Multimedia Tools and Applications manuscript No.
(will be inserted by the editor)

Enlightening the Sun

A User Interface to Explore Music Artists via Multimedia Content

Markus Schedl · Tim Pohle

Received: date / Accepted: date

Abstract This article presents an approach to browse collections of web pages about music artists by means of descriptive terms and multimedia content. To this end, a user interface called *Three-Dimensional Co-Occurrence Browser (3D-COB)* is introduced. 3D-COB automatically extracts and weights terms from artist-related web pages. This textual information is complemented with information on the multimedia content found on the web pages. For the user interface of 3D-COB, we elaborated a three-dimensional extension of the Sunburst visualization technique. The hierarchical data to be visualized is obtained by analyzing the web pages for combinations of co-occurring terms that are highly ranked by a term weighting function.
We further investigated, in a first user study, different term weighting strategies to generate the visualization. A second user study was carried out to assess ergonomic aspects of 3D-COB, especially its usefulness for gaining a quick overview of a set of web pages and for efficiently browsing within this set.

Keywords User Interface · Web Mining · Multimedia Information Retrieval · Co-Occurrence Browser · Evaluation

Markus Schedl
Department of Computational Perception
Johannes Kepler University
Linz, Austria
Tel.: +43-732-2468-1512
Fax: +43-732-2468-1520
E-mail: markus.schedl@jku.at

Tim Pohle
Department of Computational Perception
Johannes Kepler University
Linz, Austria
Tel.: +43-732-2468-1514
Fax: +43-732-2468-1520
E-mail: tim.pohle@jku.at

1 Introduction

Automatically finding descriptive terms for a given music artist is an important task in music information retrieval, a subfield of multimedia information retrieval. Such terms may describe, for example, the genre or style of the music performed by the artist under consideration and enable a wide variety of applications, e.g. enriching music players [1], recommending unknown artists based on the user's favorite artists (recommender systems) [2], enhancing user interfaces for browsing music collections [3], [4], [5], [6], [7], or automatically tagging of artists [8].
One possibility for assigning musically relevant terms to a given artist is manual annotation by music experts or communities, as it is usually employed by music information systems like *allmusic* [9] and *last.fm* [10] or interfaces for music search like *musiclens* [11]. However, this is a very labor-intensive task and barely feasible for huge music collections. An alternative way, which we follow here, is to exploit today's largest information source, the World Wide Web. Automatically deriving information about music artists from the web is advantageous since it incorporates the opinions of a large number of different people, and thus embodies a kind of cultural knowledge.
The *Three-Dimensional Co-Occurrence Browser (3D-COB)* presented here automatically indexes a set of web pages about music artists according to a dictionary of musically relevant terms and organizes these web pages by creating a number of subsets, each of which is described by a set of terms. The terms that describe a particular subset are determined by a term weighting function. The subsets are then visualized using a variant of the well-established Sunburst technique [12], [13], aka InterRing [14].
The purpose of 3D-COB is threefold. First, it facilitates getting an *overview* of the web pages related to a music artist by structuring them according to co-occurring terms. Second, since the descriptive terms that most often occur on web pages related to a music artist X constitute an individual profile of X, 3D-COB is also suited to reveal various *meta-information* about the artist, e.g. musical style, related artists, or instrumentation. Third, by visualizing the amount of *multimedia content* provided at the indexed web pages, the user is offered a means of exploring the audio, image, and video content of the respective set of web pages.

2 Related Work

This paper is mainly related to the two research fields of *web-based music information retrieval (MIR)* and *information visualization of hierarchical data*, which will be covered in the following.

2.1 Web-based Music Information Retrieval

Determining terms related to a music artist via web-based MIR has first been addressed in [15], where Whitman and Lawrence extract different term sets (e.g. noun phrases and adjectives) from artist-related web pages. Based on term occurrences, individual term profiles are created for each artist. The authors then use the overlap between the term profiles of two artists as an estimate for their similarity. A quite similar approach is presented in [16]. Knees et al. however do not use specific term sets, but create a term list directly from the retrieved web pages. Subsequently, a term selection technique is

applied to filter out less important terms. Hereafter, the TF·IDF measure is used to weight the remaining words and subsequently create a weighted term profile for each artist. Knees et al. propose their approach for artist-to-genre classification and similarity measurement. Pampalk et al. in [17] use a dictionary of about 1,400 musically relevant terms to index artist-related web pages. Different term weighting techniques are applied to describe each artist with some terms. Furthermore, the artists are hierarchically structured using a GHSOM [18], a special version of the Self-Organizing Map [19]. The authors show that considering only the terms in the dictionary for term weighting and clustering outperforms using all terms found on the extracted web pages. An approach to assign descriptive terms to a given artist is presented in [1]. Schedl et al. use co-occurrences derived from artist-related web pages to estimate the conditional probability for the artist name under consideration to be found on a web page containing a specific descriptive term and vice versa. To this end, a set of predefined genres and other attributes, like preferred tempo or mood of the artist's performance, is used. The aforementioned probabilities are then calculated, and the most probable value of the attribute under consideration is assigned to the artist. Independent of Schedl et al., Geleijnse and Korst present in [20] an approach that differs from [1] only in regard to the normalization used.
The 3D-COB proposed here uses a dictionary similar to that in [17] to extract artist-related information from web pages. However, the clustering is performed in a very different way and on a different level (for individual web pages instead of artists).

