Vocalmetrics: An interactive software for visualization and classification of music

Felix Schönfeld
Faculty of Computer Science
Technische Universität
Dresden
Dresden, Germany
felix.schoenfeld@mailbox.tu-dresden.de

Axel Berndt
Faculty of Computer Science
Technische Universität
Dresden
Dresden, Germany
axel.berndt@tu-dresden.de

Tilo Hähnel
Department of Musicology
Franz Liszt School of Music
Weimar
Weimar, Germany
tilo.haehnel@hfm-weimar.de

Martin Pfleiderer
Department of Musicology
Franz Liszt School of Music
Weimar
Weimar, Germany
martin.pfleiderer@hfm-weimar.de

Rainer Groh
Faculty of Computer Science
Technische Universität
Dresden
Dresden, Germany
rainer.groh@tu-dresden.de

ABSTRACT
Vocalmetrics is an interactive software tool that provides scientific techniques for interactive visualization and classification of musical data. The application supports the classification of music data as a pivotal aim of music education and analysis. The paper, in particular, introduces Vocalmetrics' *prototype semantics* and the *egg cell metaphor*. The former provides an intuitive and playful approach for exploring and classifying multidimensional musical data, whereas the latter is a direct manipulative interaction technique for rating features of musical data, particularly suitable for subjective assessments.

Categories and Subject Descriptors
H.5.2 [**Multimedia Information Systems**]: Multimedia Information Systems; J.5 [**Arts and Humanities**]

General Terms
Information Visualization, Interaction, Visual Analytics, Music Education, Musicology

1. INTRODUCTION
Today, talking about music is as ubiquitous as music itself. However, communicating about music, e.g., how a certain piece of music, musician or singer sounds, is still a non-trivial and challenging task for musical laymen as well as for music experts like music teachers or musicologists. Music can be classified not only according to style, time of creation etc.,

but also according to musical features like structure, sound, and expression. Therefore, learning to classify music pieces and artists in this manner is a pivotal aim of music education as well as an essential requirement for music expertise. Ratings and classifications can serve as a starting point for discussing music, its features and effects in classroom or university seminars as well as for comparing different music repertoires, artists, or styles in music research. Moreover, it is a general human demand and capacity to classify music in order to compare different pieces, artists, and styles—a capacity that is challenged today by the ubiquity and overwhelming amount of music distributed as digital audio files. For a long time content-based music information retrieval has been searching for strategies of automatic annotation of music according to features like form, meter, harmony, or sound [3]. However, it remains a difficult task to describe in detail, e.g., the differences in sound between musical instruments or the idiosyncratic ways a musician plays an instrument and a singer sings a song. In particular, describing singing styles is a very demanding and intricate task. Furthermore, music could also be rated according to more subjective features like personal preference and liking, expressed emotion, or mood.

In this paper, the conceptual design and implementation of the interactive software tool Vocalmetrics will be described. Vocalmetrics has initially been developed as a web application to visualize audio sample datasets of vocal recordings.[1] These samples were rated according to nine dimensions of vocal expression in order to show relationships between song excerpts, singers, and their ratings. In Section 2 the theoretical background and conceptual design of this first version, called Vocalmetrics v1.0, as well as its problems and shortcomings are outlined and situated in the context of computational visualization techniques in general. Then, the actual version, called Vocalmetrics v1.1, will be described in detail starting with its tools for visualization, followed by

---

[1] www.hfm-weimar.de/popvoices/vocalmetrics/main.htm

the concepts of project maintenance, and finally the facilities it offers to rate audio excerpts with the help of *prototype interaction technique* and the so-called *egg cell interaction technique*. Finally, implications of Vocalmetrics for musicological research and applications in music education are discussed.

2. BACKGROUND

Since 2011, the research project "Voice and singing in popular music in the U.S.A. (1900?1960)" investigates vocal expression with respect to different genres and stereotypes of class, race, gender, religion, and region. Besides the traditional ways of publication, the project aims at publishing some results to a non-scientific audience as well as at providing an overview over important means of vocal expression and their relationship to history, genre and so forth. This has been the starting point of Vocalmetrics. Vocalmetrics v1.0 contains a database of over 200 examples for vocal expression which any user interested in vocal expression in popular music can explore interactively. Vocalmetrics v1.0 provides some research findings, a quick overview over the collection, a summary of parameters relevant to describe vocal expression in popular music, and is capable of communicating the results to scientists as well as non-scientists. To this end, a list of representative singers of a number of different genres had been compiled. Afterwards, representative samples were used to capture representative phrases of these singers. The audio excerpts should represent typical personal styles, timbres, or other vocal features within a time-frame of ten to twenty seconds. They were selected to cover an artist's multi-faceted singing and vocal development over a longer period while providing high sound quality and a clear audibility of the singing. Each audio sample was supplied with a spectrogram view and a rating of several features of vocal expression as well as conventional meta data such as genre, name, year of recording, etc. (see Table 1). In section 3.2 the basic visualization tools of Vocalmetrics v1.0 that are still implemented in the new Vocalmetrics v1.1 will be described in detail followed by a description of the recent enhancements. In the following paragraph we focus on the rating process as it was conceived in Vocalmetrics v1.0 and its shortcomings which lead to the development of new ratings procedures (see section 3.3).

