}, \text{ where } r = \arg\max_k P_i(a_k \leqslant f_t^{(i)} < a_{k+1}). \tag{1}$$

$P_i(a_k \leqslant f_t^{(i)} < a_{k+1})$ is number of individual whose t-th feature value ($f_t^{(i)}$) falls within the area of a_k and a_{k+1}. The concentration value of a feature in a category, in simple terms, is the peak of the histogram regarding this feature value for all individuals[2]. The closer an object is to the concentration value of its category, the more representative it is of the features in this category and, therefore, the greater the value of salience for its features. Conversely, the farther away the concentration value of a feature in a category is from the concentration values of other categories, the more unique and therefore salient it is. The salience value of feature t for object i is obtained as follows:

$$s(f_t^{(i)}) = U(C_{c_I}(t)) \ \left(2 - \frac{2}{1 + e^{a \ |C_{c_I}(t) - f_t^{(i)}|}}\right). \tag{2}$$

$U(C_{c_I}(t))$ is an estimation of how "unique" the concentration value is for the entire category, which is computed as the ratio of the minimum distance between this and the concentration values of all other categories (normalised with the maximum distance of concentration values in the database):

$$U(C_{c_I}(t)) = \frac{\min_J |C_{c_I}(t) - C_{c_J}(t)|}{\max_J |C_{c_I}(t) - C_{c_J}(t)|}, \tag{3}$$

where J is the set of category indexes except index I. The later term is simply a normalised proximity measure between $f_t^{(i)}$ and $C_{c_I}(t)$. It has a

[2]The centroid instead of the concentration value has also been examined, however, the centroid was considered non-representative of the feature values behaviour since they did not follow normal distributions.

maximum value of 1 for highly salient features ($f_t^{(i)}$ close to $C_{c_I}(t)$) and a minimum of zero. The factor a defines the steepness of the "slope" at which feature values that are more distant from the concentration value become less salient.

Through this process, every feature vector describing a melody, $\vec{f^{(i)}}$ is accompanied by a salience vector, $\vec{s_{m_i}}$, with the respective number of values representing the saliences of each feature in $\vec{f^{(i)}}$ ($\vec{s_{m_i}} = \{s(f_1^{(i)}),\ s(f_2^{(i)}),\ \ldots,\ s(f_n^{(i)})\}$), where n is the number of features – $n = 6$ in the examined application.

Optimal Blend:. The algorithm for computing the single optimal blend of features of the two inputs is shown in Algorithm 1. The arrays of the two input feature vectors and their respective salience arrays are given as inputs along with the array that includes the generic space features. The algorithm outputs an array of blended features with the desired properties, i.e. this blended array incorporates the most balanced combination of the most salient features of the input spaces. The blended array of features generated by feature blending is then used to provide the target features in an evolutionary process, leading to the implementation of objects (melodies in the presented application) that incorporates the desired blended features.

The idea behind the algorithm for feature blending is to first assign the generic space features into the blend and then fill the remaining features by interchangeably selecting the most salient features from the input spaces. To this end, the indexes of the sorted saliences of each input space are stored in two arrays (through the `getSortedIndexes()` function in lines 3 and 4) and then the features of the generic space are passed into the blend, while the corresponding indexes are removed from the aforementioned arrays (lines 6-12). Until now the generic space requirements have been dealt with. The first elements of the index arrays ($i_1^{(1)}$ and $i_1^{(2)}$) will always correspond to the index of the feature with the highest saliences available in both inputs; in the following steps the indexes of the features that are selected for the blend are going to be removed from the sorted index arrays. In lines 14-18 the algorithm decides which input has the most salient available feature to begin the interchanging process in the remaining lines. During the interchanging process, the most salient available feature from each input space at each step of the loop beginning in line 20 is selected and put in the blend. Afterwards, the indexes of the selected features are removed from both index arrays and the process continues until all features of the blend have been filled up – which means that the index arrays have emptied.

3.4. Generating Melodies from Blended Features through Genetic Evolution

Several methods for generating melodies with evolutionary algorithms have been proposed in the literature (e.g. [65, 55] among others). The typical approach that most of those methodologies follow involves the employment of a set of target melodic features and an initial population/generation of melodies, which can be random sequences of notes and rests [65]. Afterwards, evolutionary operators are employed for producing new generations of melodies with features that gradually converge to the target features. Melodies are evolved according to the typical overall principles employed in evolutionary melodic generation, however specific genetic operators and population initialisation are examined that relate with the theory of CB.

In the low information level, melodies are represented as pairs of pitch and onset-time values (disregarding duration), along with information about pitch transitions in a Markov probability matrix. In an attempt to preserve elements of low-level blending and retain "veridical" [15] aspects of the input melodies, the initial population of melodies are exact copies of the input spaces and melodies are evolved according to genetic operators that ensure that the new (children) melodies incorporate material only from the two parent melodies. Additionally, the average Markov matrix of pitch transitions between the two inputs is used during fitness evaluation to encourage the recombinations of parts from the parents that are merged with transitions found in the input melodies. This process leads to the generation of melodies that exclusively include recombined material and also transitions found in either of the two input spaces.

The employed operators are different types of crossover, including: (a) "bar exchange" crossover, where two parent melodies exchange a bar selected in random; (b) "note exchange", where a single note is exchanged between parents; and (c) "pitches-to-rhythm" crossover where the pitches in one bar of one parent are fitted to the rhythm structure of the other parent and vice versa. If a different numbers of notes are included in the involved bars then the pitches of the shorter sequence are successively repeated until they match the rhythm events of the longer sequence, while only the beginning pitches of the longer sequence are used that match the rhythm events of the larger sequence. The fitness value (to be minimised) of new melodies is calculated as the Euclidean distance between their features vector and the vector of *blended target features* constructed as the optimal blend (Section 1) plus a "transition penalty" derived from the Kullback-Leibler divergence between the average input Markov matrix of pitch transitions and the one of the generated individual.

