International Journal of Art, Culture and Design Technologies
Volume 8 • Issue 2 • July-December 2019

MixMash:
An Assistive Tool for Music Mashup Creation from Large Music Collections

Catarina Maçãs, CISUC - Department of Informatics Engineering University of Coimbra, Coimbra, Portugal

Ana Rodrigues, CISUC - Department of Informatics Engineering University of Coimbra, Coimbra, Portugal

Gilberto Bernardes, INESC TEC & University of Porto, Faculty of Engineering, Porto, Portugal

https://orcid.org/0000-0003-3884-2687

Penousal Machado, CISUC - Department of Informatics Engineering, University of Coimbra, Coimbra, Portugal

ABSTRACT

This article presents MixMash, an interactive tool which streamlines the process of music mashup creation by assisting users in the process of finding compatible music from a large collection of audio tracks. It extends the harmonic mixing method by Bernardes, Davies and Guedes with novel degrees of harmonic, rhythmic, spectral, and timbral similarity metrics. Furthermore, it revises and improves some interface design limitations identified in the former model software implementation. A new user interface design based on cross-modal associations between musical content analysis and information visualisation is presented. In this graphic model, all tracks are represented as nodes where distances and edge connections display their harmonic compatibility as a result of a force-directed graph. Besides, a visual language is defined to enhance the tool's usability and foster creative endeavour in the search of meaningful music mashups.

KEYWORDS

Cross-Modal Associations, Dissonance, Force-Directed Layout, Harmonic Mixing, Information Visualisation, Music Mashup, Perceptual Relatedness, Rhythmic Density, Spectral Region, Timbral Similarity

INTRODUCTION

Mashup creation is a music composition practice strongly linked to various sub-genres of Electronic Dance Music (EDM) and the role of the DJ (Shiga, 2007). It entails the recombination of existing pre-recorded musical audio as a means of creative endeavour (Navas, 2014). This practice has been nurtured by the existing and growing media preservation mechanisms that allow users to access large collections of musical audio in digital format for their mixes (Vesna, 2007). However, the scalability of these growing audio collections also raises the issue of retrieving musical audio that matches particular criteria (Schedl, Gómez, & Urbano, 2014). In this context, both industry and academia have been devoting effort to develop tools for computational mashup creation, which streamline the time-consuming and complex search for compatible musical audio.

Early research on computational mashup creation, focused on rhythmic-only attributes, particularly those relevant to the temporal alignment of two or more musical audio tracks (Griffin, Kim & Turnbull, 2010). Recent research (Davies, Hamel, Yoshii, & Goto, 2014; Gebhardt, Davies,

DOI: 10.4018/IJACDT.2019070102

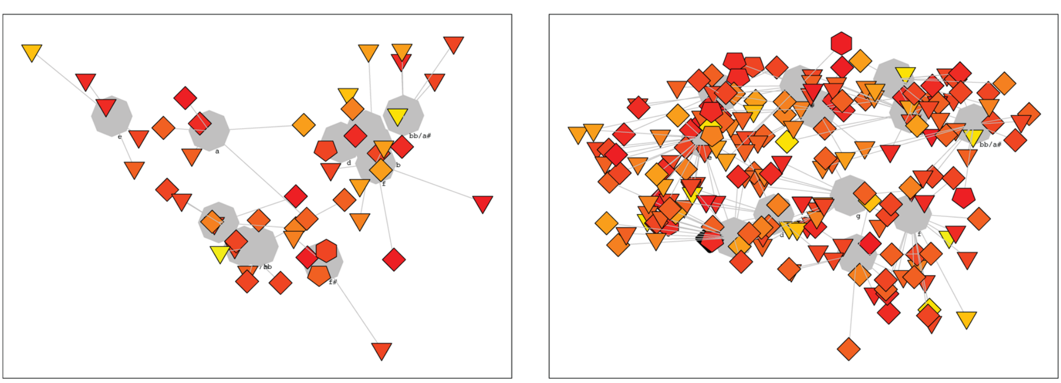
& Seebe, 2016; Bernardes, Davies, & Guedes, 2018) has expanded the range of musical attributes under consideration towards harmonic- and spectral-driven attributes. The former aims to identify the degree of harmonic compatibility in musical audio, commonly referred to as harmonic mixing. The latter aims to identify the spectral region occupied by a particular musical audio track across the frequency range (e.g., the concentration of energy in low, middle, and high frequency bands), which can then guide the spectral distribution of the mix.

The interface design of early software implementation models adopts a one-to-many mapping strategy between a user-defined track and a ranked list of compatible tracks to show the results to the user (Mixed in Key, n.d.; Native Instruments, n.d.; Davies et al., 2014). Recently, Bernardes et al. (2018) proposed an interface design which adopts a many-to-many mapping strategy, which offers a global view of the compatibility between all tracks in a music collection and promotes serendipitous navigation (Figure 1). It represents each audio track in a collection as a graphical element in a navigable 2-dimensional interface. Distances among these elements indicate harmonic compatibility and the additional graphic variables of these elements, such as colour and shape, indicate rhythmic and spectral information relevant to mashup creation. By exposing users to the compatibility between all tracks in a collection, this interface design aims to promote an overview of the relations between tracks.

In short, advances in computational mashup creation models, emphasize a gradual increase in the number of extracted data-driven attributes from musical audio and a global view of the audio collections through information visualization. This tendency acknowledges the subjective nature of the task and enhances the degree of personalization in the search for compatible audio in mashup creation. However, scalability of these audio collections now raises concerns at the usability level in many-to-many interface design (Bernardes, Davies, & Guedes, 2018; Maçãs et al., 2018). Figure 1 b) highlights the three main limitations: (i) the clutter resulting from the superposition of graphic elements; (ii) the static representation of the tracks, which does not promote a finer exploration of particular dense areas; and (iii) the reduced number of existing graphic attributes in the visual tracks representation to depict musical audio content-driven information. From these limitations it was possible to define the three main goals for the present work: (i) the prevention of overlapping graphic elements and subsequent simplification of the visualisation; (ii) the creation of a system capable of dynamically adapt to the user interactions; and (iii) the complete representation of the tracks' musical characteristics and their harmonic compatibility. To address the identified interface design limitations, it was adopted in Maçãs et al. (2018) a methodology based on the Three Cycle View of Design Science Research (Hevner, 2007). This methodology promoted the iterative implementation

Figure 1. Screenshot of the visualisation of the original MixMash interface representing (a) 50 musical audio tracks and (b) 200 musical audio tracks. Refer to Bernardes et al. (2018) for a detailed interpretation.

a) b)



of the visualisation model and the definition of the visual representations of the attributes of musical audio tracks.

The current paper extends Maçãs et al. (2018) in four aspects. First, it provides a detailed overview of computational mashup creation, from the early rhythmic-only alignments to the current multidimensional attribute spaces, which to the best of our knowledge has not been addressed elsewhere. Second, it expands the range of musical audio attributes under consideration, notably including timbre as a relevant dimension, following recent evidence on its significant impact on listening preference as listeners are able to reliably evoke changes in timbre towards their preferences (Dobrowohl, Milne, & Dean, 2019). Third, it details the signal processing underlying all metrics adopted to extract content-driven information from musical audio. Fourth, the force-directed algorithm, visual mappings and interactive interface are thoroughly described.

The remainder of the paper is organised as follows. The Background summarises the related work. MixMash: Compatibility Method describes the audio analysis methods that support a novel music visualisation system for assisting music mashup creation. The Methodology presents the development strategies for the visualisation model. MixMash: Visualisation Model presents a new approach for the visualisation of compatible musical audio tracks, a description of the established associations between graphic elements and musical audio attributes, and the interface design. The Conclusion states the conclusions of this work and its future directions.

BACKGROUND

The present section is divided into three parts, in which the authors present related work concerning the following topics: harmonic mixing methods, the visualisation methods applied to represent music, and the characterisation of force-directed graphs and their application in the visualisation of music collections.

Harmonic Mixing

There are four major harmonic mixing methods in the literature: key affinity, chroma vectors similarity, sensory dissonance minimisation, and hybrid (hierarchical) models. The initial three approaches focus on single harmonic attributes only, and the latter approach provides a hierarchical view over harmonic compatibility.

Key affinity is one of the earliest computational approaches to harmonic mixing. It has been proposed by industry (Mixed in Key, n.d., Native Instruments, n.d.) and is computed as distances in the Camelot Wheel or circle of fifths representation (Mixed in Key, n.d.). This approach enforces some degree of tonal stability and large-scale harmonic coherence of the mashup by privileging the use of the same diatonic key pitch set.

Chroma vectors similarity inspects the cosine distance between chroma vector representations of pitch-shifted versions of two given audio tracks as a measure of their compatibility (Davies et al., 2014; Lee, Lin, Yao, Li, & Wu, 2015). The similarity is typically computed at the time scale of beat durations, thus privileging small-scale alignments over large-scale harmonic structure between audio slices with highly similar pitch class content.

Sensory dissonance models follow the same local and small-scale alignment strategy as chroma vectors similarity between pitch-shifted versions of overlapping musical audio, yet adopt a more refined metric of compatibility, which aims to minimize their combined level of sensory dissonance (Gebhardt, Davies & Seebe, 2016); a motivation well-rooted in the Western musical tradition by favouring a less dissonant harmonic lexicon.