Initially, the rating procedure was inspired by the work of Alan Lomax and his *Cantometrics* approach [5, 6, 7]. Lomax defined 37 features to describe vocal expression in music in order to reveal interdependences between the singing style of a culture and its social structure. The 37 features cover aspects like number of singers and audience setting, rhythmic blend of the vocal group, melodic form, position of the final tone, and the degree of embellishment used by singers. In our research project nine dimensions of vocal expression proved to be of particular relevance: vibrato, glissando, intensity, roughness, breathiness, vocal register crossing, articulation, tempo rubato, and offbeat frequency. Some of the dimensions combined ratings of intensity and frequency of occurence of a vocal feature. Both intensity and frequency were rated on a five-step scale from very weak respectively very rare or not existing to very strong or very frequent. This very coarse discretization is extended to a (quasi-continuous) percentage scale during the further developments of Vocalmetrics.

Rating was preferred to an automatic feature extraction, for there is still no way to compute the subjective impression and rating of a human listener, even if the feature is technically well understood. The vocal technique of the vibrato serves as a good example to illustrate this discrepancy: technically, a vibrato is roughly a sinusoidal fluctuation of pitch and loudness (and timbre, as Seashore showed in the 1940s [10]), each of which has a modulation frequency and a modulation depth component which both contribute to the subjective intensity of the vibrato as a whole. Additionally, the way a singer starts and ends the vibrato—whether the vibrato is just on or off, or increases in its intensity over the course of the tone—influences the listener's subjective impression of the vibrato's intensity. Therefore, a listener's impression of a single feature depends on a large number of characteristics. More importantly, however, the relationship between all known components is not yet fully understood, and there might even be further components still unknown. Finally, the weighting of the components might depend on the listeners' individual preferences and listening biography, and the musical and singer's context, in which the feature in question is embedded.

The nine rating dimensions were derived from music analyses and the basic protocol for functional assessment of voice pathology [1, 4, 8]. As these sources indicate subjective ratings have a low inter-rater reliability but a quite high intra-rater reliability, which means that different listeners rate a vocal feature differently, but the rating of a vocal feature is constant over a long period of time for each listener. Since all raters rated all samples, the position of the samples between the extremes became similar, which finally increased the inter-rater reliability. Rating more than 200 samples for each of the rating dimension is a very time-consuming matter. An alternative, less time-consuming solution would be a pair-wise comparison that simplifies the rating process, e. g., listeners would be able to refer to existing ratings and possibly adapt and refine them. Then, they have to be provided automatically with the right samples. This became one of the objectives of the application Vocalmetrics v1.1 (see Section 3.3).

The visualizations of Vocalmetrics v1.0 (see Section 3.2) constitute a successful solution for exploring the results of the research project on vocal expression. The software is easy to use even for musical laymen and serves as a comfortable tool for data exloration as well as for presentational and educational purposes. At the same time it raised interest to use and enhance the provided tools to expand the database and assist users in the rating process of new audio samples. In particular, the following shortcomings were the starting point for a further development of the Vocalmetrics software: Firstly, the process of expanding the database is complex, time-consuming, inconsistent, and error-prone. Secondly, during the rating process the raters have no relation to existing ratings. Thirdly, exploration techniques are rather scientific (scatter plot and star plot) and partly rather difficult to be grasped intuitively by laymen, students, or pupils.

Therefore, the new software has to meet the follwing requirements: The rating process has to be intuitive, interactive, and playful—as an alternative to the direct numeric input.

	feature	data type	example
meta data	artist	String	Little Richard
	title	String	rip it up
	gender	[f,m]	m
	label	String	Specialty
	genre	String	Rock'n'Roll
files	audio sample	audio file	LR3.ogg
	spectrogram	file	LR3.png
	pdf	file	LR3.pdf
rating dimensions	vibrato	[0% ... 100%]	0
	glissando	[0% ... 100%]	32
	intensity	[0% ... 100%]	67
	roughness	[0% ... 100%]	95
	breathiness	[0% ... 100%]	25
	register	[0% ... 100%]	31
	articulation	[0% ... 100%]	67
	rubato	[0% ... 100%]	8
	off-beat	[0% ... 100%]	58

Table 1: Data model for one data record in the "Voice and Singing" database.