4. Example Applications

The effectiveness of the proposed framework for CB is demonstrated with examples that include two distinctive categories of melodies. These categories include melodies from the Eastern and Western culture and specifically a selection of pieces in four styles (two Chinese and two German) in the Essen corpus [66, 54]. A *main* set of old German songs (30 melodies the "`altdeut1`" dataset) and a set of songs from the Chinese Han culture (30 melodies in the "`han`" dataset) constitute the input material for blending, while two *secondary* sets with 30 pieces each are used for representing other styles of Chinese ("`natmin`") and German ("`zuccal`") pieces. Pieces are selected from each set that reflect the unique characteristics of the respective styles, in terms of the employed features. Figure 4 illustrates the features extracted from the selected pieces in the main sets (`han` and `altdeut1`). The selected melodies belonging to the `han` set display higher rhythm inhomogeneity and pitch range, while `altdeut1` melodies include more often smaller intervals (two semitones or less) with more complex pitch class profiles (PCP complexity). Those features reveal some basic characteristics of those styles: German melodies have more robust and predictable rhythm, while Chinese melodies use mainly notes in the pentatonic scale (thus including many intervals larger than three semitones and smaller PCP complexity).

For the remainder of this section, results are organised so as to demonstrate three possible key-applications that high-level feature blending and the statistical computation of salience could allow or enhance, namely (a) the infusion of a single high-level characteristic to an existing melody, (b) identification of exemplar melodies in a set according to listener background, leading to the generation of blended melodies based on feature salience and (c) the recommendation of new melodies based on blended features. For the (b) and (c) scenarios two "virtual" listeners are assumed with different backgrounds (using the secondary data Chinese and German datasets), similarly to the work of van der Weij et al. [67] and Pearce [21]; the Eastern listener is assumed to be exposed to sets of Chinese melodies (`han` and `natmin`), while the Western listener to German melodies (`altdeut1` and `zuccal`).

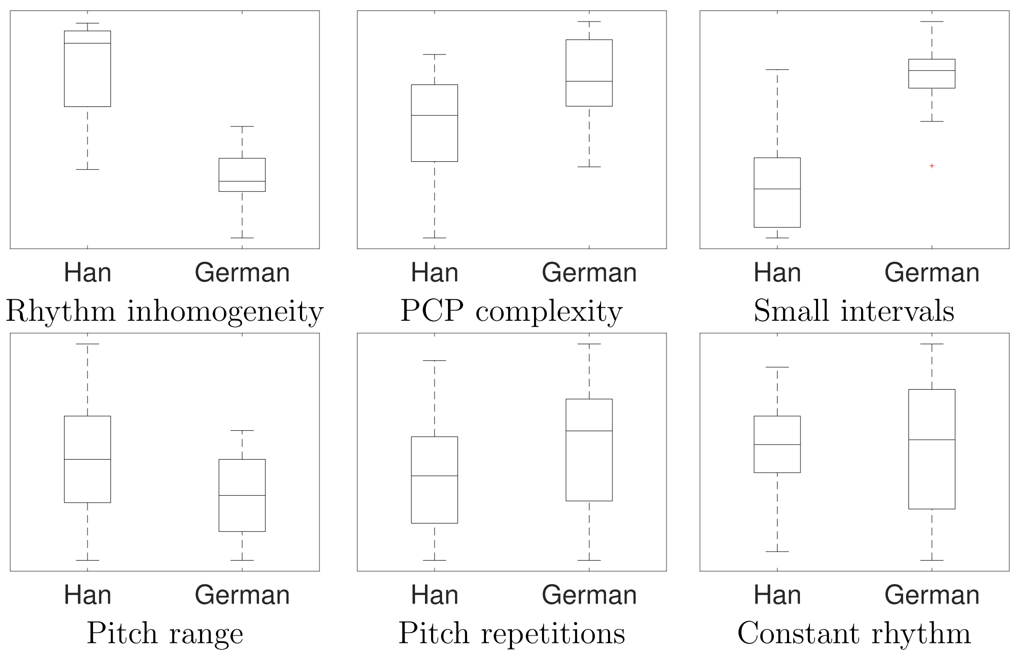


Figure 4: Extracted feature values from the "main" (Han and German) sets of pieces.

4.1. Infusion of single new characteristic through single scope blends

Feature blending allows the infusion of a single new characteristic, or a high-level feature, into a given melody. The example in Figure 5 illustrates such a scenario, where a Chinese (`han`) melody, with low pitch class profile complexity (only two notes are played) and low percentage of small intervals (only 5 and 0-semitone intervals are employed), is given as Input 1. If we want to increase the pitch class complexity of the melody, we can use a German (`altdeut1`) melody with higher value in this feature (shown as Input 2-A) and generate a blend. This blended melody, Blend-A Figure 5 is generated by substituting the pitch class complexity feature value of Input 1 with the one of Input 2-A. Indeed the rhythm characteristics of the blended melody (Blend-A) are almost identical to Input 1. The pitch class complexity is significantly increased while the percentage of small intervals is retained relatively low (0.33 while in Input 1 it is zero). Similarly, if we want to increase the percentage of small intervals to the extreme value of 1, we can blend the Chinese melody (Input 1) with the German melody labeled as Input 2-B. The blended melody (Blend-B) retains the rhythmic structure as well as other characteristics (including pitch class complexity) and adopts