Recently, a hybrid hierarchical model for harmonic mixing has been proposed by Bernardes et al. (2018). It combines previous approaches for (small- and large-scale) harmonic compatibility in a single framework. Furthermore, it proposes a novel interface design approach, which offers a global view over the harmonic compatibility of an entire music collection (many-to-many), beyond the existing one-to-many relationships between a user-defined track and an audio collection (Bernardes et al., 2018).

Music Visualisation

Historically, numerous artists have created audiovisual associations, later referred to as graphic notation and visual music. Pioneering works by Kandinsky, Pfenninger, Cage, Fischinger, and Whitney explore combinations of visual principles—mainly colour and shape—to emphasise the audiovisual experience (McDonnell, 2007). In the present work, the visual translation of a collection of musical audio tracks is approached from two different points of view: music visualisation and the visual representation of a large collection of musical audio tracks. Regarding music visualisation, some authors have tried to solve this problem by focusing on the geometry of musical structure (Bergstrom, Karahalios, & Hart, 2007), while others have focused on a solution based on mappings between a specific set of musical characteristics and some visual characteristics (Snydal & Hearst, 2005; Wattenberg, 2002; Sapp, 2001). For example, Wattenberg (2002) uses arc diagrams to connect sequences containing the same pitch content, revealing the structure of musical compositions. On the other hand, Sapp (2001) presented a multi-timescale visualisation of the harmonic structure and the key relations in a musical composition. Another experimental approach has been presented by Rodrigues, Cardoso, and Machado (2016), where a visualisation model is created to provide a perceptually relevant experience for the user. Overall, the aforementioned visualisations do not allow great modularity of data, often binding the visual clarity of each element and limiting the comparison of other musical characteristics.

Although visualisations of large collections of musical audio collections have already been addressed by some authors (Grill & Flexer, 2012; Hamasaki & Goto, 2013; Rauber, Pampalk, Merkl, 2003; Schwarz and N. Schnell, 2009; Gulik, Vignoli, & Van de Wetering, 2004), it still is an area in need of great development. Grill and Flexer (2012), similarly to the work of Bernardes et al. (2018), developed a visualisation strategy capable of representing perceptual qualities from a large collection of sounds. Although their musical focus was in sound textures, they aimed at finding an intuitive and meaningful interface. For this purpose, they built an audiovisual language based on the cross-modal mechanisms of human perception. In this project, subjects were able to successfully associate sounds with the corresponding graphic representations. Hamasaki and Goto (2013), Gulik et al. (2004) and Schwartz and Schnell (2009) also proposed interactive visualisations of several layers of information; however, they do not represent the perceptual relevance of the musical characteristics at a visual level.

Force-Directed Graphs

The visualisation of graphs handles the representation of relational structures in data, aiding in the analysis, modelling, and interpretation of complex network systems (Meirelles, 2013). Graph visualisation is characterised by the existence of two main elements: (i) nodes, representing an entity (e.g., person, cell, machine); and (ii) edges, representing the relationship between nodes. Such models are often applied to large and complex datasets, which entails a set of problems related to performance and clutter. As the graphs grow in size, the required visual space and computational resources also increase (Herman, Melançon, & Marshall, 2000). To solve this, clustering techniques are applied by grouping similar nodes by their semantics and/or position on the graph. Through clustering, it is possible to reduce visual clutter and complexity, enhance clarity and performance, and create a simplified overview of the network's structure (Herman et al., 2000; Kimelman, Leban, Roth, & Zernik, 1994).

The positioning of nodes in space can be defined through a different set of graph layouts, such as arc diagrams, treemaps, circular, and force-directed layouts. In this paper, the latter is adopted. Force-directed graphs are based on a physical system that organises the network through forces of repulsion and attraction applied continuously to each node. This technique facilitates the visual interpretation of the inherent structure of the data, improving the analysis and comprehension of the relations within complex networks (Jacomy, Venturini, Heymann, & Bastian, 2014).

Force-directed graphs are used in a variety of areas, such as biology, medicine, literature, sociology (Enright and C. A. Ouzounis, 2001; Goh et al., 2007, Chen, 2006; Heer and Boyd, 2005; Gilbert, Simonetto, Zaidi, Jourdan, & Bourqui, 2011; Fu et al., 2007), but also in the visualisation of music collections (Vavrille, n.d.; Gibney, n.d.). In the work of Muelder, Provan, and Ma (2010), a force-directed layout was applied to visualise music libraries. Each piece of music represents a node, positioned depending on its similarities with other musical audio tracks (e.g., same artist). Songrium (Hamasaki & Goto, 2013), is a music browsing service that visualises relations among original songs and their derivative works. Each node represents a video-clip and each edge connects an original work to its derivatives. In addition, songs with similar moods get stronger forces, thus are positioned closer together.

The current work expands the state-of-the-art by applying a force-directed graph to the mashup creation process. To enable the user to analyse the graph and relate musical audio tracks from large collections, the harmonic compatibility metrics are applied to the forces of the graph nodes. This improves the visual separation of distinct musical audio tracks, which can have a positive impact on user creativity (Henry, Fekete, & McGuffin, 2007).

MIXMASH: COMPATIBILITY METHOD

MixMash is a software application which aims to assist users in finding musical mashups in the context of mashup creation. It builds on metrics and methods presented in Bernardes et al. (2018) and expands the state-of-the-art of harmonic mixing by providing a greater amount of relevant information to the process of music mashup. Its main novelty lies in a hierarchical harmonic mixing method, which includes metrics for both small- and large-scale structural levels, i.e., local (e.g., beats or phrases) and global (e.g., large sections or overall musical mashup) harmonic alignments between musical audio tracks, respectively. Moreover, this method considers three additional dimensions that can help users defining the compatibility of musical audio tracks and remaining compositional goals in terms of rhythmic (onset density), spectral (region) and timbral qualities.

To promote an intuitive search for compatible tracks in a music collection, a many-to-many mapping strategy was previously introduced in the interface design. This design opposes the ranked track list to a user-defined track, adopted in previous systems harmonic mixing software (Mixed in Key, n.d.; Native Instruments, n.d.), which is (i) inefficient computationally, as it recomputes intensive audio signal analysis every time a different audio track is selected as target, and (ii) limited in promoting creative endeavor and serendipity (Bernardes et al., 2018). A flexible many-to-many mapping strategy was allowed by the adoption of novel signal processing methods for small- and large-scale harmonic compatibility metrics in a confined spatial configuration. This method is at the basis of the MixMash visually-driven interface strategy. The signal processing methods used to compute these harmonic compatibility metrics, followed by the additional rhythmic, spectral, and timbral attributes are described next.

Small- and Large-Scale Harmonic Compatibility

MixMash adopts the perceptually-motivated Tonal Interval Space (Bernardes, Cocharro, Caetano, Guedes, & Davies, 2016) for representing the harmonic content of musical audio tracks. Each track exists as a 12-dimensional (12-D) Tonal Interval Vector (TIV), $T\left(k\right)$, whose locations represent unique harmonic configurations. An audio track TIV, $T\left(k\right)$, is computed as the weighted Discrete Fourier Transform (DFT) of an L_1 normalized chroma vector, $c\left(n\right)$, such that:

$$T\left(k\right) = w\left(k\right)\sum_{n=0}^{N-1}\overline{c}\left(n\right)e^{\frac{-j2\pi kn}{N}},\ k \in \mathbb{Z} \tag{1}$$

where $N = 12$ is the dimension of the chroma vector, each of which expresses the energy of the 12 pitch classes, and $w\left(k\right) = \left\{3,8,11.5,15,14.5,7.5\right\}$ are weights derived from empirical ratings of dyads consonance used to adjust the contribution of each dimension k (Bernardes et al., 2018). k is set to $1 \leq k \leq 6$ for $T\left(k\right)$ since the remaining coefficients are symmetric. $T\left(k\right)$ uses $\overline{c}\left(n\right)$ which is $c\left(n\right)$ normalized by the DC component $T\left(0\right) = \sum_{n=0}^{N-1}c\left(n\right)$ to allow the representation and comparison of music at different hierarchical levels of tonal pitch (Bernardes et al., 2016). To represent variable-length audio tracks, the chroma vectors, $c\left(n\right)$, resulting from 16384 sample windows analysis at a 44.1 kHz sampling rate (≈ 372 ms) with 50% overlap are accumulated across the track duration.

From the audio tracks TIVs, two metrics that capture the harmonic compatibility between TIVs to be mixed are computed. Of note is the split between small- and large-scale harmonic compatibility, which roughly correspond to the sound object and meso or macro time scales of music, respectively. In other words, the small scale denotes the basic units of musical structure, from notes to beats, and the large scale inspects the structural levels between the phrase and the overall musical piece architecture (Roads, 2001). In the context of the current work, the first aims mostly at finding good harmonic matches between the tracks in a collection, and the second in guaranteeing control over the overall harmonic structure of a mix, i.e., the tonal changes at the key level across its temporal dimension.