In order to explore individual rating differences, e.g., in regard to musical effects on emotion and mood, the individual rating data of each rater should be visualized (and exported for further investigation, too). Furthermore, the software should be applicable for any musical project with any chosen rating dimensions and meta data type. Finally, there is a need for exploration techniques which are easier to understand and also suitable for laymen.

3. VOCALMETRICS

Vocalmetrics v1.1 is a software tool for interactive visualization and classification of musical data. It is suitable for a quick as well as detailed analysis of musical datasets done by music experts and also for presentational and educational purposes with regard to non-specialists. The *prototype visualization technique* (introduced in Section 3.2) allows for the classification of objects, especially where clear criteria for class boundaries are hard to specify or are even missing. The *egg cell metaphor* (introduced in Section 3.3) was developed as a visual interaction technique that supports the user when rating psycho-acoustically biased musical features. Additionally, Sections 3.1 and 3.3 include more general considerations concerning prototype theory.

This Section provides an overview of the functionalities of Vocalmetrics v1.1. It distinguishes between general characteristics, visualization techniques, and the rating process. Some details about the technical implementation complete this Section.

3.1 General characteristics

Vocalmetrics v1.1 facilitates multiple projects (different databases) and multiple users with different roles (admin, user, guest). It provides the handling of data with a multidimensional feature space. Each data record in the database represents a music excerpt, as shown in Table 1. In addition to the audio file itself, each data record can include other features like meta data (e.g., artist, title, genre etc.) and
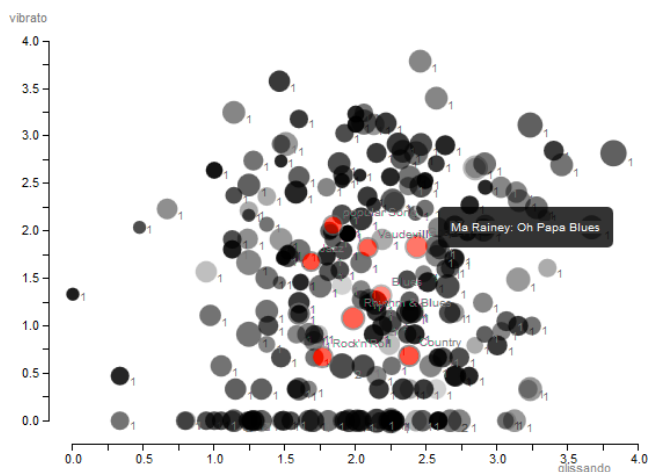

Figure 1: Vocalmetrics' *scatter view*. Red circles show average values of different musical genres. The circles are positioned within a coordinate system (y: vibrato, x: glissando).

the particular rating dimensions chosen for a certain project (e.g., vibrato, glissando etc.). Generally, the features of a project (database columns) are of one of the following data types: nominal (character string), ordinal (number or enumeration), or file reference.

3.2 Visualization

As a visualization tool Vocalmetrics implements a variety of geometric and iconic techniques encapsulated in four different views. A basic idea within the design process was to represent each data record by a geometrical circle so that the user is always aware of the amount of data and is motivated to interact with the physically attractive object (affordance). Each circle can be selected by the user. Then, all details of the corresponding data record (meta data, rating values of the dimensions etc.) are shown in the upper area of the screen. The software includes an audio player to play back the attached audio sample. Four different views are offered. While the *scatter view*, *timeline view*, and *star plot view* were already part of Vocalmetrics v1.0., the *prototype view* is a new feature of Vocalmetrics v1.1.

The *scatter view*, demonstrated in Figure 1, is a two-dimensional scatter plot enriched with some interactive functionalities. The position of a data record depends on two of its features that are associated with the x- and y-coordinates. Besides those two, a circle's size and transparency can encode two additional features. The view can therefore visualize up to four different features of an audio sample at the same time. The assignment between visual variable and data feature can be customized by the user, causing each data record to change its position, implemented as an animated transition. Additionally, aggregated information, like average values per feature, can be added as extra circles and are color coded for better differentiation.