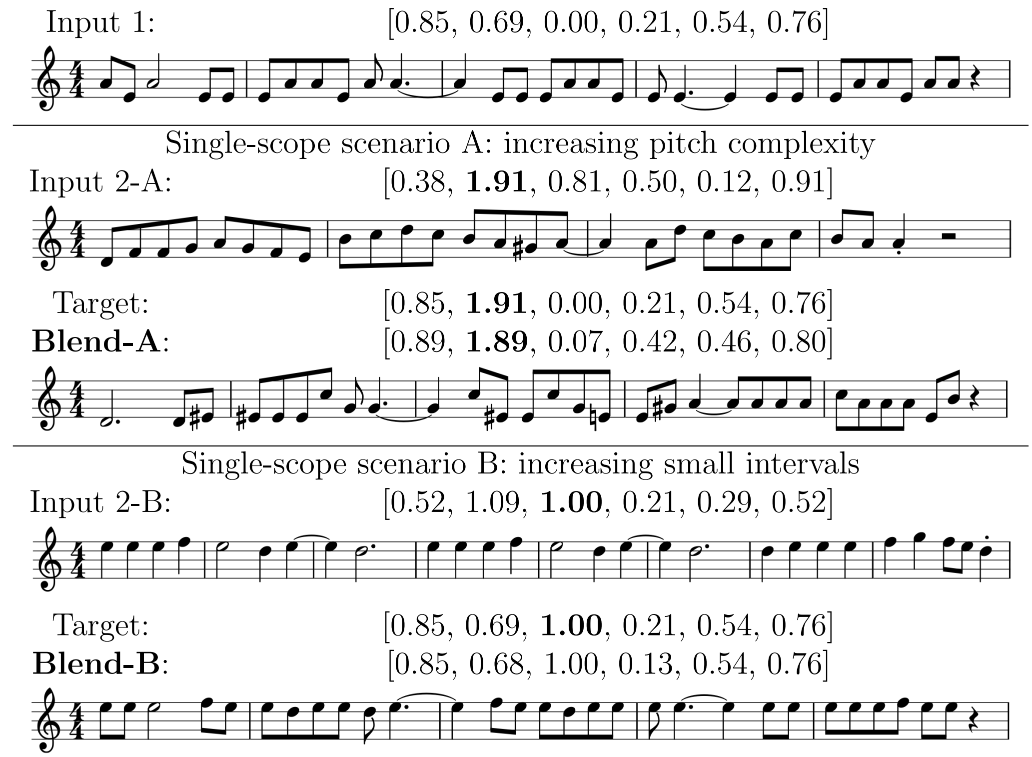


Figure 5: Two examples of "single scope" blends using as the first input (Input 1) a Chinese (`han`) melody with low pitch class complexity and low percentage of small intervals. In the first example, the other input is a German melody (`altdeut1`) with higher pitch class complexity (Input 2-A) and in the second (Input 2-B) is a German melody (`altdeut1`) with high percentage of small intervals. In both cases the blended melodies retain the rhythm features of the Han melody while adjusting to increased pitch class complexity and percentage of small intervals.

the extreme value for the feature of small intervals.

4.2. The Role of Feature Salience: Exemplar Objects, Double-Scope Blends and Listener Background

The salience value of a feature, according to the employed statistical approach, depends on the available background knowledge. Different interpretations might be given regarding which features are more salient in a specific object. In the shark-zebra example, if an isolated tribe of people never had witnessed another grey animal but the shark, then the grey zebra would be a meaningful blend for them, since it would encompass the grey colour which would be extremely "salient" for the shark category. The qualitative characteristics of the proposed statistical definitions of salience are examined on melodies through simple scenarios that involve the Eastern (with background the `han` and `natmin` styles) and Western (`altdeut1` and `zuccal`) "virtual" listeners. The Eastern and Western listeners are assumed to be exposed to a "new" style of the other culture: the Eastern is exposed to the `altdeut1` style and the Western to `han` style melodies. After being exposed to the new styles, listeners are assumed to adjust their understanding about what feature values are unique for each style. Figure 6 shows the concentration value uniqueness as computed in Equation 3.

For the Eastern listener (top row of figures) the most unique feature of the new style (`altdeut1`) is the PCP complexity (highest bar in the right-most figure). This means that the concentration value, as computed using Equation 1 on the respective datasets, in the `altdeut1` set (1.87) is distinctively higher than the respective concentration values in the Eastern sets (1.57 and 1.58 in the `han` and `natmin` sets respectively). It should be noted that the Shannon Information Entropy of a discrete uniform distribution with 7 out of 12 possibilities is 1.95 and with 5 out of 12 is 1.61, which is a good indication about the fact that the Eastern styles incorporate pentatonic scales and Western diatonic; therefore Eastern listeners familiar with pentatonic scales would find the diatonic nature of Western melodies unique. The high uniqueness value for this feature constitutes German melodies categorically salient regarding their PCP complexity and melodies with the PCP complexity feature closer to 1.87 (concentration value) are highly salient for the German idiom to the ears of an Eastern listener. On the other hand, the Western listeners in this hypothetic example find most aspects of the `han` melodies unique (except from the note range and note repetition) and should therefore find most aspects of the `han` melodies salient.

Before examining how salience values affect blending in the presented framework, we note that the concept of salience can be employed for identifying the *"exemplar"* object in a category. The "exemplar" object is the one that gathers the most typical characteristics of the category it belongs to and, according to psychology theories, "when an unfamiliar stimulus is encountered, its similarity is computed to the memory representation of every previously seen exemplar from each potentially relevant category" [68]. An alternative to the exemplar categorisation model is the "prototypical" model, where objects are categorised according to whether their features are close enough to some "prototypical" feature values that do not necessarily describe