Equation 2 computes the small-scale harmonic compatibility, $S_{i,p}$, between two given audio tracks, i and p, represented by their TIVs, $T_i\left(k\right)$ and $T_p\left(k\right)$, as the combination of two indicators: perceptual relatedness, $R_{i,p}$, (Equation 3) and dissonance, $D_{i,p}$, (Equation 4). The smaller the perceptual relatedness, $R_{i,p}$, the greater the affinity between two given tracks, as shown by Bernardes et al. (2016). The smaller the degree of dissonance, $D_{i,p}$, the greater their compatibility, as shown by the empirical data in Gebhardt et al. (2016).

$$S_{i,p} = R_{i,p} \cdot D_{i,p} \tag{2}$$

where:

$$R_{i,p} = \sqrt{\sum_{k=1}^{6}\left|T_i\left(k\right) - T_p\left(k\right)\right|^2} \tag{3}$$

$$D_{i,p} = 1 - \left(\frac{a_i T_i\left(k\right) + a_p T_p\left(k\right)}{a_i + a_p w\left(k\right)}\right) \tag{4}$$

a_i and a_p are the amplitudes of $T_i\left(k\right)$ and $T_p\left(k\right)$, respectively.

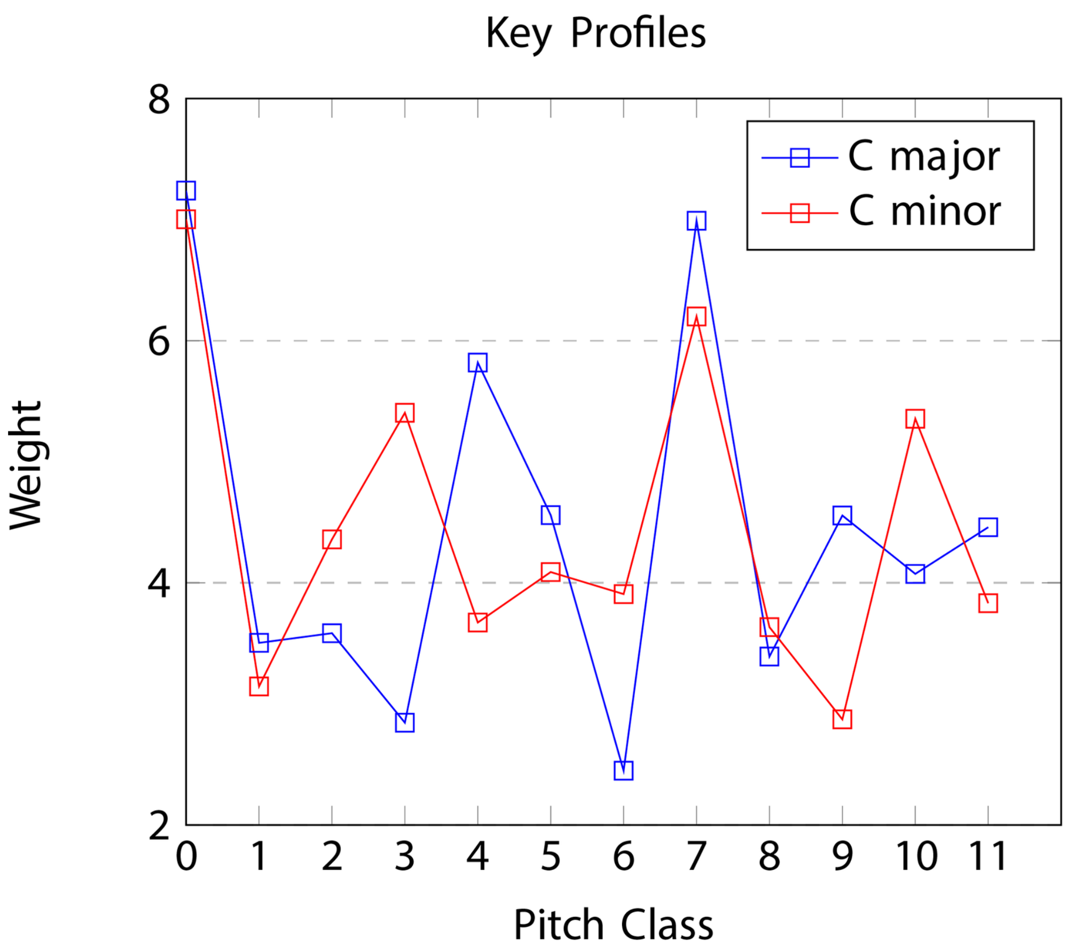
Large-scale harmonic compatibility, $L_{i,p}$, is a derivation of the perceptual relatedness, $R_{i,p}$, indicator, as it expresses the relatedness of a given track TIV from the $m = 24$ major and minor key TIVs, and can be interpreted as the degree of association of a given track to a musical key (Bernardes, Davies, & Guedes, 2017). As such, the large-scale harmonic compatibility can be computed by interpreting $T_i(k)$ and $T_p(k)$ in Equation 3 as a track TIV and a key TIV, respectively. The $m = 24$ major and minor keys TIV are computed by adopting the 12 shifts of the C major and C minor keys chroma vectors, $c(n)$, shown in Figure 2, in Equation 1.

Rhythmic, Spectral and Timbral Attributes

Three additional descriptions of rhythmic, spectral, and timbral audio track attributes are computed. They are subsidiary of the primary small- and large-scale harmonic compatibility metrics and aim to refine the search among compatible audio tracks. Next, a description of the metric and their musical interpretation in the context of music mashup creation, most notably in MixMash, is provided. Readers can refer to Brent (2010) for a comprehensive description of their computation.

Each audio track's rhythmic content is described by its note onset density, O_i, of a musical track i, and is computed by a threefold stage. First, a spectral flux function for each track is computed,

Figure 2. Sha'ath's (2011) key profiles for the C major and C minor keys



using the timbreID library (Brent, 2010) within Pure Data. This function describes the amount of novelty from a windowed power spectrum representation of the audio signal (window size ≈ 46 ms with 50% overlap). Second, the peaks from the function above a user-defined threshold, t, are extracted and interpreted as note onset locations. Prior to the peak detection stage, a bi-directional low-pass IIR filter, with a cut-off frequency of 5 Hz, was applied to avoid spurious detections. The note onset density, O_i, is then computed as the ratio between the number of onsets and the entire duration of the track (in seconds).

An indicator of the spectral region, B_i, of a musical track i, is given by the centroid of the accumulated Bark spectrum, b, across the duration of an audio track (Equation 5). The Bark spectrum, b, is computed by the timbreID library (Brent, 2010), which balances the frequency resolution across the human hearing range, by warping a power spectrum representation to the 24 critical bands, h, of the human auditory system (i.e., Bark bands).

$$B_i = \frac{\sum_{h=1}^{24} b_h \cdot h}{\sum_{h=1}^{24} b_h} \tag{5}$$

where b_h is the energy of the Bark band h. The B_i indicator can range from 1 to 24 Bark bands.

Following Pachet and Aucouturier (2004), the timbral similarity, $C_{i,p}$, between two tracks, i and p, is given by the cosine similarity between their mel-frequency spectrum coefficients (MFCC), M_i and M_p (Equation 6). MFCC vectors, M, include 38 components resulting from applying a 100 mel-scaled filter bank spacing in the timbreID library (Brent, 2010).

$$C_{i,p} = \frac{M_i \cdot M_p}{M_i M_p} \tag{6}$$

The timbral similarity metric, $C_{i,p}$, ranges between 1 and -1, which corresponds to tracks with equal timbre and the most dissimilar timbre, respectively.

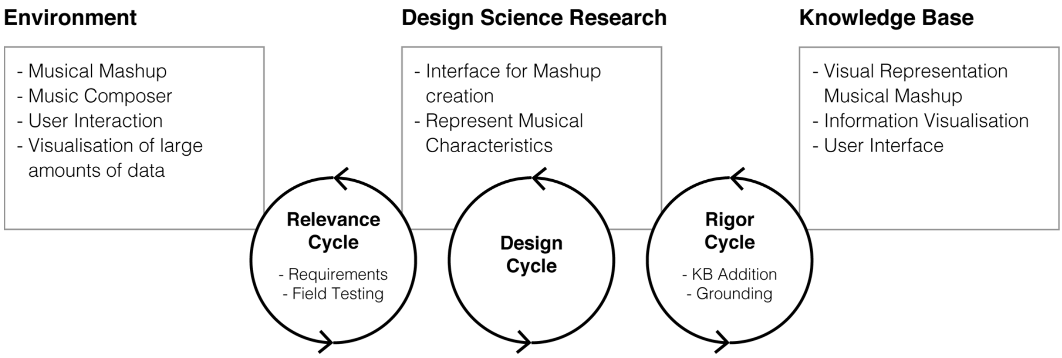
METHODOLOGY

The methodology used for the development of the present work is based on *A Three Cycle View of Design Science Research* (Hevner, 2007). This methodology facilitates the development of interactive applications and promotes quicker iterations between the several phases of the application development, such as the visualisation implementation, the visual component's validation, and the guidelines for its refinement (Figure 3).

The present methodology is divided into three parts: The Relevance Cycle, the Design Cycle, and the Rigor Cycle. Firstly, in the Relevance Cycle, the research context is defined. The requirements and problems from the previous system were highlighted, leading to the definition of the key objectives for the new system (see Introduction).