The *time line view*, see Figure 2, gives an overview of the average values of each feature per year. This facilitates the recognition of long-term developments in musical style over
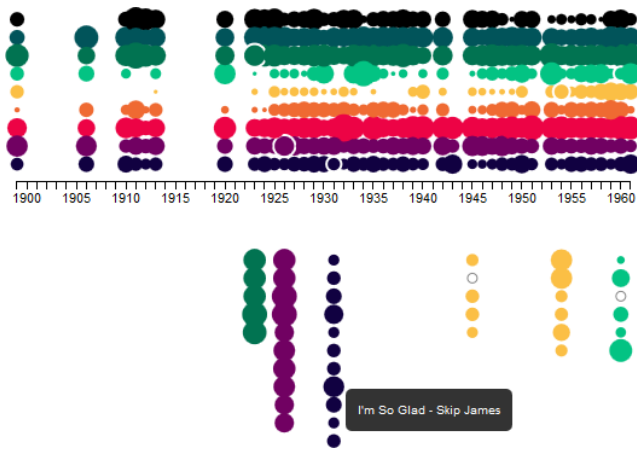
Figure 2: The *time line view* gives a chronological overview of the feature distribution.

several decades. One circle above the timeline represents the average value of one feature in one year, visually encoded by the size of the circle. The selection of such a circle unrolls a list of representative audio samples underneath the timeline.

The *star plot view*, see Figure 3, an iconic visualization of all features, facilitates the detailed comparison of chosen data records. The data records of interest can be selected within the other views. Hence, only a subset of all data records is shown, but with an excessive abundance of information. This reflects the "details-on-demand" part of Shneiderman's visual information seeking mantra [11].

Compared to the other views, the *prototype view*, see Figure 4, supports a more playful technique for data exploration and is well suited to non-specialists. It emerged from the idea that a common human technique for organizing things would be to group similar objects. The user can organize the data records (circles) by creating prototypes that embody a certain feature configuration. All objects similar to a prototype are attracted and move towards the prototype, just like magnetic particles would move towards a magnet. With the help of the *magnetism metaphor* and *metaphor of attraction* respectively, the similarity between musical pieces is applied to the geometric distance in a two-dimensional space. The interpretation and use of interface metaphors for software development should not be too strict in order
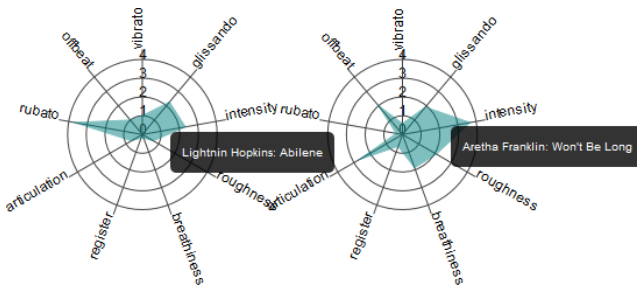

Figure 3: The *star plot view* showing two audio samples and their dominant features for easy comparison.
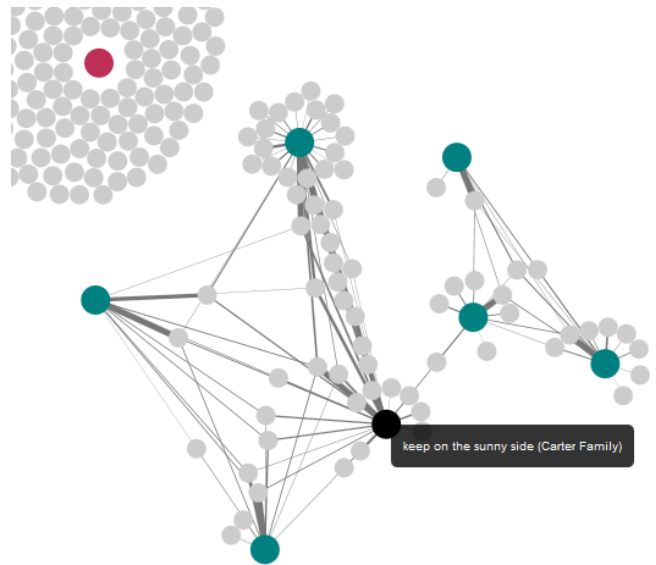

Figure 4: The *prototype view* visualizes the similarity relations between data records of the *Voice and Singing* project. Gray circles (data records) are attracted by cyan circles (prototypes), unless they are not similar enough to any of the seven prototypes.

to avoid hindering the actual interaction purpose, i. e. sticking completely to the real world counterpart of a metaphor can be less gainful than interpreting a metaphor more freely and take account of the specific use case the metaphor is used for [2]. The here formulated distance constraint. I. e. objects that are more similar to the prototype are closer to the prototype than others, differs from the physical model of attraction where an attracted object would move towards the attracting object as close as possible.