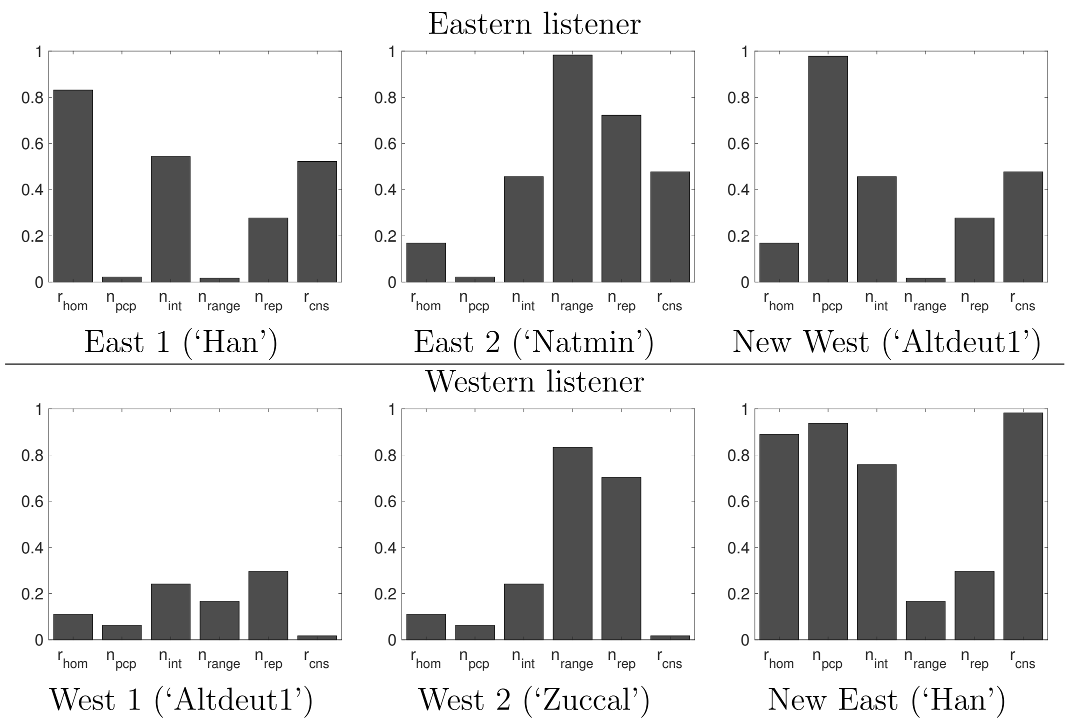


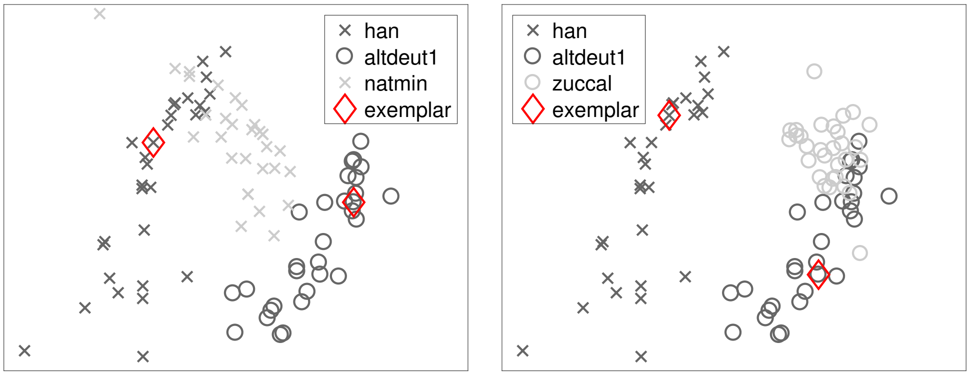
Figure 6: Salience uniqueness values for a new style presented to the "virtual" Eastern (top) and Western (bottom) listener. Eastern listeners, assumed "trained" two sets of Eastern melodies are exposed to one set of Western melodies of pieces and vice versa.

a concrete object; humans use either or both models in different classification tasks [69]. The exemplar object of a category can be assumed as the one that encompasses the highest sum of categorical salience in its category. Based on the definition of concentration value uniqueness ($U(C_{c_I}(t))$) in Equation 2, the "exemplar" object could be the one that encompasses features with high proximity to the respective unique concentration values of the category[3].

Figure 7 shows the placement of the melodies computed as exemplars for the `han` and `altdeut1` styles according to the background of the assumed Eastern (left) and Western (right) listeners. This illustration is produced by projecting the six melodic features on the 2-dimensional plane of maximum variance produced by Principle Component Analysis (PCA) on the main sets of melodies; the explained variance accounts for the 85.9% of variance in these two sets. Even though this graph does not show precise information, it illustrates how the background knowledge influences the perception of what an exemplar is. For example, for the Eastern listener (left graph in Figure 7), the exemplar in the `altdeut1` style is a melody placed higher on the y-axis in comparison to the Western listener (right). The location of the `altdeut1` exemplar for the Eastern listener is "occupied" by melodies in the `zuccal` style in the background of the Western listener (light-grey circles), therefore, for the Western listener the exemplar melody is placed further down – further away from the `zuccal` area.

According to Algorithm 1, using the computed salience values for the input melody features generates target-feature blends that incorporate a balanced combination of the most salient features from both blends. Figure 8 shows the blended melodies generated by using the aforementioned algorithm on Input 1 and Input 2-A in Figure 5. It is reminded that in Figure 5 only the increased pitch class profile complexity feature was taken from the German melody and used to generate a Chinese melody with increased pitch class profile complexity. In the current example of Figure 8, the salience perceived by the Eastern and Western listeners are used to generate the "best" blend accordingly. As shown in the "Target" feature vector, the Eastern listener is primarily attracted by the increased pitch class complexity of the Ger-

[3]In fact, summing the categorical salience values in Equation 2 for all features (for $t \in \{1, 2, \cdots, 6\}$ in the melodies example) produces the inner product between the feature uniqueness vector and the respective feature-to-concentration value proximities for an object; higher values of this inner product means better "alignment" between uniqueness of features and proximities to the respective concentration values.



Eastern listener Western listener

Figure 7: Exemplar melodies in the main styles for the Eastern (left) and Western (right) listeners, based on the a 2-dimensional PCA. The melodies in the background knowledge of the Eastern and the listeners are shown with a lighter grey shade.

man melody (value 1.91 of the second feature) and secondarily, based on the feature balancing that Algorithm 1 attempts, by the percentage of note repetitions (fifth feature – 0.12) and note range in the German melody (fourth feature – 0.50). Similarly, the Western listener finds salient for the Chinese melody the decreased pitch class profile (0.38 in the "Target" features), the non-existence of small intervals (third feature) and the percentage of constant intervals (0.76 in the sixth feature).

In both aspects of double-scope blends (for the Eastern and the Western listeners) in Figure 8, the underlying genetic algorithm materialised the blends ("Target" features and accompanied average Markov matrix of pitch transitions) into melodies that to some extent encompass the desired blended features (shown in the "Blend" arrays for the respective listener). A notable deviation concerns the blend generated for the Western listener, where the "target" 0.38 value in the rhythm inhomogeneity percentage (first feature) was not achieved and the melody that was actually generated had a value of 0.74 for this feature, which is closer to the feature value of the Chinese input. This could have happened in order to approach the desired value (0.76) for the conflicting feature of constant rhythm (feature 6)[4]. Even though this

[4]The problem of finding a melody that optimally satisfies all (potentially conflicting) criteria is a multi-objective optimisation problem that can be more efficiently addressed

	Input 1	[0.85, 0.69, 0.00, 0.21, 0.54, 0.76]
	Input 2	[0.38, 1.91, 0.81, 0.50, 0.12, 0.91]
Eastern	Target:	[0.85, 1.91, 0.00, 0.50, 0.12, 0.76]
	Blend:	[0.83, 1.92, 0.13, 0.50, 0.12, 0.68]

Western	Target:	[0.38, 0.69, 0.00, 0.50, 0.12, 0.76]
	Blend:	[0.74, 0.76, 0.05, 0.54, 0.21, 0.78]

Figure 8: Double-scope blending between Input 1 and Input 2-A in Figure 5 for the "virtual" Eastern and Western listeners.

is a strictly case-dependent situation, it should be noted that the targeted features generated in double-scope blends (through Algorithm 1) will not necessarily be satisfiable by a melody since features can be directly conflicting or even, in some cases, the genetic algorithm might fail to capture the targeted features.