2.2 Information Visualization of Hierarchical Data

Related work on visualizing hierarchical data primarily focuses on the *Sunburst* approach, as the 3D-COB extends the Sunburst in various aspects. The Sunburst (aka InterRing) as proposed in [12], [13], [14] is a circular, space-filling visualization technique. The center of the visualization represents the highest element in the hierarchy, whereas elements on deeper levels are illustrated by arcs further away from the center. Child elements are drawn within the angular borders of their parent, but at a more distant position from the center. In almost all publications related to the Sunburst, its usual application scenario is browsing the hierarchical tree structure of a file system. In this scenario, directories and files are represented by arcs whose sizes are proportional to the sizes of the respective directories/files. In the case of the 3D-COB, however, some *constraints for the size of the visualization* are necessary. Furthermore, the arc sizes are determined by the *term weighting function*, which is applied to select the most important terms (for the clustering). Moreover, the 3D-COB allows for *encoding an additional data dimension* in the height of each arc. This dimension is used to visualize the amount of multimedia content provided by the analyzed web pages. As three different types of multimedia content are taken into account (audio, image, and video), the Sunburst stack of the 3D-COB consists of three individual Sunburst visualizations. Other space-filling visualization techniques for hierarchically structured data include the *Treemap* [21] and the *Hyperbolic Browser* [22]. In contrast to the Sunburst, the Treemap uses a rectangular layout and displays elements further down in the hierarchy embedded in the rectangle of their parent element. The Sunburst, however, displays all elements that reside on the same hierarchy level on the same torus, which facilitates getting a quick overview of the hierarchy. The Hyperbolic Browser lays out the tree representing the hierarchical data on a hyperbolic plane, which has the nice property that the circumference of a circle grows exponentially with its radius, like the number of nodes in a tree with its depth. To project the visualization to the Euclidean space, a Poincaré mapping, e.g. [23], is then applied. This mapping is a conformal mapping that preserves angles, but distorts lines on the hyperbolic plane to arcs on the Euclidean unit disk. Having applied the conformal mapping, the root node of the tree is represented by the center of the unit disk, while nodes at deeper hierarchy levels perspectively diminish – the more, the farther away from the center they are located.

3 The Three-Dimensional Co-Occurrence Browser (3D-COB)

To get a first impression of the appearance of the 3D-COB user interface, the reader is invited to take a look at Figure 2. This figure shows a stack of three three-dimensional Sunburst visualizations created from 161 web pages of the band *Iron Maiden*. Details on the information gathering process, the creation of the visualization, and the user interaction possibilities are provided in the following subsections.
3D-COB has been implemented using the *processing* environment [24] and the *CoMIRVA* framework [25]. CoMIRVA already contained an implementation of the two-dimensional version of the Sunburst. We heavily extended this version by heaving it to the third dimension and incorporating various multimedia content in the visualization. To this end, we particularly had to elaborate new data representation and user interaction models.

3.1 Retrieval and Indexing

Given the name of an artist, we first query *Google* with the scheme *"artist name" +music +review* to obtain the URLs of up to 1,000 web pages related to the artist, whose content we then retrieve. Subsequently, a term analysis step is performed. To this end, we use a dictionary of musically relevant terms, which are searched in all web pages of every artist, yielding an inverted file index. For the conducted experiments, a manually compiled dictionary that resembles the one used in [17] was utilized. It was assembled using various sources such as *Wikipedia* [26], *allmusic.com*[9], *Yahoo! Directory* [27] and contains music genres and styles, instruments, moods, and other terms which are somehow related to music. We further added the names of all artists in the collection used for our experiments, which comprises 112 quite popular artists (14 genres, 8 artists each). Altogether, the dictionary contains 1,506 terms and can be downloaded from `http://www.cp.jku.at/people/schedl/music/cob_terms.txt`.
As for indexing the multimedia content of the web pages, we first extract a list of common file extensions for audio, image, and video files from *Wikipedia* [28]. We then search the HTML code of each web page for links to files whose file extension occur in one of the extracted lists. Finally, we store the URLs of the found multimedia files and the inverted file index gained by the term analysis in an XML data structure. An example excerpt of such an XML file describing a music band is given in Figure 1.

Fig. 1 Excerpt from the XML file describing the indexed web content of *Iron Maiden*.

3.2 Creation of the Visualization

Using the inverted index of the web pages of an artist X, we can easily extract subsets $S_{X,\{t_1,\dots,t_r\}}$ of the web page collection of X which have in common the occurrence of all terms $t_1, \dots, t_r$.
Starting with the entire set of web pages $S_{X,\{\}}$ retrieved for X, we use a term weighting function (e.g. document frequency, term frequency, TF·IDF) to select a maximum number N of terms with highest weight, which are used to create N subsets $S_{X,\{t_1\}}, \dots, S_{X,\{t_N\}}$ of the web-page-collection. These subsets are visualized as arcs $A_{X,\{t_1\}}, \dots, A_{X,\{t_N\}}$ around a centered cylinder which represents the root arc $A_{X,\{\}}$, and thus the entire set of web pages retrieved for X. The angular extent of each arc is proportional to the weight of the associated term t_i, e.g. to the number of web pages containing t_i when using document frequencies for term weighting. To avoid very small, thus hardly perceivable, arcs, we omit arcs whose angular extent is smaller than a fixed threshold E, measured in degrees. Furthermore, each arc is colored with respect to the relative weight of its corresponding term t_i (relative to the maximum weight among all terms). The term selection and the corresponding visualization step are recursively performed for all arcs, with a maximum R for the recursion depth. This eventually yields a complete Sunburst visualization, where each arc at a specific recursion depth r represents a set of web pages $S_{X,\{t_1,\dots,t_r\}}$ in which all terms $t_1, \dots, t_r$ co-occur.
As for representing the multimedia content found on the web pages, in each layer of the Sunburst stack, the amount of a specific category of multimedia files is depicted. To this end, we encode the relative number of audio, image, and video files in the height of the arcs (relative to the total number represented by the root node of the respective layer). For example, denoting the audio layer as L_A, the image layer as L_I, and the video layer as L_V and focusing on a fixed arc A, the height of A in L_I shows the relative number of image files contained in the web pages that are represented by arc A, the height of A in L_V illustrates the relative number of video files, and the height of A in L_A the relative number of audio files, cf. Figure 2. The corresponding multimedia files can easily be accessed via the user interface of 3D-COB.