Secondly, in the Design Cycle, the new system was developed through a loop of research between the implementation of the visualisation components and their assessment. This iterative process promotes the analysis and improvement of previous steps and the experimentation and refinement of different approaches. In particular, this cycle was divided into four main phases: (i) the data analysis, which is related to the computation of the harmonic compatibility matrix (see MixMash: Compatibility

Figure 3. The Three Cycle View of design science research



Method); (ii) the design of the visualisation model, which relates to the iterative process between definition of the graphical variables to represent the musical characteristics and the visualisation model and its validation with experts in Information Visualisation; (iii) the implementation of the model, where the ForceAtlas2 algorithm was studied and implemented; and (iv) the evaluation of the system. The definition of such phases is aligned with the methodologies proposed by Wilkins (2003) and Fry (2004) from the Information Visualisation field. The validation of the system was conducted through an initial informal evaluation (Lam et al., 2011), in which the visualisation components were discussed between the authors and external experts from the visualisation field to assess the visualisation intuitiveness and usability.

Finally, the Rigor Cycle connects with the central Design Cycle through an iterative exchange of knowledge, both from scientific foundations as from the visualisation validation. This final cycle is characterised by fine-tuning the visualisation model through knowledge acquired from the evaluations and related work.

MIXMASH: VISUALISATION MODEL

The visual representation of the relationships between audio tracks were guided by three objectives (see Introduction): (i) the implementation of an adaptive visualisation model (see subsection Force-directed System); (ii) the distinction between the different sound characteristics of the several musical audio tracks (see subsection Graphic Variables and Audiovisual Mappings); and (iii) the conception and implementation of a simple and intuitive interface (see subsection Interface Design and Interaction).

To improve the scalability, interaction, and visualisation of the interface previously presented in Bernardes et al. (2018), a force-directed graph layout based on the ForceAtlas2 algorithm (Jacomy, Venturini, Heymann, & Bastian, 2014) was implemented. This visualisation technique enabled the creation of an emergent, organic, and appealing environment for the user. Furthermore, it improves the readability of the previous interface of Bernardes et al., by preventing overlaps, arranging the tracks in space by their harmonic compatibility, and by enabling the user to explore and interact with tracks of interest easily. Additionally, to characterise, distinguish and improve the readability of the tracks and to augment the graphic attributes that represent and distinguish the tracks, a graphical representation for the musical audio tracks was also studied and applied.

To enable the user to filter the track collection according to harmonic, rhythmic, spectral, and timbral attributes, a graphic interface was also implemented (Figure 4). With the force-directed algorithm, the user can detect the most harmonically compatible tracks through their visual proximity in the canvas. This is caused by the forces applied to each track, which depend directly on the harmonic compatibility. However, the user can manipulate the impact of the forces of attraction and repulsion between tracks, easing the comprehension of more cluttered zones. Additionally, highly compatible tracks are clustered to reduce undesired clutter. These clusters are visually distinguished from the tracks and can be expanded or withdrawn through interaction. The force-directed algorithm and each component of the interface are described in more detail in the following subsections.

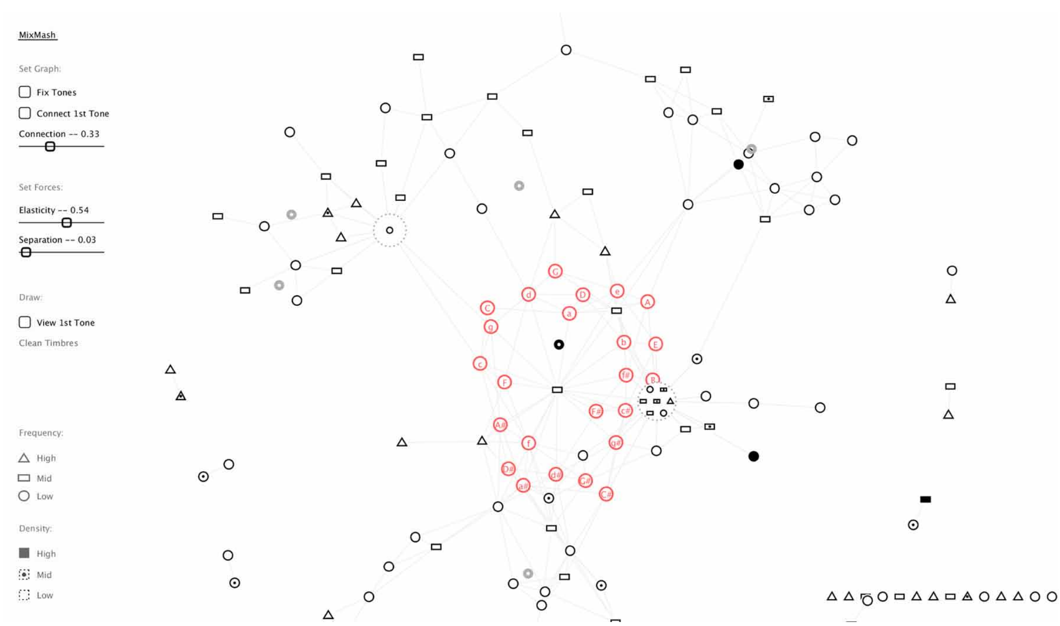
Force-Directed System

The ForceAtlas2 algorithm (Jacomy, Venturini, Heymann, & Bastian, 2014) can be characterised by its ability to place the nodes within a graph according to their connections weight. The algorithm simulates a physical system that spatially arranges the network's nodes in an automatic form. The nodes have forces of repulsion to prevent them from overlapping, and the edges between nodes apply forces of attraction to bring the nodes closer. These edges have different force values according to the similarity between nodes (e.g., their harmonic compatibility). By applying continuously, the different forces, the graph converges to a balanced state that aids the semantic interpretation of the network. The ForceAtlas2 algorithm was fully implemented, thus for a detailed description of the algorithm, refer to Jacomy et al. (2014).

In this project, two types of nodes are defined: the ones representing each musical audio track in the collection, t, and the ones representing each one of the m = 24 major and minor musical keys. The visual distinction between the nodes is discussed at length in subsection *Graphic Variables and Audiovisual Mappings*. By representing the musical keys, m, through nodes, and consequently their relation to the musical audio tracks, the most compatible key of each musical audio is indicated. This enables the user to visually detect the tracks and sets of tracks more harmonically compatible with the different keys and relate tracks according to this compatibility.

The weight of the edge between two nodes is mapped according to the compatibility value in the $t + m$ square matrix. This square matrix is computed through a given a user-defined collection of t audio tracks and m = 24 major and minor keys and expresses the metrics for small-scale ($S_{i,p}$) and large-scale ($L_{i,p}$) scale harmonic compatibility (see MixMash: Compatibility Method). As in ForceAtlas2, the weight defines the force of attraction between nodes. For this project, the more

Figure 4. Screenshot of the visualisation interface. On the left, the interaction panel; at the centre, the graph.



compatible the nodes, the higher the force of attraction, and consequently the closer the nodes. A force of repulsion is applied to all nodes to avoid them to overlap. These forces are applied to the nodes independently of their type, facilitating the interpretation of the graph and interaction when searching for harmonically compatible tracks, *t*.

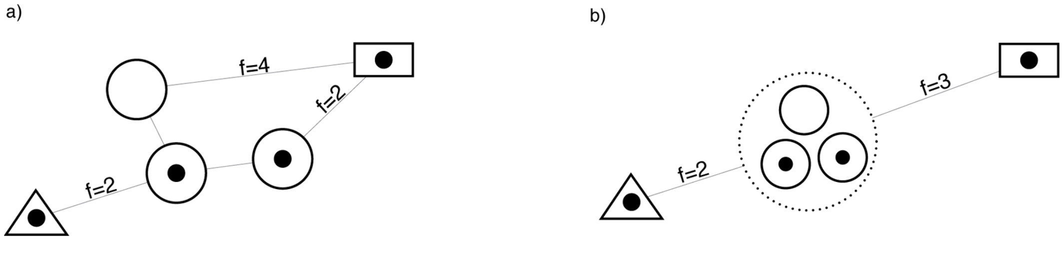
Two mechanisms were implemented to enable the user to refine the graph layout: (i) a connectivity threshold; and (ii) musical key restriction. Both mechanisms can be explored by the user through the left panel of the interface (Figure 4). Through the connectivity threshold, the user can define a threshold value to determine whether two nodes are connected. For each node, the connections to other nodes only occur when their harmonic compatibility value is lower than the predefined value. The second mechanism limits the connections between nodes and keys. Through this mechanism, the user can define whether a track is connected to all compatible musical keys or only to its most probable key. This mechanism is intended to reduce clutter and enhance the association between keys and musical audio tracks.

As the number of tracks in a collection can vary greatly, an agglomerative clustering algorithm (Rokach & Maimon, 2005) is implemented to prevent cluttered graphs. This algorithm aggregates the nodes by their compatibility values. A minimum number of three nodes per cluster is required to prevent small clusters, which wouldn't enhance the clarity of the visualisation. Each cluster is also connected to the compatible nodes and clusters, thus exposing their harmonic relation to the neighbourhood elements. These compatible nodes are retrieved from the list of compatible nodes of each node within the cluster. If an outer node is connected to an inner node of a cluster, a connection between the node and the cluster is established. The attraction force between a cluster and a compatible node is equal to the average force of all forces between inner nodes and the connected outer node (as depicted in Figure 5).