There are two kinds of prototypes. Any musical piece (data record) can be transformed into a *live prototype*, adapting the features of the chosen data record and pulling in similar data records. It is also possible to create a custom prototype and configure its features individually, the *virtual prototype*. A line between a prototype and an attracted object indicates the similarity: the thicker the line the stronger the similarity. The user can freely arrange the resulting clusters on the screen by drag interaction. If two or more prototypes exert attractive forces to a musical piece it either goes to the strongest prototype or it is placed in-between the prototypes at a relative position that reflects the similarity relations. Both modes are implemented and can be selected. The similarity is determined by different geometric distance calculations (euclidean distance by default) on the basis of the project-specific rating dimensions (e. g., vibrato, glissando etc.). The user can define a maximum distance to limit the amount of attracted data objects.

Our approach is related to the *prototype theory* classification method formulated, amongst others, by Eleanor Rosch [9]. A prototype represents a class of the classification system and the belonging of a musical piece to a class derives from the distance to the corresponding prototype. This differs from the conventional way of thinking, namely that some-

thing does or does not belong to a certain class without gradual differentiation. The *prototype view* enables users to organize and classify the data collection in very individual ways and at the same time supports them with an automatic feature-based positioning mechanism.

A further new functionality is the rating mechanism that builds upon the prototype semantics and magnetism metaphor. It is introduced in the succeeding Section.

3.3   Rating process

Rating as a process of setting the feature values of a data record can also be referred to as *attribution*. Traditionally, the attribution is implemented by simple text boxes for direct keyboard input. There are, however, use cases demanding another, more intuitive solution. This, for instance, is the case when it comes to certain types of features that require a subjective human perception-based evaluation, or rating. Musicology frequently has to cope with psychoacoustically biased, subjective ratings, be it by expert listeners or non-experts. This is also the situation with the example database "Voice and Singing" as measurement methods for the rated features (vibrato, glissando, intensity, roughness, breathiness, register, articulation, rubato, off-beat) are not available or too extensive to be applied to the whole corpus of musical material. The bigger and denser the dataset is the harder an objective judgement becomes because of the necessity of increasingly finer differentiations.

The input technique should therefore support the user to rate as objectively as possible. This claim is, of course, only valid for the rating of objective features and does not apply to subjective, e.g. emotional, ratings. This leads to the central question: How can the interaction concept and certain auxiliary functions promote a rating with a maximum of objectivity? In general, data input process should be low in complexity, error-resilient, user-friendly, and efficient. We were looking specifically for an alternative, more visual, and directly manipulative, mouse-based input technique.

Vocalmetrics v1.1 supports the user with direct and indirect attribution to rate the features of data records. The direct attribution allows users to do this by entering explicit values (e.g., a number between 0 and 100 or the title string of a musical piece) and is implemented by sliders and simple text boxes for keyboard input. The indirect attribution defines the values by referencing existing data records and provide them for a pair-wise comparison and comparative listening respectively. The user can ensure that his rating is in balance with existing ratings, which serve as a frame of reference. The design process towards a visual implementation of this approach is described in the following.

We examined common user interface controls for data input that could satisfy our requirements for the comparative listening method: Spinner, Rating Controls, Dial Widgets, Sliders, interactive Carpet Plots. Spinners are numeric text boxes that are edited by mouse drag gestures or arrow clicks but feature a limited numeric resolution and a time-consuming configuration. The numeric resolution is even coarser with rating controls (e.g., four stars). Dial widgets allow a suitable numeric resolution but require arc gestures, which is inconvenient for mouse input, and are better suited

to represent angular values. Sliders seem very interesting, especially as several data objects can be placed along the slider, thus offering references for comparison. Interactive carpet plots, derived from Tufte's visualization published in [12], offer a comfortable way of precise, multidimensional numeric input and create characteristic geometric shapes that can easily be compared to others. Reading and input, however, require some experience because of the ever-changing directions.

In order to evaluate those interface controls in the context of our software purposes, an exploratory experiment with an interdisciplinary group of ten students aged from 19 to 28 (probands) examined techniques for describing the attributes of a familiar object by the use of visual elements exclusively. They were given portraits of 11 well-known public persons and a pen. The probands had to describe the personality of one specific celebrity and were allowed to refer visually to the other 10 persons and use drawings to communicate their assessment of the celebrity. The probands used pie diagrams, weighted graphs, iconic annotations and spatial distance (proximity) to indicate close relations or likeliness. It is notable that all approaches put the target person in the center.