3.3 User Interface and User Interaction

Figure 2 depicts a screenshot of 3D-COB's user interface for 161 web pages retrieved for *Iron Maiden*. The constraints were set to the following values: $N = 6$, $R = 8$, $E = 5.0$ (cf. Subsection 3.2). Document frequencies were used for term weighting. Each arc $A_{X,\{t_1,\dots,t_r\}}$ is labeled with the term t_r that subdivides the web pages represented by the arc's parent node $A_{X,\{t_1,\dots,t_{r-1}\}}$ into those containing t_r and those not containing t_r. Additionally, the weight of the term t_r is added in parentheses to the label of each arc $A_{X,\{t_1,\dots,t_r\}}$. The topmost layer illustrates the amount of video files found on the web pages, the middle one the amount of image files, and the lower one the amount of audio files. In the screenshot shown in Figure 2, the arc representing the web pages on which all of the terms "Iron Maiden", "guitar", and "metal" co-occur is selected. Since document frequencies were used for this screenshot, determining the exact number of web pages represented by a particular arc is easy: 74 out of the complete set of 161 web pages contain the mentioned terms.
User interaction is provided in several ways. First, the mouse can be used to *rotate* the Sunburst stack around the Y-axis, i.e. the vertical axis going through the root nodes of

all Sunbursts in the stack. We implemented this by moving the mouse in the horizontal direction while pressing an arbitrary mouse button. *Zooming* in/out (within predefined boundaries) is provided as well as *changing the inclination of the stack*, which is limited to angles between a front view and a bird's eye view. This function is supported by moving the mouse upwards/downwards while holding the right mouse button pressed. To select a particular arc, e.g. to access the multimedia content of the corresponding web pages, the arrow keys can be used to navigate in the hierarchy. Using the keys *arrow down* and *arrow up*, the hierarchy is browsed in a vertical manner. More precisely, with the *arrow down* key, the first child arc of the currently selected arc is chosen, while the *arrow up* key selects the parent of the currently selected arc. Using the keys *arrow left* and *arrow right*, the user can navigate within the elements on the same hierarchy level which are grouped by the selected parent arc. The currently selected arc is highlighted by means of drawing a white border around it and coloring its label in white. So are all previously selected arcs at higher hierarchy levels. This facilitates tracing the selection back to the root arc and quickly recognizing all co-occurring terms on the web pages represented by the selected arc.
In addition to the basic interaction capabilities described so far, the following functionalities are provided.

– Creating a new visualization based on the subset of web pages given by the selected arc.
– Restoring the original visualization that incorporates all web pages in its root node.
– Showing a list of web page URLs which are represented by the selected arc.
– Displaying and browsing a list of audio, image, or video files, which are found on the web pages of the currently selected arc.
– Opening the web pages or the available multimedia files represented by the selected arc.
– Toggling the data dimension encoded in the arcs' color between the number of web pages and the amount of multimedia data.

Alternative Visualization Strategies

The proposed visualization approach is obviously not the only solution to the problem. We were, for example, also thinking about presenting the labels only on the top Sunburst to overcome the redundancy of drawing the same labels on each Sunburst layer. However, this would necessitate connecting the corresponding arcs among the different layers by other means, e.g. by drawing auxiliary lines between them. As a result, users would probably be distracted more than they are when labeling each arc on each layer. Furthermore, displaying labels for each layer facilitates orientation when the user zooms in. If labels are only shown on the top layer instead, the user will no longer be able to see the label of the arcs when zooming in to one of the lower layers. It would be also possible to put each layer directly on top of its lower neighbor, thereby leaving no space between the layers. On the one hand, this would allow for a more compact representation of the whole stack. However, the user would have to spend a lot of time adjusting inclination and zooming factor in order to bring the desired arc to a well-perceivable position, in particular, if this arc does not reside on the top layer.
Using different colormaps for different layers is also an option. This would support distinguishing the individual data dimensions, e.g. by using bluish color tones for audio, greenish tones for images, and reddish tones for video.

Fig. 2 The user interface of 3D-COB for a web-page-collection of the band *Iron Maiden*.

4 Evaluation of the Term Weighting Functions

We experimented with three different term weighting functions (document frequency, term frequency, TF·IDF) for term selection in the Sunburst creation step, cf. Subsection 3.2. Given a set of web pages S of an artist, the *document frequency* DF_t of a term t is defined as the absolute number of pages in S on which t appears at least once. The *term frequency* TF_t of a term t is defined as the sum of all occurrences of t in S. The *term frequency · inverse document frequency* measure $TF \cdot IDF_t$ of t is calculated as $TF_t \cdot \ln \frac{|S|}{DF_t}$.
To assess the influence of the term weighting function on the quality of the hierarchical clustering, the hierarchical layout, and thus on the visualization of 3D-COB, we conducted a user study as detailed in the following.