As in ForceAtlas2, all nodes gravitate around the centre of the canvas. This effect results from the attraction force applied to all nodes towards the canvas centre point. The gravitational force is significantly weaker than the others and, as it is applied equally to all nodes, it does not interfere with the distance between nodes, and only prevents the nodes from dispersing in the canvas.

By default, the musical key nodes are positioned by the force-directed layout, depending on their relations with the other tracks. This causes the key nodes to have different positions at every run, which can create some user fatigue while searching for a specific musical key and its connected tracks. To facilitate this search, a mechanism that positions the key nodes according to the Camelot Wheel or circle of fifths representation was implemented (Figure 6). With this mechanism, all musical keys have a fixed position in the canvas, and only the track nodes are influenced by the attraction and repulsion forces of the algorithm.

Figure 5. Scheme of the forces representing outer and inner nodes of a cluster. The forces between clusters and nodes (b) are computed through the average force between the inner and outer nodes (a). The lines depict compatible nodes.



Graphic Variables and Audiovisual Mappings

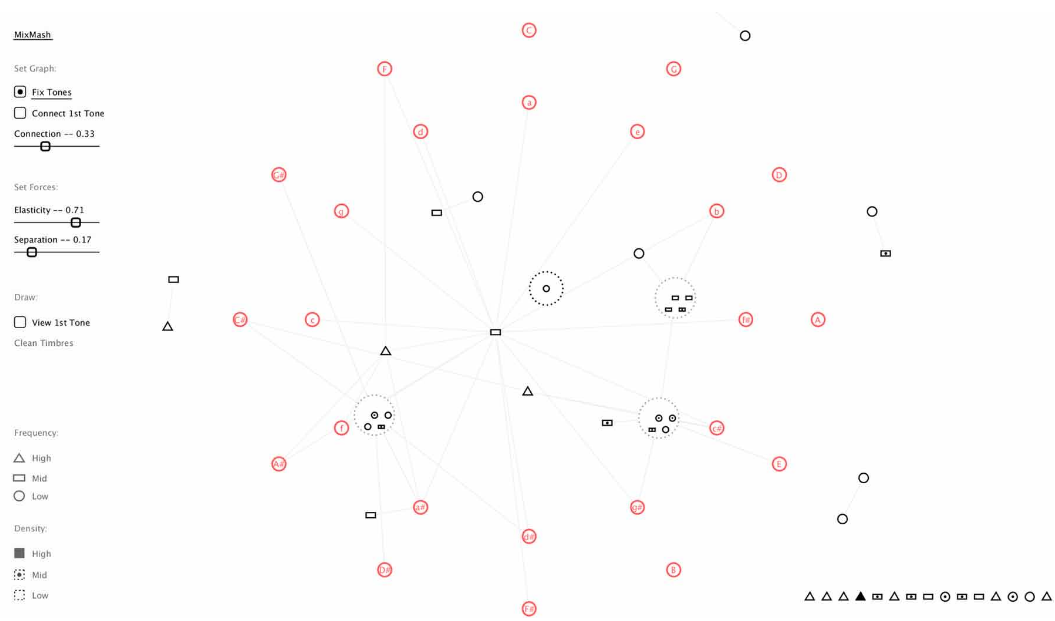
The visual representation of data elements strongly impacts the visualisation of large amounts of data. In this section, the adopted visual representation of musical concepts is discussed in light of a carefully designed interactive visualisation.

The development of the visualisation model complies with the following guidelines and subsequent challenges: (i) to each track there is a corresponding visual representation based on its musical attributes, creating a consistent visual feedback between similar tracks; (ii) perceptual foundations are used to guide the aesthetical choices concerning visual representation of tracks; (iii) a natural and intuitive interaction with the tool is promoted allowing the user to easily navigate among tracks to create his/her mashup.

For each track, the spectral region, B_i, onset density, O_i, and timbral similarity, $C_{i,p}$, are mapped to a corresponding visual variable. Additionally, both spectral region, B_i, and onset density, O_i, are subdivided into three levels of magnitude, allowing a more accurate interpretation and analysis of musical data.

The spectral region attribute, B_i, is split into high, medium, and low regions, and its corresponding visual representation is primarily characterised by shape. Although there is a tendency to associate colour with pitch levels in similar audiovisual mapping problems (McDonnell, 2007; Mudge, 1920; Ward, Huckstep, & Tsakanikos, 2006), it has also been proposed that the use of shape or monochrome colour is more efficient than hue to distinguish several layers of information for the human eye (Arnheim, 1974; Chatterjee, 2013). Based on such studies (Arnheim, 1974; Chatterjee, 2013; Ramachandran 2003; Spence, 2011), high-frequency regions of the spectrum are associated with sharp contours (triangle), low frequencies were associated with rounded contours (circle), and medium frequencies are associated with neutral contours, achieved through the use of straight lines of the rectangle (Figure 7a). Colour hue, is used to reinforce these associations. Higher frequencies are related to a cold colour (blue and low frequencies to a warm colour (orange; see Figure 7c).

Figure 6. Fixing the keys in the circle of fifths. By clicking on the Fix Tones button, all nodes that represent a key are fixed on the canvas according to their positioning in the circle of fifths.



Onset density, O_i, is split into high, medium and low-density levels, and then conveyed in the visual domain as a parameter of shape-filling, as it relates to the number of notes in a track. Low density corresponds to an empty shape (only contours visible), medium density to a half-filled shape, and high density to a completely filled shape (Figure 7b).

Timbre is often referred to in the literature as the colour of instruments (Mudge, 1920; Ward et al., 2006). Inspired by this definition, timbral similarity, $C_{i,p}$, is conveyed in the visual domain through coloured circles in the top side of the nodes, which are only visible when a certain node is selected. The choice of colours does not rely on any type of association, and for this reason, are randomly selected from a predefined set of colours. They simply aim to provide a clear distinction between different timbres when multiple timbres are selected. By selecting one track node, a coloured circle and a line connecting the former to the node are drawn (Figure 8a). Then, all tracks with similar timbres will also gain a circle coloured with the same colour code as the clicked one (Figure 8b). If a track node is similar to multiple timbres, multiple coloured circles will be drawn over the node (Figure 8c).

The visual distinction between nodes representing musical audio tracks and keys is highlighted by the coloured outline of the key nodes. The latter has a red outline and include the typographic representation of the key's tonic pitch, which allows the direct reading of the key (Figure 4).

The visual representation of a cluster is defined by the respective inner nodes to avoid the use of additional (and potentially, more complex) visual elements. More specifically, all nodes that belong to a cluster are represented and positioned within a circular shape outlined with black dots. The size of the circular shape depends on the number of elements within the cluster (Figure 9). As such, the user can differentiate the clusters from the nodes, and, simultaneously, get an overview of their inner nodes.

Figure 7. Audiovisual mappings: (a) Shape—Spectrum Frequency from low frequencies (circle) to higher ones (triangle); (b) Object Fill—Onset Density. Empty shapes correspond to low density, half-filled shapes to normal/mid density, and full shape corresponds to high density; (c) Colour—Spectrum Frequency from low frequencies (orange) to higher ones (blue).

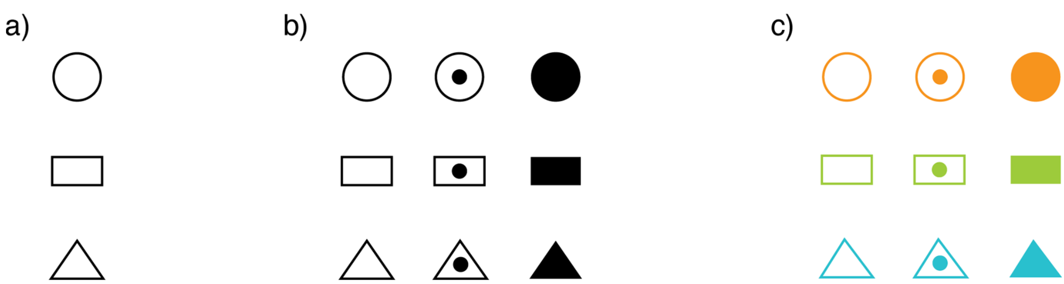
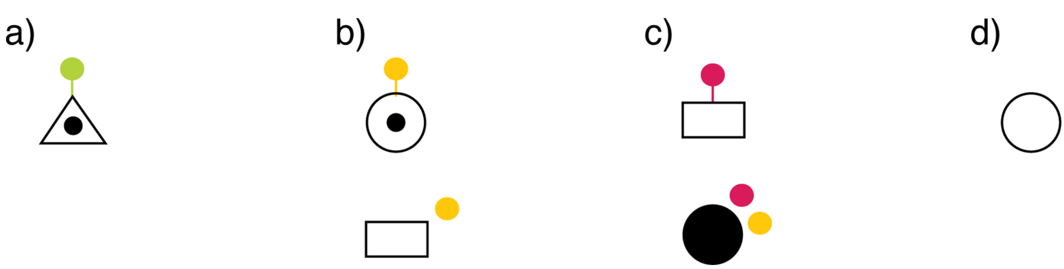


Figure 8. Timbre representations. (a) One track is selected and there are no other tracks with similar timbres. (b) One track is connected timbre wise to the selected track. (c) The track can have more than one timbre similarities (d) The track is not related to any of the selected tracks.