4.3. Blending-based recommendation

Music recommendation has been extensively studied during the last years and it is employed in services offered by big companies for recommending new content (specifically music) to wide audiences every day. Purpose of music recommendation is to recommend to users new musical pieces they have never heard before and they may like. The approaches to music recommendation that have been studied can be divided in two broad categories: *collaborative filtering* [70] and *content-based* recommendation [71]. Collaborative filtering is based solely on user preferences, aiming to implicitly group users based on their preferences, regardless of content, and recommend new material according to the estimated group of each user: users placed closer together in the space of preferences are assumed to have similar preferences. Content-based

using pertinent techniques (e.g. employing the notion of the Pareto front); in this paper we intent to focus on the general ideas presented and detailing such techniques extend beyond the scope of this paper.

recommendation involves recommending new material to a user regardless of what other users prefer, based on the assumption that the new material should incorporate similar features to the material that this user prefers [72]. Content-based music recommendation employs distance measures between features of music pieces and proposes new music that belongs to the same "cluster of preference" of a user.

Feature blending potentially offers new possibilities for content-based music recommendation. Hitherto proposed approaches aim to recommend new music based on which unknown pieces are clustered together with highly-rated pieces of the user. This task shares the basic principles with genre or style classification [73], since both cluster pieces according to their content. However, in many occasions users prefer more than one styles which potentially incorporate completely different features; current content-based approaches are able to propose new music within the cluster of each user-preferred style but are unable to allow exploration to new styles. According to the hypotheses made about the computation of salience, this value for music features can be computed for listeners based on the music pieces and genres they prefer. The most salient features in user-preferred pieces reflect the most common and special characteristics of the categories that the user prefers. Therefore, by blending the most salient features of preferred pieces, new pieces can be retrieved and recommended that are not necessarily included in the styles known to the user.

If we suppose that an Eastern and a Western listener have rated high the Chinese and German melodies used as Input 1 and Input 2-A in Figure 5, then, based on the computed salience values and by applying Algorithm 1, we would get the blended target blended features demonstrated for each listener in Table 1 (which are the same blends used as target features in the examples shown in Figure 8). For the recommendation example, the most basic form of content-based recommendation is used: recommendations are new pieces with features close to the target (blended) features in terms of the Euclidean distance. The best five results (five new pieces closest to the target features) are returned for each listener, as shown in Table 1. Even though the Essen dataset is biased towards Western and Chinese melodies, it is obvious that this simple recommendation approach returns pieces in new styles to the listener. The Eastern listener receives recommendations for Austrian and `zuccal`-style melodies, while the Western listener gets `natmin`-style melodies.

Table 1: Recommendations based on blended features for the Eastern and Western listener.

	Input 1	[0.85, 0.69, 0.00, 0.21, 0.54, 0.76]
	Input 2	[0.38, 1.91, 0.81, 0.50, 0.12, 0.91]
Eastern	Target:	[0.85, 1.91, 0.00, 0.50, 0.12, 0.76]
	Recommendations	
	oesterrh/oestr039	[0.82, 1.86, 0.08, 0.58, 0.07, 0.46]
	han/han0719	[0.89, 1.71, 0.23, 0.58, 0.02, 0.20]
	zuccal/deut5021	[0.51, 1.84, 0.08, 0.42, 0.17, 0.75]
	oesterrh/oestr058	[0.67, 1.83, 0.22, 0.71, 0.31, 0.68]
	oesterrh/oestr041	[0.75, 1.66, 0.12, 0.71, 0.26, 0.32]
Western	Target:	[0.38, 0.69, 0.00, 0.50, 0.12, 0.76]
	Recommendations	
	han/han0791	[0.43, 0.97, 0.00, 0.50, 0.34, 0.48]
	natmin/natmn010	[0.46, 1.00, 0.00, 0.67, 0.18, 0.35]
	zuccal/deut4637	[0.51, 1.07, 0.00, 0.50, 0.54, 0.64]
	han/han0529	[0.42, 1.07, 0.17, 0.50, 0.33, 0.73]
	natmin/natmn204	[0.35, 1.09, 0.00, 0.38, 0.20, 0.76]

5. Conclusions

In this paper a new framework for generative Conceptual Blending (CB) has been presented that allows blending quantitative high-level features along with low-level information employing feature salience values for determining which features of the inputs should be included to the blend. Current approaches for generative CB act only on a low level of information, combining basic elements of the inputs and disregarding high-level information that captures meaning; we have referred to this problem as the *representation problem*. Additionally, in current approaches the identification of which features are important for each input is either performed ad-hoc during the definition of the input spaces, or is not considered at all, leading to the generation of many blends that need to be filtered at subsequent steps. This is problematic since it raises scalability issues: either all objects need to get hand-crafted annotations regarding the importance of their features, or overwhelmingly many blends will be generated; we have referred to this problem as the *salience problem*. The aforementioned problems have been addressed in the paper at hand by developing a simple methodology that blends high-level features of objects and employs Genetic Algorithms (GA) to combine low level information, leading to the construction of blends that adhere to the blended high-level features. Each feature is accompanied by a salience value that is computed based on the statistical layout of feature values in a dataset that represents the background of the listener. The proposed framework relies on (a) the availability of data in different categories; (b) a basic low-level representation of objects that GA can manipulate; (c) and the definition of some meaningful features that quantify high-level aspects of objects.

Three test-cases have been presented based on melody blending, with melodies derived from the Essen Corpus: (a) generation of single-scope blends, where a single characteristic from a melody is imported to another (two examples: increase in pitch class profile complexity and increase in small intervals); (b) examination of the role of the proposed statistical approach to salience in identifying "exemplar" objects in a category and in generating double-scope blends; and (c) recommendation of new music based on blended features from preferred music in different styles. The latter two cases present results based on assumptions about the background of two "artificial" listeners, an Eastern and a Western, who are assumed to have acquaintance only with sets of Chinese and German melodies respectively. The computed feature salience values for the two listeners is affected by their background according to the proposed model, while the presented examples verify, on an intuitive level, that the proposed framework makes sense.