4.1 Setup

For the user study, we chose a collection of 112 well-known artists (14 genres, 8 artists each). Indexing was performed as described in Subsection 3.1. To create the evaluation data, for each artist, we calculated on the complete set of his/her retrieved and indexed web pages, the 10 most important terms using each of the three term weighting functions. To avoid biasing of the results, for each artist, the 10 terms obtained by applying every weighting function were then merged. Hence, every participant was presented a list of 112 artist names and, for each name, a set of associated terms (as a mixture of the terms obtained by the three weighting functions). Since the authors had no a priori knowledge of which artists were known by which participant, the participants were told to evaluate only those artists they were familiar with. Their task was then to rate the associated terms with respect to their appropriateness for describing the artist or his/her music. To this end, they had to associate every term to one of the three classes *+ (good description)*, *– (bad description)*, and $\sim$ *(indifferent or not wrong, but not a description specific for the artist)*.
Due to time constraints, we had to limit the number of participants in the user study to five. Three were postgraduate students and two scientific staff, and all were from the computer science department. All participants were male and stated to listen to music often.

4.2 Results and Discussion

We received a total of 172 assessments for sets of terms assigned to a specific artist. 92 out of the 112 artists were covered. To analyze the results we calculated, for each artist and weighting function, the sum of all points obtained by the assessments. As for the mapping of classes to points, each term in class + contributes 1 point, each term in class – gives -1 point, and each term in class $\sim$ yields 0 points.
Summing up the points over all assessments of each artist, for the three term weighting functions, gives the results shown in the columns labeled *TF*, *DF*, and *TF·IDF* of Table 3. Only the 92 artists that were assessed at least once are depicted. The column

labeled Ass shows the number of assessments made, i.e. the number of test persons who evaluated the respective artist. The next three columns reveal, for each weighting function, the summed up ratings over all terms, in points. Since the performance of the term weighting functions is hardly comparable between different artists using the summed up points, columns TF_{avg}, DF_{avg}, and $TF{\cdot}IDF_{avg}$ illustrate the averaged scores, which are obtained by dividing the summed up points by the number of assessments. These averaged points reveal that the quality of the terms vary strongly between different artists. Nevertheless, it can be stated that, for most artists, the number of descriptive terms exceeds the number of the non-descriptive ones. To investigate the overall performance of the term weighting functions, the arithmetic mean of the averaged points over all artists were calculated. These were 2.22, 2.43, and 1.53 for TF, DF, and TF·IDF, respectively. Due to the performed mapping from classes to points, these values can be regarded as the average excess of the number of good terms over the number of bad terms. Hence, overall, the document frequency measure performed best, the term frequency second best, and the TF·IDF worst for this specific task of finding descriptive terms for a music artist based on a dictionary of musically relevant terms.

To test for the significance of the results, we performed Friedman's two-way analysis of variance [29], [30]. This test is similar to the two-way ANOVA, but does not assume a normal distribution of the data. It is hence a non-parametric test, and it requires related samples (ensured by the fact that for each artist all three measures were rated). The outcome of the test is summarized in Table 1. Due to the very low p value, we can

Table 1 Results of Friedman's test to assess the significance of the differences in the term weighting measures.

N	92
df	2
χ^2	16.640
p	0.00000236

state that the variance differences in the results are significant with a very high probability. To assess which term weighting measures produce significantly different results, we conducted pairwise comparison between the results given by the three weighting functions. To this end, we employed the Wilcoxon signed ranks test [31] and tested for a significance level of 0.01. The test showed that TF·IDF performed significantly worse than both TF and DF, whereas no significant difference could be made out between the results obtained using DF and those obtained using TF. This result is quite surprising as TF·IDF is a well-established term weighting measure and commonly used to describe text documents according to the vector space model, cf. [32]. A possible explanation for the worse performance of TF·IDF is that this measure assigns high weights to terms that are very specific for a certain artist (high TF and low DF), which is obviously a desired property when it comes to distinguish one artist from another, for example, in artist classification tasks where finding the most discriminative terms of an artist usually increases classification accuracy. In our application scenario, however, we aim at finding the most descriptive terms – not the most discriminative ones – for a given artist. This kind of terms seems to be better determined by the simple TF and DF measures.

The laborious task of combining and analyzing the different assessments of the participants in the user study further allowed the author to take a qualitative look at the terms. Although the majority of the terms was judged descriptive, some interesting flaws were discovered. First, the term "musical" occurred on quite a lot of web pages and was therefore often contained in the set of the top-ranked terms. However, no participant judged this term as descriptive for any artist. A similar observation could be made for the term "real". In this case, however, one participant stated that this is a term commonly used in the context of hip-hop music and may therefore be descriptive to some extent. Furthermore, the term "christmas" was associated occasionally to some artists. These associations seem quite random since none of the artists is known for his/her performance of Christmas carols. Another reason for erroneously assigning a term to an artist is terms that are part of artist, album, or song names, but are not suited well to describe the respective artist. Examples for this problem category are "infinite" for the band *Smashing Pumpkins* and "human" as well as "punk" for the band *Daft Punk*.

5 Evaluating the User Interface

To investigate the usefulness of 3D-COB for gaining a quick overview of a set of artist-related web pages and efficiently browsing within this set, we conducted a second user study that primarily focuses on ergonomic aspects of 3D-COB. To this end, we employed a task-oriented evaluation scheme. We measured task completion times as well as correctness and also asked the participants for their satisfaction with the interface and possible suggestions for improvement. A comprehensive elaboration on different approaches to evaluating user interfaces can be found, for example, in [33].