Interface Design and Interaction

MixMash enables the user to explore the graph by allowing him/her to (i) listen and select individual node tracks of interest (Figures 10a, 10b, 10c); (ii) highlight nodes within clusters (Figure 10d) or according to their sound characteristics (Figure 11); (iii) modify its organisation (e.g., fixing the keys according to the circle of fifths) (Figures 10e, 10f); (iv) change how the track nodes connect to the key nodes (Figures 10g, 10h); (v) change the connection threshold between nodes; and (vi) adjust the forces of attraction and repulsion. This is all accessible through an interactive panel on the left side of the interface (Figure 4) and through mouse interactions. A video regarding these interactions and possibilities can be accessed at https://vimeo.com/270076175.

Once the system has been initialised, the user will see the nodes and clusters establishing their position in the centre of the canvas. In addition to the functionalities present in the left panel, the user can interact with the visualisation through mouse interactions. The user can zoom and pan the graph to view more details. Then, the user can interact with the tracks individually. To listen to the tracks, the user has to move the mouse over each node. To select a node, the user needs to left-click. With this action, he/she will listen again to the track sound. To view the compatible timbres of a certain track, the user has to right-click. By doing so, the closest tracks (in terms of timbre) are complemented with a coloured circle as explained in subsection *Graphic Variables and Audiovisual Mappings*. Interactions with the clusters were also implemented. To expand a cluster, the user has to left-click over it. All nodes inside the cluster will be affected by the forces according to the ForceAtlas2 algorithm, and a doughnut shape figure will be made visible. The doughnut shape, positioned in the centre of all corresponding nodes, behaves like a button and, when clicked, it closes the cluster. If the user does mouseover on this latter shape, all the nodes belonging to the cluster will be highlighted through a magenta stroke (Figure 10d). Finally, all the tracks that have no similarity to others will be placed at the bottom right side of the window.

Figure 9. Representation of two clusters with different sizes

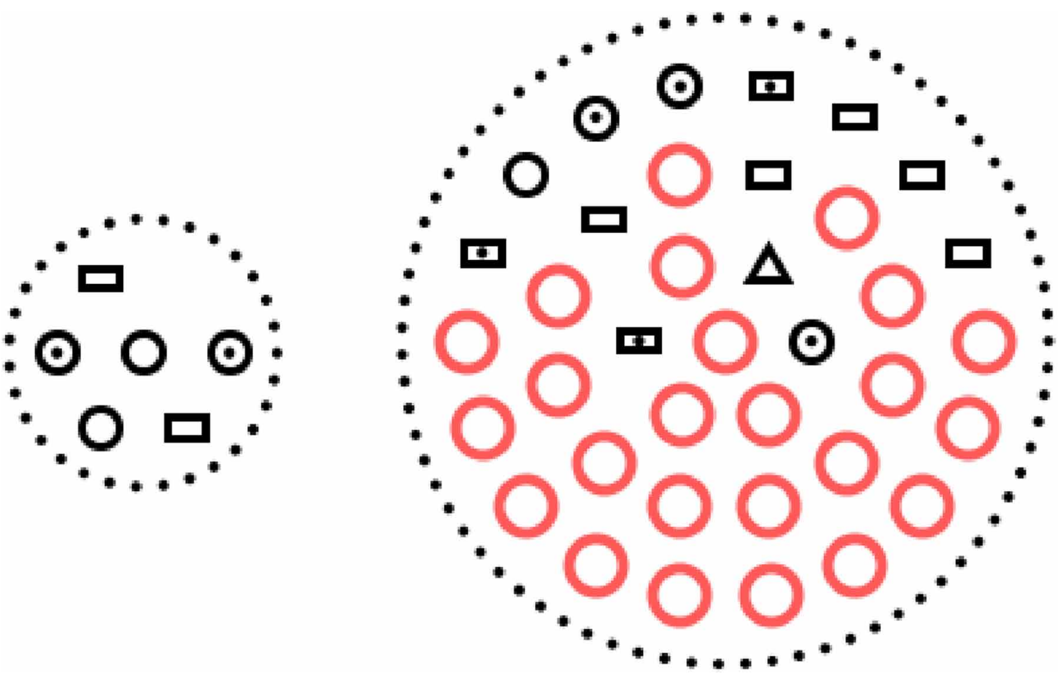
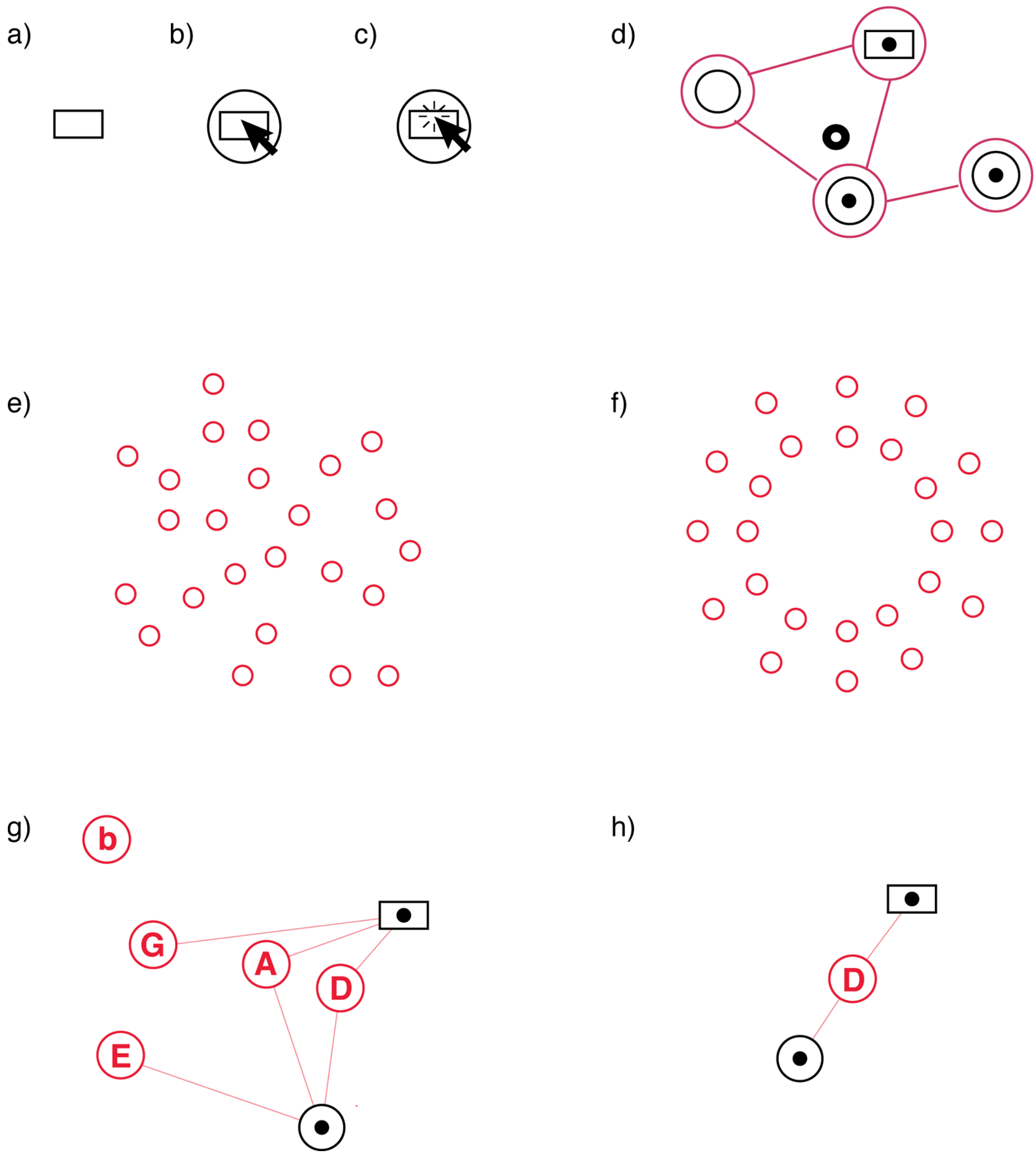