Intuitive insights have been presented that support the plausibility of the proposed methodology; empirical experiments, however, will be necessary to reveal whether human listers indeed perceive blending and salience as the system predicts. Even though it is outside the scope of this paper to discuss evaluation in detail, we briefly refer to previous work that on empirical evaluation of blending methodologies that might be pertinent. In [74] listers rated how dissimilar pairs of cadences were, where cadences where either the two inputs or their blending products, leading to a space of perceived distance among all cadences. In a similar manner, the perceived distances of a set of blended melodies could be estimated, leading to conclusions regarding which features play a more important role in defining melodic distance; such tests could allow estimations of feature salience for specific groups of listeners when rating melodic distances. Another methodology for empirical testing could be similar to [75], where harmonisations (blends or non-blends) were given as stimuli and listeners had to rate whether they sounded like tonal or jazz. Such tests could reveal whether some feature values are decisive for classifying melodies as Chinese or German, leading to assumptions about the importance of features. It has to be noted that listers with diverse, possibly non-Western, backgrounds should be included in such test, which is all the more difficult with the current global abundance of Western-related music.

Future projection of this work may be relevant for a recent paradigm of methodologies for human-computer communication: argumentation systems. Such systems engage in "dialogues" with the user, exchanging arguments toward creating a satisfactory output. Argumentation systems have been studied in the context of conceptual blending in [13], but the level of communication was deteriorated by the fact that user choices concerned only low-level properties. Enabling high-level concepts and relations between them will allow more intuitive queries by the user and more informative responses by the system, leading to more meaningful dialogues. The methodology proposed in the paper at hand allows the incorporation of such high-level, quantitatively-expressed concepts in the framework of generative Conceptual Blending, allowing the aforementioned improvement and expansion of the current framework.

Acknowledgments

This work has been supported by the Greek National Scholarships Foundation (Grant Agreement Number: 2016-050-0503-8138).

References

[1] J. W. Charnley, A. Pease, S. Colton, On the notion of framing in computational creativity., in: Proceedings of the International Conference on Computational Creativity ICCC, 2012, pp. 77–81.

[2] M. A. Boden, Computer Models of Creativity, AI Magazine 30 (3) (2009) 23.

[3] M. A. Boden, The creative mind: Myths and mechanisms, Psychology Press, 2004.

[4] G. Fauconnier, M. Turner, The Way We Think: Conceptual Blending And The Mind's Hidden Complexities, reprint edition Edition, Basic Books, New York, 2003.

[5] C. Tsougras, D. Stefanou, Conceptual Blending and Meaning Construction: A Structural/hermeneutical Analysis of the 'Old Castle' from Musorgsky's 'Pictures at an Exhibition', in: Proceedings of the 9th Triennial Conference of the European Society for the Cognitive Science of Music (ESCOM), 2015.

[6] L. M. Zbikowski, Metaphor and music, The Cambridge handbook of metaphor and thought (2008) 502–524.

[7] L. M. Zbikowski, Conceptualizing Music: Cognitive Structure, Theory, and Analysis, Oxford University Press, 2002.

[8] D. Gómez-Ramírez, Conceptual blending as a creative meta-generator of mathematical concepts: prime ideals and dedekind domains as a blend, in: Proceedings of the 4th international workshop on computational creativity, concept invention, and general intelligence (C3GI), Publications of the Institute of Cognitive Science, Vol. 2, 2015, pp. 1–11.

[9] F. Bou, M. Schorlemmer, J. Corneli, D. Gómez-Ramírez, E. Maclean, A. Smaill, A. Pease, The role of blending in mathematical invention, in: Proceedings of the International Conference on Computational Creativity (ICCC), 2015, pp. 55–62.

[10] M. Eppe, R. Confalonier, E. Maclean, M. Kaliakatsos-Papakostas, E. Cambouropoulos, M. Schorlemmer, M. Codescu, K.-U. Kühnberger, Computational invention of cadences and chord progressions by conceptual chord-blending., in: International Joint Conference on Artificial Intelligence (IJCAI) 2015, 2015.

[11] A. Zacharakis, M. Kaliakatsos-Papakostas, E. Cambouropoulos, Conceptual blending in music cadences: A formal model and subjective evaluation, in: Proceedings of the 4th International Conference on Music Information Retrieval (ISMIR 2015), Malaga, Spain, 2015.

[12] K. Hirata, S. Tojo, M. Hamanaka, Melodic morphing algorithm in formalism, Mathematics and Computation in Music (2011) 338–341.

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

[14] M. Kaliakatsos-Papakostas, D. Makris, A. Zacharakis, C. Tsougras, E. Cambouropoulos, Learning and blending harmonies in the context of a melodic harmonisation assistant, in: IFIP International Conference on Artificial Intelligence Applications and Innovations, Springer, 2016, pp. 520–527.

[15] D. B. Huron, Sweet anticipation: Music and the psychology of expectation, MIT press, 2006.

[16] C. L. Krumhansl, Cognitive foundations of musical pitch, Oxford University Press, 2001.

[17] D. Temperley, The cognition of basic musical structures, MIT press, 2004.

[18] S. M. Demorest, S. J. Morrison, Quantifying culture: The cultural distance hypothesis of melodic expectancy, The Oxford Handbook of Cultural Neuroscience (2016) 183.

[19] E. J. Kessler, C. Hansen, R. N. Shepard, Tonal schemata in the perception of music in bali and in the west, Music Perception: An Interdisciplinary Journal 2 (2) (1984) 131–165.

[20] C. L. Krumhansl, Tonality induction: A statistical approach applied cross-culturally, Music Perception: An Interdisciplinary Journal 17 (4) (2000) 461–479.

[21] M. T. Pearce, Statistical learning and probabilistic prediction in music cognition: mechanisms of stylistic enculturation, Annals of the New York Academy of Sciences.

[22] E. Cambouropoulos, A. Smaill, Similarity and categorisation inextricably bound together: The unscramble machine learning algorithm, in: Proceedings of the Interdisciplinary Workshop on Similarity and Categorisation, University of Edinburgh, 1997.

[23] E. Cambouropoulos, Towards a general computational theory of musical structure.

[24] E. Cambouropoulos, G. Widmer, Melodic clustering: Motivic analysis of schumanns träumerei, Proceedings of JIM.

[25] G. Widmer, Getting closer to the essence of music: The con espressione manifesto, ACM Transactions on Intelligent Systems and Technology (TIST) 8 (2) (2016) 19.