5.1 Setup

We formulated the following tasks, which we believe are important for the mentioned purposes, and evaluated them in a quantitative manner:

1. Which are the five top-ranked terms that occur on the web pages mentioning "Iron Maiden"?
2. Indicate the number of web pages containing all of the terms "Iron Maiden", "metal", and "guitar".
3. Show a list of web pages that contain the terms "Iron Maiden" and "british".
4. Considering the complete set of web pages, which are the three terms that co-occur on the highest number of web pages?
5. How many web pages contain the terms "Iron Maiden" and "metal", but not the term "guitar"?
6. Display a list of audio files available at web pages containing the term "Iron Maiden".
7. Which terms co-occur on the set of web pages that contains the highest number of image files in hierarchy level three?
8. Indicate the URL of one particular web page that contains image files but no video files.
9. How many web pages does the complete collection contain?
10. Find one of the deepest elements in the hierarchy and select it.

11. Generate a new visualization using only the web pages on which the terms "bass" and "heavy metal" co-occur.

The tasks 1–8 are general ones that are likely to arise when analyzing and browsing collections of web pages. In particular, tasks 1–5 address the co-occurring terms, whereas tasks 6–8 deal with the multimedia content extracted from the web pages. In contrast, the tasks 9–11 relate to the structure of the Sunburst tree.
After having explained the interaction functionalities provided by 3D-COB to our participants, they had five minutes to explore the user interface themselves with a visualization gained for *Britney Spears*. During this warm-up, the participants were allowed to ask questions. After the exploration phase, we presented them the visualization obtained when using the web page collection of *Iron Maiden*, cf. Figure 2. We consecutively asked them each of the questions and measured the time they needed to finish each task. Each participant had a maximum time of three minutes to complete each task. The constraints were set as follows: $N = 8$, $R = 8$, and $E = 3.0$ (cf. Subsection 3.2).
Due to time limitations, we had to restrict the number of participants in the user study to six (five males, one female). All of them were computer science or business students at the *Johannes Kepler University Linz* and all stated to have a moderate or good knowledge of user interfaces and to be very interested in music. All participants performed the user study individually, one after another. The experiments were carried out on a *Pentium 4* 3GHz with 2GB RAM and a *nVidia GeForce 6600 GT* graphics card running *ubuntu Linux*.

5.2 Results and Discussion

As for the results of the study, Table 2 shows the time, in seconds, needed by the participants (A–F) to finish each task. In general, the tasks related to structural questions were answered in a shorter time than those related to browsing the collection. Among the structural questions, solely task 11 required a quite high average time. This is explained by the fact that the term "bass" was not easy to find on all layers. The same holds for the term "british" requested in task 3.
For the questions related to browsing in the hierarchy, it was observed that tasks requiring intensive rotation of the Sunburst stack (1, 3, 4, 5, 7) yielded worse results than those for which this was not the case (2, 6). In general, users spent a lot of time rotating the Sunburst stack to a position at which the label of the selected arc was readable. This process will be automatized in future versions of 3D-COB by providing an option to lock the position of the selected arc where it is well perceivable.
The relatively high average time required to perform the first task may be attributed to the fact that most participants needed some seconds to get used to the new visualization of *Iron Maiden* after having explored the web pages of *Britney Spears* in the exploration phase. Task 2 was successfully finished quite fast (in 11 seconds on average) by all participants. This may be explained by the fact that the combination of the terms "Iron Maiden", "metal", and "guitar" was one of only two term combinations that formed a third hierarchy level in the visualization. In spite of the fact that task 3 was solved in only 37 seconds on average, we realized that some participants had problems locating the arc "british" since it was hardly perceivable due to its position behind a much higher arc. As both task 4 and 2 required finding the same arc, it was quite interesting that the averaged times differed considerably. As for task 5, two participants

were not sure which number to subtract from which other. Except for one participant, who chose a correct but time-consuming solution, task 6 was generally solved quickly. Solving task 7 took the second highest average time since it required finding and navigating to the Sunburst that illustrates the amount of image files and comparing the heights of all arcs in hierarchy level three of this Sunburst. Task 8 yielded the worst results as no arc on the video layer had a height of zero, which confused most of the participants. It was obviously not clear that a positive height of an arc on the video layer does not necessarily mean that each web page represented by this arc offers video content.

To conclude, the user study assessing ergonomic aspects showed that 3D-COB can be efficiently used for tasks related to browsing sets of web pages. Although barely comparable to the user study on similar tree visualization systems conducted in [34], due to a different application scenario, a very rough comparison of the average total performance times for the tasks shows that this time is much shorter for 3D-COB (45 sec) than for the best performing system of [34] (101 sec). Therefore, our results seem promising. As for user satisfaction, in general, users liked the idea of the 3D-COB to browse sets of web pages via important terms and multimedia content. However, it must be stated that the user interaction functionalities provided by 3D-COB need some improvements. In particular, users criticized that rotating the Sunburst stack was required too often when performing the tasks. Some participants were also irritated by the redundancy among arcs. According to the used algorithm, cf. Subsection 3.2, the pages that contain both terms “guitar” and “metal” can be reached by selecting “metal” in the first torus (tree node) and then “guitar” in the second one. Alternatively, the same set of web pages can be reached by first selecting “guitar” and subsequently selecting “metal”. Moreover, two participants disliked that the multimedia content opened in an external web browser. They suggested to incorporate viewers and players for the multimedia content directly into the user interface. One user requested to define the search term(s) by himself. For example, he wanted to be able to create a visualization containing the web pages that mention both “Britney Spears” and “lovers”. This user also suggested to highlight arcs that are labeled with other artist names (and maybe link them to the 3D-COB describing the corresponding artist). We will investigate if and how these suggestions can be addressed in future versions of the implementation.

Table 2 For each participant, the time (in seconds) needed to finish each task of the user study on ergonomic aspects. Inverse numbers indicate that the given answer was wrong. The mean was calculated excluding the wrong answers.