The results of the experiment encouraged us to develop the *egg cell metaphor* as a formalization of a proband's approach of describing similarities by the use of spatial distance. We developed an input technique for the rating of musical data and integrated it into the *prototype view* as shown in Figure 5. The rating process starts within the *prototype view* and all existing prototypes are preserved and stay available for comparison. The nine green prototypes (*pure virtual prototypes*) are created automatically and can be used only within the rating process. They are referred to as *pure* because each of them represents only one of the features to be rated. They serve as circularly arranged sliders with their dragging direction oriented towards the center. In addition to the data records attracted by the user-created prototypes, the *pure virtual prototypes* also provide typical representatives of one specific feature that can be used for comparative listening. The visualization consists of the *cell nucleus* (the audio sample to be rated, placed in the center), the *cell envelope* and the *sperms* (other existing data records and prototypes). The user can rate an audio sample by either direct numeric input (using text boxes) or indirectly by referring to existing data objects to indicate relations between those and the audio sample that is to be rated. This means that the user can set the features of a data record (cell nucleus) by moving existing data objects (sperms) into the area of inheritance (within the cell envelope). The closer a data object is moved towards the cell nucleus, the more does the cell nucleus inherit the data object's feature values. The resulting feature values of the rated data record are a combined inheritance from existing data records, *live prototypes* and *virtual prototypes*. It performs an interpolation of the features of all objects within the cell envelope, weighted by their distance to the core. However, a green *pure virtual prototype* inside the cell envelope dominates the rating of its particular dimension and causes the value of all other data objects to be ignored. Thus it defines an absolute value for its respective feature solely. The user would rate an audio sample by starting the rating process, uploading
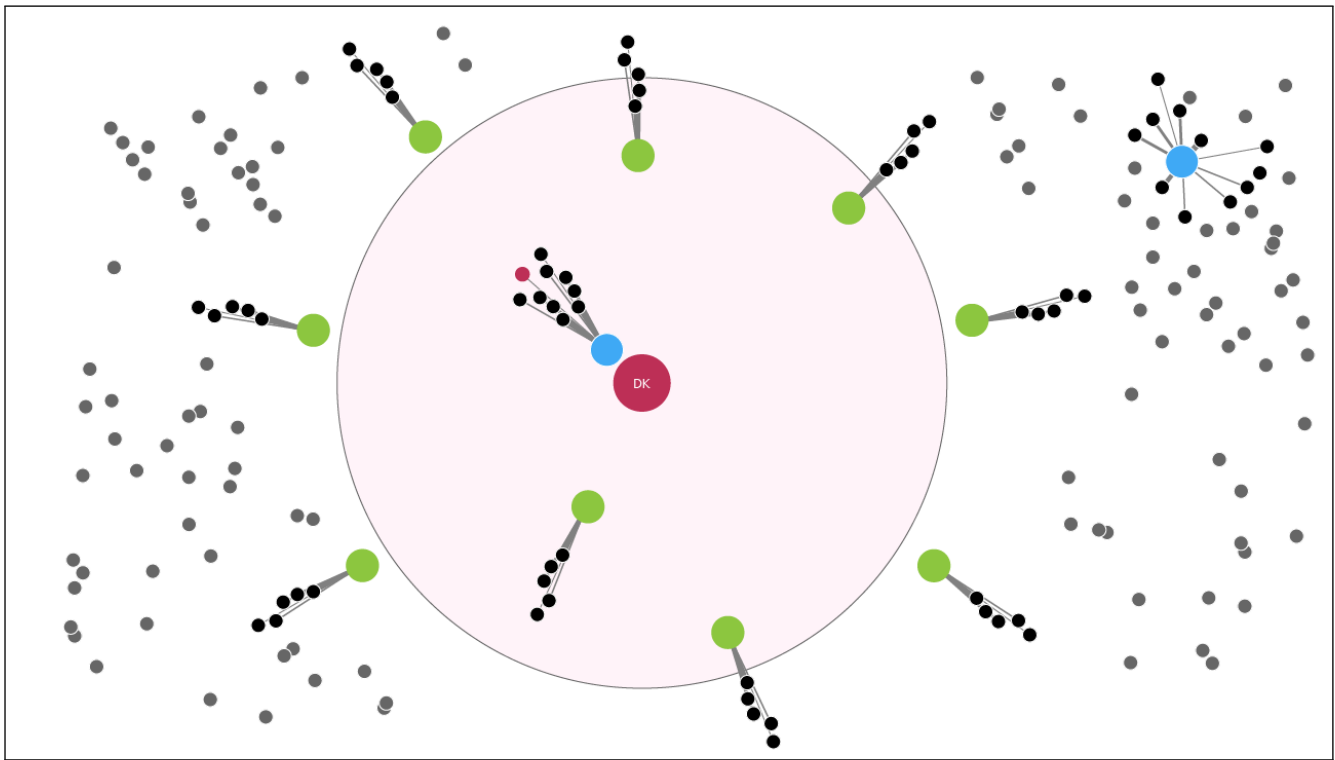
Figure 5: Vocalmetrics' input technique for rating the features of a data record, which is represented as the core of the circular area (egg cell metaphor). Proximity to *pure virtual prototypes* (green) and *live prototypes* (blue) defines the weighted inheritance of their attribution.