[26] M. A. Kaliakatsos-Papakostas, A. Floros, M. N. Vrahatis, Intelligent music composition, in: X.-S. Yang, Z. Cui, R. Xiao, A. H. Gandomi, M. Karamanoglu (Eds.), Swarm Intelligence and Bio-Inspired Computation, Elsevier, Oxford, 2013, pp. 239 – 256.

[27] M. Kaliakatsos-Papakostas, A. Floros, M. N. Vrahatis, Evodrummer: deriving rhythmic patterns through interactive genetic algorithms, in: International Conference on Evolutionary and Biologically Inspired Music and Art, Springer, 2013, pp. 25–36.

[28] C. O. Nuanáin, P. Herrera, S. Jorda, Target-based rhythmic pattern generation and variation with genetic algorithms, in: Sound and Music Computing Conference, 2015.

[29] R. Vogl, M. Leimeister, C. Ó. Nuanáin, S. Jordà, M. Hlatky, P. Knees, An intelligent interface for drum pattern variation and comparative evaluation of algorithms, Journal of the Audio Engineering Society 64 (7/8) (2016) 503–513.

[30] A. Eigenfeldt, P. Pasquier, Evolving structures for electronic dance music, in: Proceedings of the 15th annual conference on Genetic and evolutionary computation, ACM, 2013, pp. 319–326.

[31] A. J. Milne, R. T. Dean, Computational creation and morphing of multilevel rhythms by control of evenness, Computer Music Journal.

[32] M. Kaliakatsos-Papakostas, Generating drum rhythms through data-driven conceptual blending of features and genetic algorithms, in: Evolutionary and Biologically Inspired Music, Sound, Art and Design, Lecture Notes in Computer Science, Springer Berlin Heidelberg, 2018, pp. 145–160.

[33] A. Piata, Genre "out of the box": A conceptual integration analysis of poetic discourse, Genre in Language, Discourse and Cognition 33 (2016) 225.

[34] C. P. Cánovas, M. Valverde, Interactions with the beloved in greek literature: Conceptual blending and levels of representation, Trends in Classics 9 (1) (2017) 85–112.

[35] A. Joy, J. F. Sherry, J. Deschenes, Conceptual blending in advertising, Journal of business research 62 (1) (2009) 39–49.

[36] M. Turner, G. Fauconnier, A mechanism of creativity, Poetics today 20 (3) (1999) 397–418.

[37] H. Ling, A study on news headline translation from the cognitive perspective-based on conceptual blending theory, English Language and Literature Studies 5 (4) (2015) 82.

[38] L. M. Zbikowski, Conceptual models and cross-domain mapping: New perspectives on theories of music and hierarchy, Journal of music theory 41 (2) (1997) 193–225.

[39] M. Antovic, Musical metaphor revisited: Primitives, universals and conceptual blending, Universals and Conceptual Blending.

[40] M. Spitzer, Metaphor and musical thought, University of Chicago Press, 2004.

[41] M. L. Gick, K. J. Holyoak, Schema induction and analogical transfer, Cognitive psychology 15 (1) (1983) 1–38.

[42] M. M. Hedblom, O. Kutz, F. Neuhaus, Image schemas in computational conceptual blending, Cognitive Systems Research 39 (2016) 42–57.

[43] M. Leman, Music and schema theory: Cognitive foundations of systematic musicology, Vol. 31, Springer Science & Business Media, 2012.

[44] I. Shmulevich, D.-J. Povel, Measures of temporal pattern complexity, Journal of New Music Research 29 (1) (2000) 61–69.

[45] G. Madison, G. Sioros, What musicians do to induce the sensation of groove in simple and complex melodies, and how listeners perceive it, Frontiers in psychology 5.

[46] G. Sioros, M. Miron, M. Davies, F. Gouyon, G. Madison, Syncopation creates the sensation of groove in synthesized music examples, Frontiers in psychology 5.

[47] M. Kaliakatsos-Papakostas, M. Queiroz, C. Tsougras, E. Cambouropoulos, Conceptual blending of harmonic spaces for creative melodic harmonisation, Journal of New Music Research 43 (3) (2017) 1–25.

[48] A. Koestler, The act of creation, London Hutchinson, 1964.

[49] V. Goel, Creative brains: designing in the real world, Frontiers in human neuroscience 8 (2014) 241.

[50] J. McCarthy, P. J. Hayes, Some philosophical problems from the standpoint of artificial intelligence, Readings in artificial intelligence (1969) 431–450.

[51] J. Goguen, Mathematical Models of Cognitive Space and Time, in: D. Andler, Y. Ogawa, M. Okada, S. Watanabe (Eds.), Reasoning and Cognition, Vol. 2 of Interdisciplinary Conference Series on Reasoning Studies, Keio University Press, 2006.

[52] J. Goguen, D. F. Harrell, Style: A Computational and Conceptual Blending-Based Approach, in: S. Argamon, S. Dubnov (Eds.), The Structure of Style: Algorithmic Approaches to Understanding Manner and Meaning, Springer, Berlin, 2010, pp. 147–170.

[53] F. Neuhaus, O. Kutz, M. Codescu, T. Mossakowski, Fabricating Monsters is Hard. Towards the Automation of Conceptual Blending, in: Proceedings of the Workshop Computational Creativity, Concept Invention, and General Intelligence (C3GI 2014), 2014.

[54] H. Schaffrath, Beyond midi, MIT Press, Cambridge, MA, USA, 1997, Ch. The Essen Associative Code: A Code for Folksong Analysis, pp. 343–361.

[55] M. Towsey, A. Brown, S. Wright, J. Diederich, Towards melodic extension using genetic algorithms, Educational Technology & Society 4 (2) (2001) 54–65.

[56] M. A. Kaliakatsos-Papakostas, A. Floros, M. N. Vrahatis, N. Kanellopoulos, Genetic evolution of L and FL–systems for the production of rhythmic sequences, in: In Proceedings of the 2nd Workshop in Evolutionary Music (GECCO 2012), Philadelphia, USA, 2012, pp. 461–468.

[57] F. Gómez, A. Melvin, D. Rappaport, G. Toussaint, Mathematical measures of syncopation, in: BRIDGES: Mathematical Connections in Art, Music and Science, 2005, pp. 73–84.

[58] H. C. Longuet-Higgins, C. S. Lee, The rhythmic interpretation of monophonic music, Music Perception: An Interdisciplinary Journal 1 (4) (1984) 424–441.