Task	1	2	3	4	5	6	7	8	9	10	11
A	28	13	45	47	36	61	172	*180*	2	12	*25*
B	69	23	46	52	14	15	68	76	*6*	12	62
C	15	3	39	27	22	3	34	68	1	9	31
D	132	1	57	30	117	14	43	*180*	5	12	40
E	110	9	16	8	*163*	7	12	148	2	38	74
F	36	14	21	46	44	12	79	*180*	3	5	61
Mean	**65**	**11**	**37**	**35**	**47**	**19**	**68**	**97**	**3**	**15**	**54**

6 Conclusions and Future Work

In this paper, we presented the *Three-Dimensional Co-Occurrence Browser (3D-COB)*, a user interface for browsing collections of music artist-related web pages in a novel way. 3D-COB automatically extracts musically relevant terms from web pages about artists, applies a term weighting function, organizes the web pages according to co-occurring terms, and finally employs a variant of the Sunburst visualization technique to illustrate not only the extracted terms, but also the amount of multimedia files, grouped in different categories (audio, image, video).
Moreover, we conducted two user studies: one to evaluate the performance of different term weighting strategies for finding descriptive artist terms, the second to assess ergonomic aspects of 3D-COB's user interface. From the first study, we learned that using TF·IDF yielded significantly worse results than the simple TF and DF measures with respect to the appropriateness of the terms to describe the music artists used in our experiments. In contrast, comparing the measures TF and DF, no significant difference in their performance was detected. The second user study showed that 3D-COB offers valuable additional information about web pages that cannot be discovered by the standard list-based representation of search results, which is commonly used by web search engines.
The proposed user interface is not limited to the music domain. It can be used indeed for visualizing and browsing all kinds of hierarchically organized data, where every element in the hierarchy is assigned a set of attributes. Other possible application areas include the one proposed in [13] of illustrating the hierarchical tree structure of a file system. In this case, the attributes assigned to each element (file/directory) may be *time elapsed since the last change of the file* or *number of file accesses*. Furthermore, the proposed technique may be employed to visualize the product hierarchy of (web) shops offering a large range of products or of online auction systems like *eBay* [35]. In this case, one Sunburst layer for each of the attributes *price of the product* or *current bid*, *time remaining until the end of the auction*, *different feedback levels (positive, neutral, negative)* or *distance from the item location to the user's own domicile* may be included in the visualization. The proposed technique could also be applied to domains like movies, literature, and news.
As for future work, we will elaborate alternative ways to navigate in the visualization, e.g. using alternative input devices. We are also developing a focused web crawler in combination of which 3D-COB may be used to browse a set of web pages related to a certain topic (not necessarily related to music) without relying on existing search engines. Moreover, we will improve the user feedback provided by 3D-COB, e.g. by showing all terms that co-occur in the selected set of web pages, independently from the labels of the arcs. Automatically rotating the Sunburst stack to display the selected arc at a position at which its label is readable well would certainly also improve the usability of the interface. Finally, smoothly embedding the multimedia content directly in the user interface instead of opening it in external applications would be a desirable feature for future versions of 3D-COB.

Acknowledgements This research is supported by the Austrian Fonds zur Förderung der Wissenschaftlichen Forschung (FWF) under project number L511-N15.

Table 3 Results of the user study on different term weighting functions.

Artist	Ass	TF	DF	TF·IDF	TF_{avg}	DF_{avg}	$TF{\cdot}IDF_{avg}$
50 Cent	3	17	16	19	5.67	5.33	6.33
ABBA	3	10	11	5	3.33	3.67	1.67
Al Green	1	-2	0	-4	-2.00	0.00	-4.00
Alice Cooper	3	8	5	1	2.67	1.67	0.33
Alice in Chains	2	10	12	7	5.00	6.00	3.50
Alpha Blondie	1	-10	-8	-8	-10.00	-8.00	-8.00
Anthrax	2	6	9	5	3.00	4.50	2.50
Antonin Dvorak	2	6	9	9	3.00	4.50	4.50
Aphex Twin	2	13	13	9	6.50	6.50	4.50
Aretha Franklin	3	9	8	9	3.00	2.67	3.00
Bad Religion	3	4	17	8	1.33	5.67	2.67
Basement Jaxx	1	7	8	7	7.00	8.00	7.00
BB King	3	-1	0	-1	-0.33	0.00	-0.33
Beck	3	-4	-6	0	-1.33	-2.00	0.00
Belle & Sebastian	2	-1	-3	-2	-0.50	-1.50	-1.00
Big Bill Broonzy	1	4	4	3	4.00	4.00	3.00
Billie Holiday	2	9	8	7	4.50	4.00	3.50
Black Sabbath	3	10	10	11	3.33	3.33	3.67
Bob Dylan	3	4	8	10	1.33	2.67	3.33
Bob Marley	3	-5	-3	1	-1.67	-1.00	0.33
Britney Spears	3	10	18	15	3.33	6.00	5.00
Carl Cox	1	8	7	8	8.00	7.00	8.00
Chemical Brothers	3	5	8	6	1.67	2.67	2.00
Chuck Berry	1	1	1	3	1.00	1.00	3.00
Cypress Hill	2	6	2	6	3.00	1.00	3.00
Daft Punk	2	6	9	3	3.00	4.50	1.50
Dave Brubeck	2	5	4	1	2.50	2.00	0.50
Dead Kennedys	1	5	6	4	5.00	6.00	4.00
Deep Purple	3	6	7	3	2.00	2.33	1.00
Dixie Chicks	1	6	5	6	6.00	5.00	6.00
Django Reinhardt	2	9	9	8	4.50	4.50	4.00
Dolly Parton	1	4	4	1	4.00	4.00	1.00
Dr. Dre	2	11	12	3	5.50	6.00	1.50
Duke Ellington	3	11	10	5	3.67	3.33	1.67
Elvis Presley	4	-3	-4	-5	-0.75	-1.00	-1.25
Eminem	4	22	15	15	5.50	3.75	3.75
Faith Hill	1	4	4	2	4.00	4.00	2.00
Fatboy Slim	2	5	6	1	2.50	3.00	0.50
Frederic Chopin	3	4	-1	0	1.33	-0.33	0.00
Garth Brooks	1	3	3	2	3.00	3.00	2.00
Glenn Miller	1	0	0	0	0.00	0.00	0.00
Grandmaster Flash	1	1	3	3	1.00	3.00	3.00
Hank Williams	1	4	3	2	4.00	3.00	2.00
Howlin' Wolf	1	1	1	-2	1.00	1.00	-2.00
Iron Maiden	3	10	11	11	3.33	3.67	3.67
James Brown	2	-1	1	-1	-0.50	0.50	-0.50
Janet Jackson	2	3	5	1	1.50	2.50	0.50
Jimmy Cliff	1	-1	-2	1	-1.00	-2.00	1.00
Joan Baez	1	7	7	5	7.00	7.00	5.00
J. S. Bach	1	4	4	4	4.00	4.00	4.00
Johannes Brahms	2	11	11	11	5.50	5.50	5.50
John Lee Hooker	1	0	0	2	0.00	0.00	2.00
John Mayall	1	-1	-1	-3	-1.00	-1.00	-3.00
Johnny Cash	2	11	11	7	5.50	5.50	3.50
Justin Timberlake	3	-2	0	-2	-0.67	0.00	-0.67