the corresponding audio file in case the audio sample is a new data record, and rate the features by listening to the audio sample of interest and compare it to existing audio samples. Thereby, it is easy to define a basic attribution by dragging in similar data objects and refine certain features by using the *pure virtual prototypes*.

The concept has further improvements and interaction patterns. So far, it supports the rating of musical data with the following advantages:

- The whole dataset is present. Each data object can be used for comparative listening and can be referenced to actively affect the rating.

- The use of prototype semantics allows for a weighted inheritance of feature values. This facilitates the rating, because similar audio samples can be adopted and complex feature combinations are applied much quicker than by rating each feature individually. Moreover, prototype semantics automatically provides reference pieces for comparative listening.

- Direct input of numeric values is possible, but largely avoided. For questions like "What is a maximum vibrato?" or "When is it medium?", absolute values are inappropriate. Instead, the focus lies on a more relational rating which complies better with the object of analysis, music.

- The slider-like dragging of objects closer to the core

reflects an intuitive direct relation of proximity and similarity.

3.4 Notes on Implementation

Vocalmetrics is a web application working in any modern browser, either locally in the user's private environment or on the server-side for public purposes. It is built on HTML5, CSS3 and JavaScript and makes use of the JavaScript frameworks *MooTools* and *d3*. Except of some general UI-controls, all visualizations use SVG. The graph structure of the *prototype view* is implemented with the help of d3's *force layout*. The data (users, projects, datasets and feature values) and corresponding files (e.g., audio files, PDF files) are stored persistently within the application's environment (based on JSON, SQLite or a browser's local storage). The data can additionally be exported as a CSV-file for any further processing. The HTML5 Audio API is used to play back audio files encoded in free Ogg-format for compatibility reasons since some browsers, like Mozilla Firefox, do not support the patented MP3-format.

3.5 Discussion

The combination of the *prototype view* and the *egg cell metaphor* lets users benefit from their individual classification of a music repertoire when rating new audio samples to expand the repertoire. They can classify new audio samples according to their own classification methods as represented by the prototypes they create and then use them for adaptation. However, the result can be a complex set of inheriting objects inside the cell envelope. It therefore needs

further decisions concerning inheritance dominances. Also the interplay of direct and indirect attribution should be improved towards a better handling and user experience, i.e., the balance of direct numeric input of feature values and the indirect adaptation by means of inheritance.

The *prototype visualization technique* resembles the *Dust & Magnet* software of Yi et. al. [13]. It was developed independently and emerged as an exemplary implementation of the prototype theory. Our model further includes a distance constraint which is added to the physical model. This causes attracted objects not to move towards the magnetic object as close as possible, but to stop at a certain distance depending on their similarity. The similarity-distance-relation can not be guaranteed generally, which is inherent to every dimension reduction method. Therefore, the introduced *prototype technique* does not reflect the complex similarity relations among all data records. Nevertheless, it can serve as an alternative to other techniques for visualizing similarity, like dimension reduction, which is accompanied by projection errors and also requires the user to have advanced skills for a useful understanding.