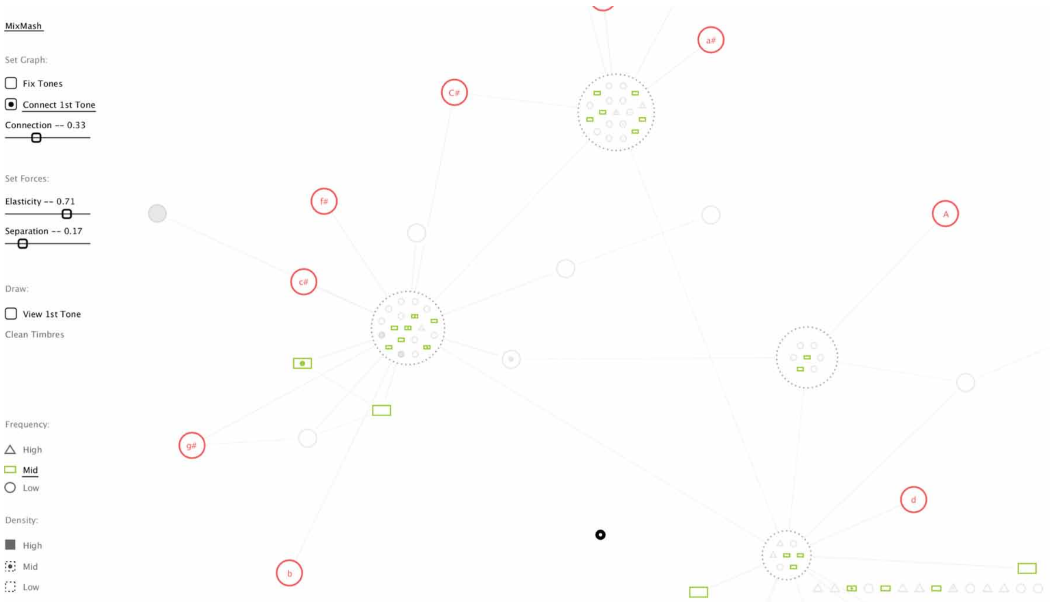
Figure 10. Visual representation of the interactions with the model: (a) no selection, (b) mouse over, (c) mouse click, (d) clusters' nodes highlight, (e) force-directed layout, (f) circle of fifths layout, (g) connections to all compatible keys, (h) connection to the most compatible key (1st tone option)



Additional keyboard interactions were implemented. When listening to an audio track, the user can click the S key on the keyboard, and the music will stop playing. By continuously selecting nodes, the user is saving the tracks, which can be heard at the same time by using the space key.

Overall, this set of interaction techniques are important to achieve an intuitive and meaningful interactive visualisation tool in the context of musical mashup creation. With this, the authors aim to enhance the understanding of the track's harmonic compatibility and foster user creativity, by allowing the user to efficiently explore a large musical audio collection towards specific composition goals.

Figure 11. Frequency highlight by color



CONCLUSION

The authors proposed a novel visualisation system which relies on forces of attraction and repulsion to position the tracks depending on their harmonic compatibility. The visualisation development was guided by a methodology proposed by Hevner (2007), which consists of three cycles and lead to clear and consistent interactions between the music mashup and the information visualisation model.

Regarding the force-directed algorithm, controlled levels of attraction and repulsion allow the reduction of clutter in the visualization of large music collections (of roughly 50 musical audio tracks or more). Clutter was also minimized by the adoption of clustering techniques, which enhance the visualization of combined hierarchical levels of harmonic compatibility in the same representation and the user-control over clustering quality (i.e., distance threshold).

A fluid re-organization of the visualization was achieved by dragging connected elements in the interface, thereby enhancing highly dense areas of particular interest to the user. On the other hand, the Fix Tones strategy explores a static visualization of the musical audio collection by fixing the location of keys in the Tonal Interval Space. The resulting representation is one of the most familiar maps of tonal regions in Western music.

The authors were able to expand the number of content-based musical audio attributes under consideration to cover both rhythmic, harmonic, spectral, and timbral attributes. The development of a specific graphic representation supported the music visualisation by providing a perceptual association, and therefore, the intuitive association between visual and musical attributes. In general, the presented visual solution was able to promote a more fluid visualisation. However, it still has some limitations due to the high number of samples that are being displayed in real time to the user.

As future work, the authors intend to improve the clustering algorithm by giving the user the possibility to cluster nodes according to different audio content-based attributes, e.g., key, onset density, spectral region or timbral similarity. To improve readability, different solutions to the nodes' size depending on their compatibility will be studied, e.g., nodes with higher compatibility, grow in size, emphasising highly compatible tracks. As the number of tracks can increase depending on the user, a fish-eye zoom technique will be implemented so the user can have detail in certain areas without losing context of the surrounding tracks. Finally, a timeline will be designed, so that users can arrange selected tracks in time, thus enabling the composition of musical mashups with complex temporal structures.

ACKNOWLEDGMENT

The work is supported by Portuguese National Funds through the FCT-Foundation for Science and Technology, I.P., under the grants SFRH/BD/129481/2017 and SFRH/BD/139775/2018. Experimentation in music in Portuguese culture: History, contexts, and practices in the 20th and 21st centuries—Project co-funded by the European Union, through the Operational Program Competitiveness and Internationalization, in its ERDF component, and by national funds, through the Portuguese Foundation for Science and Technology.

REFERENCES

Arnheim, R. (1974). *Art and visual perception*. University of California Press.

Bergstrom, T., Karahalios, K., & Hart, J. (2007). Isochords: Visualizing structure in music. In Proceedings of Graphics Interface (pp. 297–304). ACM.

Bernardes, G., Cocharro, D., Caetano, M., Guedes, C., & Davies, M. (2016). A multi-level tonal interval space for modelling pitch relatedness and musical consonance. *Journal of New Music Research*, *45*(4), 281–294. doi:10.1080/09298215.2016.1182192

Bernardes, G., Davies, M., & Guedes, C. (2017). Automatic musical key estimation with adaptive mode bias. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing* (pp. 316–320). IEEE Press. doi:10.1109/ICASSP.2017.7952169

Bernardes, G., Davies, M., & Guedes, C. (2018). A hierarchical harmonic mixing method. In M. Aramaki et al. (Eds.), International Symposium on Computer Music Multidisciplinary Research (pp. 151–170). Springer Nature Switzerland.

Brent, W. (2010). A timbre analysis and classification toolkit for Pure Data. In *Proceedings of the International Computer Music Conference*. Academic Press.

Chatterjee, A. (2013). *The aesthetic brain: How we evolved to desire beauty and enjoy art*. Oxford University Press. doi:10.1093/acprof:oso/9780199811809.001.0001

Chen, C. (2006). CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the Association for Information Science and Technology*, *57*(3), 359–377.

Davies, M., Hamel, P., Yoshii, K., & Goto, M. (2014). Automashupper: Automatic creation of multi-song music mashups. *IEEE Transactions on Audio, Speech, and Language Processing*, *22*(12), 1726–1737. doi:10.1109/TASLP.2014.2347135

Davies, M., Stark, A., Gouyon, F., & Goto, M. (2014). Improvasher: A real-time mashup system for live musical input. *Proceedings of the International Conference on New Interfaces for Musical Expression,* London, UK (pp. 541–544). Academic Press.

Dobrowohl, F. A., Milne, A. J., & Dean, R. T. (2019). Timbre preferences in the context of mixing music. *Applied Sciences*, *9*(8), 1695. doi:10.3390/app9081695

Enright, A., & Ouzounis, C. (2001). Biolayout—An automatic graph layout algorithm for similarity visualization. *Bioinformatics (Oxford, England)*, *17*(9), 853–854. doi:10.1093/bioinformatics/17.9.853

Fry, B. J. (2004). *Computational information design* [PhD thesis]. Massachusetts Institute of Technology.

Fu, X., Hong, S.-H., Nikolov, N., Shen, X., Wu, Y., & Xuk, K. (2007). Visualization and analysis of email networks. *Proceedings of the International Asia-Pacific Symposium on Visualization* (pp. 1–8). IEEE.

Gebhardt, R., Davies, M., & Seeber, B. (2016). Psychoacoustic approaches for harmonic music mixing, *Applied Sciences, 6*(5).

Gibney, M. (n.d.) Music-map, the tourist map of music. Retrieved from https://www.music-map.com

Gilbert, F., Simonetto, P., Zaidi, F., Jourdan, F., & Bourqui, R. (2011). Communities and hierarchical structures in dynamic social networks: Analysis and visualization. *Social Network Analysis and Mining*, *1*(2), 83–95. doi:10.1007/s13278-010-0002-8

Goh, K.-I., Cusick, M., Valle, D., Childs, B., Vidal, M., & Barabási, A.-L. (2007). The human disease network. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(21), 8685–8690. doi:10.1073/pnas.0701361104

Griffin, G., Kim, Y., & Turnbull, D. (2010). Beat-sync-mash-coder: A web application for real-time creation of beat-synchronous music mashups. *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing* (pp. 437–440). Academic Press. doi:10.1109/ICASSP.2010.5495743

Grill, T., & Flexer, A. (2012). Visualization of perceptual qualities in textural sounds. *Proceedings of the International Computer Music Conference*. Academic Press.

Hamasaki, M., & Goto, M. (2013). Songrium: A music browsing assistance service based on visualization of massive open collaboration within music content creation community. *Proceedings of the International Symposium on Open Collaboration*. ACM. doi:10.1145/2491055.2491059

Heer, J., & Boyd, D. (2005). Vizster: Visualizing online social networks. *Proceedings of the IEEE Symposium on Information Visualization* (pp. 32–39). IEEE.

Henry, N., Fekete, J.-D., & McGuffin, M. (2007). Nodetrix: A hybrid visualization of social networks. *IEEE Transactions on Visualization and Computer Graphics*, *13*(6), 1302–1309. doi:10.1109/TVCG.2007.70582

Herman, I., Melançon, G., & Marshall, M. (2000). Graph visualization and navigation in information visualization: A survey. *IEEE Transactions on Visualization and Computer Graphics*, *6*(1), 24–43. doi:10.1109/2945.841119

Hevner, A. R. (2007). A three cycle view of design science research. *Scandinavian Journal of Information Systems*, *19*(2), 4.