Artist	Ass	TF	DF	TF·IDF	TF_{avg}	DF_{avg}	$TF{\cdot}IDF_{avg}$
Kraftwerk	1	6	4	2	6.00	4.00	2.00
Little Richard	2	-3	-1	-3	-1.50	-0.50	-1.50
Louis Armstrong	2	-3	-4	-3	-1.50	-2.00	-1.50
L. van Beethoven	1	5	6	1	5.00	6.00	1.00
Madonna	3	13	6	7	4.33	2.00	2.33
Marvin Gaye	1	3	4	0	3.00	4.00	0.00
Megadeth	1	0	3	-2	0.00	3.00	-2.00
Michael Jackson	2	-9	-9	-10	-4.50	-4.50	-5.00
Miles Davis	1	-2	-3	0	-2.00	-3.00	0.00
Missy Elliot	2	9	11	11	4.50	5.50	5.50
Moloko	2	11	9	7	5.50	4.50	3.50
Muddy Waters	1	0	-2	-2	0.00	-2.00	-2.00
N'Sync	4	5	6	4	1.25	1.50	1.00
Nirvana	1	1	0	3	1.00	0.00	3.00
NoFX	2	15	15	-6	7.50	7.50	-3.00
Patti Smith	1	1	4	4	1.00	4.00	4.00
Prince	2	-1	-1	1	-0.50	-0.50	0.50
Public Enemy	2	10	12	7	5.00	6.00	3.50
Radiohead	1	6	6	6	6.00	6.00	6.00
Ramones	1	3	6	-1	3.00	6.00	-1.00
Run DMC	3	9	1	1	3.00	0.33	0.33
Sepultura	2	11	5	4	5.50	2.50	2.00
Sex Pistols	2	6	8	4	3.00	4.00	2.00
Shaggy	2	3	-2	3	1.50	-1.00	1.50
Sid Vicious	1	-1	1	1	-1.00	1.00	1.00
Slayer	1	-2	0	-3	-2.00	0.00	-3.00
Smashing Pumpkins	2	-2	-2	-2	-1.00	-1.00	-1.00
Solomon Burke	1	2	2	3	2.00	2.00	3.00
Sonic Youth	1	4	7	5	4.00	7.00	5.00
Suzanne Vega	2	4	6	2	2.00	3.00	1.00
The Animals	1	-4	-4	-4	-4.00	-4.00	-4.00
The Clash	1	2	0	-2	2.00	0.00	-2.00
The Kinks	1	1	0	1	1.00	0.00	1.00
The Rolling Stones	4	-1	5	-3	-0.25	1.25	-0.75
Tracy Chapman	1	2	4	1	2.00	4.00	1.00
W. A. Mozart	2	12	12	8	6.00	6.00	4.00
Ziggy Marley	1	1	1	4	1.00	1.00	4.00
Sum	**172**	**386**	**413**	**271**	**2.22**	**2.43**	**1.53**

References

1. Schedl, M., Pohle, T., Knees, P., Widmer, G.: Assigning and Visualizing Music Genres by Web-based Co-Occurrence Analysis. In: Proceedings of the 7th International Conference on Music Information Retrieval. Victoria, Canada (2006)
2. Zadel, M., Fujinaga, I.: Web Services for Music Information Retrieval. In: Proceedings of the 5th International Symposium on Music Information Retrieval. Barcelona, Spain (2004)
3. Knees, P., Schedl, M., Pohle, T., Widmer, G.: An Innovative Three-Dimensional User Interface for Exploring Music Collections Enriched with Meta-Information from the Web. In: Proceedings of the 14th ACM Conference on Multimedia 2006. Santa Barbara, CA, USA (2006)
4. Pampalk, E., Goto, M.: MusicSun: A New Approach to Artist Recommendation. In: Proceedings of the 8th International Conference on Music Information Retrieval. Vienna, Austria (2007)
5. Goto, M., Goto, T.: Musicream: New Music Playback Interface for Streaming, Sticking, Sorting, and Recalling Musical Pieces. In: Proceedings of the 6th International Conference on Music Information Retrieval. London, UK (2005)