While up to today, the usability of the software is tested only from an expert's view the software tool will be tested and evaluated by different user groups in the near future, e. g., in musicological research projects, in university courses, and in music classes at schools. Creating new datasets within the application itself is very useful. The user immediately experiences an improvement of the rating process and appreciates the comfort of listening and rating simultaneously. The instant feedback of how a newly added dataset will be positioned among the existing amount of datasets is easily comprehensible and leads to a very positive user experience. It enables the user to stay focused on the actual task of rating a musical piece and increases the motivation to keep on working. The minimal design is good enough to support a diversion-free interaction and perception process. The visual focus lies on the data rather than on auxiliary functionalities. The *scatter view* enriches a well-known visualization technique with useful functionality, which is gladly accepted by the user, e.g., linking data to one of four visual variables or to show automatically calculated averages. The *prototype view* and its strong agility of lively moving circles can be very comfortable to interact with. However, the logical connections behind it can be hard to understand. On the one hand, it is an easily understandable tool for finding similar objects visualized by their geometric distances and the metaphor of attraction. On the other hand, the user can quickly create a very complex graph structure with n-to-m relations which are hard to differentiate and understand clearly. The benefit of the indirect attribution with the help of the *egg cell technique* can only be evaluated within a real use case. The *egg cell technique* itself should be intuitive enough to describe the concept of inheritance.

4. CONCLUSIONS

Vocalmetrics v1.1 offers many valuable possibilities to visualize similarities and differences between music pieces or audio samples relating to various features of musical structure, sound, and performance as well as relating to meta data. On the one hand, Vocalmetrics can help to classify, explore, and compare large repertoires of music and music of differing provenience. On the other hand, it offers a quick and intuitive approach for visualizing features of and relations between music excerpts and, hence, to communicate about music in various settings. Listeners reflect on these features and relations during the rating process which might also exert an educational gain.

Users can easily create and manage new projects choosing individual audio samples and any rating and meta data dimensions. A main advantage of Vocalmetrics is the intuitive rating procedure with the help of the *egg cell metaphor* and the *prototype technique*. Ratings of different subjects can be compared or merged. Moreover, ratings (individual and merged) and meta data can be exported for further statistical analysis. The data could serve as a starting point for various research issues, e.g. how individuals listen to music and how effects of music on emotion and mood differ between individual listeners. Therefore, Vocalmetrics could serve as a valuable data collection tool for research in music psychology and music pedagogy.

Vocalmetrics furthermore provides techniques that are suitable for non-musical data, in fact any multidimensional data. Vocalmetrics could thus be generalized to a universal tool for data visualization and comfortable data maintenance. Users could benefit from this application as a quick and easy to use analysis tool for any data of their interest.

5. ACKNOWLEDGEMENTS

6. REFERENCES

[1] C. C. Bergan and I. Titze. Perception of Pitch and Roughness in Vocal Signals with Subharmonics. *Journal of Voice*, 15:165–175, 2001.

[2] A. F. Blackwell. The reification of metaphor as a design tool. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 13(4):490–530, 2006.

[3] M. A. Casey, R. Veltkamp, M. Goto, M. Leman, C. Rhodes, and M. Slaney. Content-based music information retrieval: Current directions and future challenges. *Proceedings of the IEEE*, 96(4):668–696, 2008.

[4] P. H. Dejonckere, P. Bradley, P. Clemente, G. Cornut, L. Crevier-Buchman, G. Friedrich, P. Van De Heyning, M. Remacle, and V. Woisard. A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques. *European Archives of Oto-Rhino-Laryngology*, 258(2):77–82, 2001.

[5] A. Lomax. Song Structure and Social Structure. *Ethnology*, 1(4):425–451, 1962.

[6] A. Lomax. *Folk song style and culture*. Transaction Publishers, 3 edition, 1968. Reprint of the 1968 ed. published by American Association for the

Advancement of Science (2000).

[7] A. Lomax, R. Rudd, V. Grauer, N. Berkowitz, B. L. Hawes, and C. Kulig. *Cantometrics : a handbook and training method.* Extension Media Center, University of California, Berkeley, 1976.

[8] K. Omori, H. Kojima, R. Kakani, D. H. Slavid, and S. M. Blaugrund. Acoustic Characteristics of Rough Voice: Subharmonics. *Journal of Voice*, 11(1):40–47, 1997.

[9] E. Rosch. Prototype classification and logical classification: The two systems. *New trends in conceptual representation: Challenges to Piaget's theory*, pages 73–86, 1983.

[10] C. E. Seashore. *In Search of Beauty in Music.* The Ronald Press Company, New York, 1947.

[11] B. Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. In *IEEE Symposium on Visual Languages 1996*, pages 336–343. IEEE, 1996.

[12] E. R. Tufte. *Envisioning information.* Graphics Press, Cheshire and Conn, 13th print., may 2011 edition, 2011, c1990.

[13] J. S. Yi, R. Melton, J. Stasko, and J. A. Jacko. Dust & magnet: multivariate information visualization using a magnet metaphor. *Information Visualization*, 4(4):239–256, 2005.