Jacomy, M., Venturini, T., Heymann, S., & Bastian, M. (2014). ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software. *PLoS One*, *9*(6), e98679. doi:10.1371/journal.pone.0098679

Kimelman, D., Leban, B., Roth, T., & Zernik, D. (1994). Reduction of visual complexity in dynamic graphs. *Proceedings of the International Symposium on Graph Drawing* (pp. 218–225). Springer.

Lam, H., Bertini, E., Isenberg, P., Plaisant, C., & Carpendale, S. (2011). *Seven guiding scenarios for information visualization evaluation*. University of Calgary.

Lee, C., Lin, Y., Yao, Z., Li, F., & Wu, J. (2015). Automatic mashup creation by considering both vertical and horizontal mashabilities. *Proceedings of the International Society for Music Information Retrieval Conference,* Malaga, Spain (pp. 399–405). Academic Press.

Maçãs, C., Rodrigues, A., Bernardes, G., & Machado, P. (2018). MixMash: A visualisation system for musical mashup creation. *Proceedings of the 2018 22nd International Conference Information Visualisation (IV)* (pp. 471–477). IEEE. doi:10.1109/iV.2018.00088

McDonnell, M. (2007). Visual music. In *Visual Music Marathon*. Boston Cyberarts Festival Programme.

Meirelles, I. (2013). *Design for information*. Rockport publishers.

Mixed in Key. (n.d.). Mashup 2 [software]. Retrieved from http://mashup.mixedinkey.com

Mudge, E. (1920). The common synaesthesia of music. *The Journal of Applied Psychology*, *4*(4), 342–345. doi:10.1037/h0072596

Muelder, C., Provan, T., & Ma, K.-L. (2010). Content based graph visualization of audio data for music library navigation. *Proceedings of the IEEE International Symposium on Multimedia* (pp. 129–136). IEEE Press. doi:10.1109/ISM.2010.27

Native Instruments. (n.d.) Traktor Pro 2 [software]. Retrieved from https://www.native-instruments.com/en/products/traktor/dj-software/traktor-pro-2/

Navas, E. (2014). *Remix theory: The aesthetics of sampling*. Birkhäuser.

Pachet, F., & Aucouturier, J.-J. (2004). Improving timbre similarity: How high is the sky. *Journal of Negative Results in Speech and Audio Sciences*, *1*(1), 1–13.

Ramachandran, V. S., & Hubbard, E. M. (2003). Hearing colors, tasting shapes. *Scientific American*, *288*(5), 52–59. doi:10.1038/scientificamerican0503-52

Rauber, A., Pampalk, E., & Merkl, D. (2003). The SOM-enhanced jukebox: Organization and visualization of music collections based on perceptual models. *Journal of New Music Research*, *32*(2), 193–210. doi:10.1076/jnmr.32.2.193.16745

Rodrigues, A., Cardoso, A., & Machado, P. (2016). Harmonic constellation: An audiovisual environment of living organisms. *Proceedings of the Fourth International Workshop on Musical Metacreation*. Academic Press.

Rokach, L., & Maimon, O. (2005). Clustering Methods. In O. Maimon & L. Rokach (Eds.), *Data mining and knowledge discovery handbook*. Boston: Springer. doi:10.1007/0-387-25465-X_15

Sapp, C. (2001). Harmonic visualizations of tonal music. *Proceedings of the International Computer Music Conference* (pp. 419–422). Academic Press.

Schedl, M., Gómez, E., & Urbano, J. (2014). Music information retrieval: Recent developments and applications. *Foundations and Trends in Information Retrieval*, *8*(2–3), 127–261. doi:10.1561/1500000042

Schwarz, D., & Schnell, N. (2009). Sound search by content-based navigation in large databases. *Proceedings of the Sound and Music Computing Conference*. Academic Press.

Shaath, I. (2011). *Estimation of key in digital music recordings* [Master's thesis]. Birkbeck College, University of London.

Shiga, J. (2007). Copy-and-persist: The logic of mash-up culture. *Critical Studies in Media Communication*, *24*(2), 93–114. doi:10.1080/07393180701262685

Snydal, J., & Hearst, M. (2005). Improviz: Visual explorations of jazz improvisations. In CHI'05 Extended Abstracts on Human Factors in Computing Systems (pp. 1805–1808). ACM.

Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception & Psychophysics*, *73*(4), 971–995. doi:10.3758/s13414-010-0073-7

Van Gulik, R., Vignoli, F., & Van de Wetering, H. (2004). Mapping music in the palm of your hand, explore and discover your collection. In *Proceedings of the International Conference on Music Information Retrieval*. Retrieved from http://www.ee.columbia.edu/~dpwe/ismir2004/CRFILES/paper153.pdf

Vavrille, F. (n.d.). Liveplasma. Retrieved from http://www.liveplasma.com/artist-Hard-Fi.html

Vesna, V. (2007). *Database aesthetics: Art in the age of information overflow*. University of Minnesota Press. doi:10.5749/j.cttts7q7

Ward, J., Huckstep, B., & Tsakanikos, E. (2006). Sound-colour synaesthesia: To what extent does it use crossmodal mechanisms common to us all? *Cortex*, *42*(2), 264–280. doi:10.1016/S0010-9452(08)70352-6

Wattenberg, M. (2002). Arc diagrams: Visualizing structure in strings. *Proceedings of the IEEE Symposium on Information Visualization* (pp. 110–116). IEEE.

Wilkins, B. (2003). *MELD: A pattern supported methodology for visualisation design* [Doctoral thesis]. University of Birmingham.

ENDNOTE

[1] In the audio domain, 12-element chroma vectors report the energy of the twelve pitch classes, i.e., all chromatic notes of the equal-tempered scale, by wrapping the spectral energy content of an audio signal into a single octave.

Catarina Maçãs has an Undergraduate and Master degree in Design and Multimedia from University of Coimbra. Currently is a computational designer and researcher at the Computational Design and Visualization Lab, is enrolled in the Doctoral Program of Information Science and Technology of the University of Coimbra, and is also a teaching assistant at the Department of Informatics Engineering of the University of Coimbra. In the past two years, started to work on Information Visualization and had already published in conferences such as, IJCAI and IVAPP.

Ana Rodrigues is a Computational Designer and Ph.D. Student. Aside from being currently enrolled in the Program of Information Science and Technology at the University of Coimbra, working on her thesis entitled "Cross-Modal Associations for Design Languages", she is also doing research at the Computational Design and Visualization Lab. At the same institution she has previously completed a B.Sc. degree and a M.Sc. degree in Design and Multimedia. Currently, she also performs the role of invited assistant lecturer. She is particularly interested in exploring bridges that may emerge from crossing domains of visual communication, music, neurocognitive science, and computational creativity.

Gilberto Bernardes has a multifaceted activity as a saxophonist and researcher in sound and music computing. He holds a PhD in digital media from the University of Porto under the auspices of the University of Texas at Austin; a Master of Music, cum laude, from the Amsterdamse Hogeschool voor de Kunsten and a Degree in music performance from the Superior School of Music and Performing Arts of the Polytechnic Institute of Porto and from the ENM d'Issy-les-Moulineaux (France). Bearnardes pursues a research agenda focused on sampling-based synthesis techniques and tonal pitch spaces, whose findings have been shared to the scientific and artistic communities in over 30 scientific publications. His artistic activity counts with regular concerts in venues and festivals with recognized merit, such as Asia Culture Center (Korea); New York University (USA); Concertgebouw and Bimhuis (Holland); Teatro Monumental de Madrid (Spain) and Casa da Música (Portugal). He is a member of the Portuguese Symphonic Band and Oficina Musical. Bernardes is currently an Assistant Professor at the University of Porto, Faculty of Engineering and an fellow researcher at INESC TEC.

Penousal Machado is Associate Professor at the Department of Informatics of the University of Coimbra in Portugal. He is a deputy director of the Centre for Informatics and Systems of the University of Coimbra (CISUC), the coordinator of the Cognitive and Media Systems group and the scientific director of the Computational Design and Visualization Lab. of CISUC. His research interests include Evolutionary Computation, Computational Creativity, Artificial Intelligence and Information Visualization. He is the author of more than 200 refereed journal and conference papers in these areas, and his peer-reviewed publications have been nominated and awarded multiple times as best paper. As of February 28, 2018, his publications gathered over 1828 citations, an h-index of 21, and an i10-index of 40. He is also the chair of several scientific events, including, amongst the most recent, PPSN XV, EvoStar 2016, and EuroGP 2015; member of the Programme Committee and Editorial Board of some of the main conferences and journals in these fields; member of the Steering Committee of EuroGP, EvoMUSART and Evostar; and executive board member of SPECIES. He is the recipient of several scientific awards, including the prestigious EvoStar Award for outstanding Contribution to Evolutionary Computation in Europe, and the award for Excellence and Merit in Artificial Intelligence granted by the Portuguese Association for Artificial Intelligence. Penousal Machado has been invited to perform keynote speeches in a wide set of domains, from evolutionary computation to visualization and art. His work was featured in the Leonardo journal, Wired magazine and presented in venues such as the National Museum of Contemporary Art (Portugal) and the "Talk to me" exhibition of the Museum of Modern Art, NY (MoMA). He has been the Principal Investigator and Coordinator of several national and international projects, both at an academic and industry level.