[59] G. Sioros, C. Guedes, Complexity driven recombination of midi loops., in: Proceedings of the 12th International Society for Music Information Retrieval Conference (ISMIR), University of Miami, Miami, USA, 2011, pp. 381–386.

[60] C. E. Shannon, Prediction and entropy of printed english, Bell Labs Technical Journal 30 (1) (1951) 50–64.

[61] M. Schorlemmer, A. Smaill, K.-U. Kühnberger, O. Kutz, S. Colton, E. Cambouropoulos, A. Pease, Coinvent: Towards a computational concept invention theory, in: 5th International Conference on Computational Creativity (ICCC) 2014, 2014.

[62] M. Eppe, E. Maclean, R. Confalonieri, O. Kutz, M. Schorlemmer, E. Plaza, ASP, Amalgamation, and the Conceptual Blending Workflow, in: Logic Programming and Nonmonotonic Reasoning, Vol. 9345 of Lecture Notes in Computer Science, Springer International Publishing, 2015, pp. 309–316.

[63] R. Confalonier, M. Schorlemmer, E. Plaza, M. Eppe, O. Kutz, R. Peñaloza, Upward Refinement for Conceptual Blending in Description Logic – An ASP-based Approach and Case Study in $\mathcal{EL}^{++}$, in: International Workshop on Ontologies and Logic Programming for Query Answering, International Joint Conference on Artificial Intelligence (IJCAI) 2015, 2015.

[64] H. Ait-Kaci, A. Podelski, Towards a meaning of life, The Journal of Logic Programming 16 (3) (1993) 195–234.

[65] G. Wiggins, G. Papadopoulos, A genetic algorithm for the generation of jazz melodies, in: Proceedings of the Finnish Conference on Artificial Intelligence (STeP'98), Jyväskylä, Finnland, 1998.

[66] H. Schaffrath, D. Huron, The essen folksong collection in kern format.[computer database], Menlo Park, CA: Center for Computer Assisted Research in the Humanities.

[67] B. van der Weij, M. T. Pearce, H. Honing, A probabilistic model of meter perception: Simulating enculturation, Frontiers in psychology 8 (2017) 824.

[68] F. G. Ashby, W. T. Maddox, Human category learning, Annu. Rev. Psychol. 56 (2005) 149–178.

[69] M. L. Mack, A. R. Preston, B. C. Love, Decoding the brains algorithm for categorization from its neural implementation, Current Biology 23 (20) (2013) 2023–2027.

[70] X. Su, T. M. Khoshgoftaar, A survey of collaborative filtering techniques, Advances in artificial intelligence 2009.

[71] Y. Song, S. Dixon, M. Pearce, A survey of music recommendation systems and future perspectives, in: 9th International Symposium on Computer Music Modeling and Retrieval, Vol. 4, 2012.

[72] J. Kaitila, A content-based music recommender system, Master's thesis, University of Tampere (2017).

[73] G. Tzanetakis, P. Cook, Musical genre classification of audio signals, IEEE Transactions on speech and audio processing 10 (5) (2002) 293–302.

[74] A. Zacharakis, M. Kaliakatsos-Papakostas, C. Tsougras, E. Cambouropoulos, Creating musical cadences via conceptual blending: empirical evaluation and enhancement of a formal model, Music Perception: An Interdisciplinary Journal 35 (2) (2017) 211–234.

[75] A. Zacharakis, M. Kaliakatsos-Papakostas, C. Tsougras, E. Cambouropoulos, Musical blending and creativity: An empirical evaluation of the CHAMELEON melodic harmonisation assistant, Musicae Scientiae 22 (1) (2018) 119–144.

Algorithm 1 Computation of the best blended set of features of two input spaces.

Require: arrays of the two input features, $\vec{m_1}$ and $\vec{m_2}$, the arrays of their corresponding saliences, $s_{\vec{m_1}}$ and $s_{\vec{m_2}}$ and an array of the generic space features, $\vec{g}$.
Ensure: array, $\vec{b}$, including the optimal set of blended features.

$\vec{b} \leftarrow \vec{\emptyset}$ {% initialise best blend as an empty array}
{% get the sorted indexes of the saliences in both inputs}
$\vec{i}^{(1)} \leftarrow$ `getSortedIndexes(`$s_{\vec{r_1}}$`)`
$\vec{i}^{(2)} \leftarrow$ `getSortedIndexes(`$s_{\vec{r_2}}$`)`
{% clear out the indexes that correspond to features of the generic space}
for $j \in \{1, 2, \ldots, 32\}$ **do**
    if $g_j \neq \emptyset$ **then**
        $b_j \leftarrow g_j$
        $\vec{i}^{(1)} \leftarrow$ `removeElement(`j`,`$\vec{i}^{(1)}$`)`
        $\vec{i}^{(2)} \leftarrow$ `removeElement(`j`,`$\vec{i}^{(2)}$`)`
    end if
end for
{% select input with the highest salience in any feature}
if $s_{\vec{r_1}}(i_1^{(1)}) > s_{\vec{r_2}}(i_1^{(2)})$ **then**
    $c \leftarrow 1$
else
    $c \leftarrow 2$
end if
{% fill the non-generic space features by picking up the most salient ones from each input interchangeably}
while `isNotEmptyArray(`$\vec{i}^{(1)}$`)` **do**
    if $c == 1$ **then**
        $b_{i_1^{(1)}} = f_{i_1^{(1)}}$ {% get most salient feature available from input space 1 and remove its index from both arrays of indexes}
        $\vec{i^{(1)}} \leftarrow$ `removeElement(`$i_1^{(1)}$`, `$\vec{i^{(1)}}$`)`
        $\vec{i^{(2)}} \leftarrow$ `removeElement(`$i_1^{(1)}$`, `$\vec{i^{(2)}}$`)`
        $c \leftarrow 2$
    else
        $b_{i_2^{(2)}} = f_{i_2^{(2)}}$ {% get most salient feature available from input space 2 and remove its index from both arrays of indexes}
        $\vec{i^{(1)}} \leftarrow$ `removeElement(`$i_1^{(2)}$`, `$\vec{i^{(1)}}$`)`
        $\vec{i^{(2)}} \leftarrow$ `removeElement(`$i_1^{(2)}$`, `$\vec{i^{(2)}}$`)`
        $c \leftarrow 1$
    end if
end while