6. Mörchen, F., Ultsch, A., Nöcker, M., Stamm, C.: Databionic Visualization of Music Collections According to Perceptual Distance. In: Proceedings of the 6th International Conference on Music Information Retrieval. London, UK (2005)
7. Vignoli, F., van Gulik, R., van de Wetering, H.: Mapping Music in the Palm of Your Hand, Explore and Discover Your Collection. In: Proceedings of the 5th International Symposium on Music Information Retrieval. Barcelona, Spain (2004)
8. Eck, D., Bertin-Mahieux, T., Lamere, P.: Autotagging Music Using Supervised Machine Learning. In: Proceedings of the 8th International Conference on Music Information Retrieval. Vienna, Austria (2007)
9. http://www.allmusic.com (access: 11-2007) (2007)
10. http://last.fm (access: 12-2007) (2007)
11. http://finetunes.musiclens.de (access: 03- 2008) (2008)
12. Andrews, K., Heidegger, H.: Information Slices: Visualising and Exploring Large Hierarchies using Cascading, Semi-Circular Discs. In: Proceedings of IEEE Information Visualization 1998. Research Triangle Park, NC, USA (1998)
13. Stasko, J., Zhang, E.: Focus+Context Display and Navigation Techniques for Enhancing Radial, Space-Filling Hierarchy Visualizations. In: Proceedings of IEEE Information Visualization 2000. Salt Lake City, UT, USA (2000)
14. Yang, J., Ward, M.O., Rundensteiner, E.A.: InterRing: An Interactive Tool for Visually Navigating and Manipulating Hierarchical Structures. In: Proceedings of the IEEE Symposium on Information Visualization 2002. Boston, MA, USA (2002)
15. Whitman, B., Lawrence, S.: Inferring Descriptions and Similarity for Music from Community Metadata. In: Proceedings of the 2002 International Computer Music Conference. Göteborg, Sweden (2002)
16. Knees, P., Pampalk, E., Widmer, G.: Artist Classification with Web-based Data. In: Proceedings of the 5th International Symposium on Music Information Retrieval. Barcelona, Spain (2004)
17. Pampalk, E., Flexer, A., Widmer, G.: Hierarchical Organization and Description of Music Collections at the Artist Level. In: Proceedings of the 9th European Conference on Research and Advanced Technology for Digital Libraries. Vienna, Austria (2005)
18. Dittenbach, M., Rauber, A., Merkl, D.: Uncovering Hierarchical Structure in Data Using the Growing Hierarchical Self-Organizing Map. Neurocomputing **48**(1–4), 199–216 (2002)
19. Kohonen, T.: Self-Organizing Maps, *Springer Series in Information Sciences*, vol. 30, 3rd edn. Springer, Berlin, Germany (2001)
20. Geleijnse, G., Korst, J.: Web-based Artist Categorization. In: Proceedings of the 7th International Conference on Music Information Retrieval. Victoria, Canada (2006)
21. Johnson, B., Shneiderman, B.: Tree-Maps: A Space-Filling Approach to the Visualization of Hierarchical Information Structures. In: Proceedings of the 2nd IEEE Conference on Visualization. San Diego, CA, USA (1991)
22. Lamping, J., Rao, R., Pirolli, P.: A Focus+Context Technique Based on Hyperbolic Geometry for Visualizing Large Hierarchies. In: CHI'95: Proceedings of the ACM SIGCHI Conference on Human factors in Computing Systems, pp. 401–408. Denver, CO, USA (1995)
23. Walter, J.A., Ritter, H.: On Interactive Visualization of High-Dimensional Data Using the Hyperbolic Plane. In: Proceedings of the 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 123–131. ACM Press, Edmonton, Canada (2002)
24. http://www.processing.org (access: 03-2008) (2008)
25. Schedl, M., Knees, P., Seyerlehner, K., Pohle, T.: The CoMIRVA Toolkit for Visualizing Music-Related Data. In: Proceedings of the 9th Eurographics/IEEE VGTC Symposium on Visualization. Norrköping, Sweden (2007)
26. http://www.wikipedia.org (access: 12-2007) (2007)
27. http://dir.yahoo.com/Entertainment/Music/Genres (access: 03-2008) (2008)
28. http://en.wikipedia.org/wiki/List_of_file_formats (access: 12-2007) (2007)
29. Friedman, M.: A Comparison of Alternative Tests of Significance for the Problem of m Rankings. The Annals of Mathematical Statistics **11**(1) (1940)
30. Sheskin, D.J.: Handbook of Parametric and Nonparametric Statistical Procedures, 3rd edn. Chapman & Hall/CRC, Boca Raton (2004)
31. Wilcoxon, F.: Individual Comparisons by Ranking Methods. Biometrics Bulletin **1**(6), 80–83 (1945)

32. Salton, G., Wong, A., Yang, C.S.: A Vector Space Model for Automatic Indexing. Communications of the ACM **18**(11), 613–620 (1975)
33. Stone, D., Jarrett, C., Woodroffe, M., Minocha, S.: User Interface Design and Evaluation. Morgan Kaufmann (2005)
34. Kobsa, A.: User Experiments with Tree Visualization Systems. In: Proceedings of the 10th IEEE Symposium on Information Visualization 2004. Austin, TX, USA (2004)
35. http://www.ebay.com (access: 07-2009) (2009)