Similarity knowledge formalisation for audio engineering

Similarity Knowledge formalisation for Audio Engineering

Christian Severin Sauer, Thomas Roth-Berghofer,
Nino Auricchio, and Sam Proctor

School of Computing and Technology, University of West London,
St Mary's Road, London W5 5RF, United Kingdom,
{christian.sauer|thomas.roth-berghofer|
nino.auricchio|sam.proctor}@uwl.ac.uk

Abstract. The work presented in this paper demonstrates the externalisation of tacit knowledge encoded in a special vocabulary used by experienced audio engineers to affectively describe emotions evoked by a sound or piece of music. We created an approach to formalising the adjectives describing the timbre of a sound as well as their relationships and composing this special vocabulary. The known problems associated with the attempt to formalise and quantise emotions and descriptors of emotions are discussed. The problems are mainly the vagueness of emotions and the variation in the emotions the same single percept can trigger in different humans. To amend these problems we used the Case-based reasoning (CBR) approach. In particular we demonstrate our project work to use CBR's ability to process fuzzy and incomplete queries to emulate the vagueness and differentiation associated with the emotions triggered by a sound percept. We then describe our way of capturing the experience of audio engineers by mapping the formalised vocabulary of timbre-describing adjectives to workflows. These workflows describe the actions to be performed to change the spectral shaping of a sound to change its emotional effect in reflection to the descriptive adjectives provided.

Keywords: Case-based reasoning, audio engineering, similarity measures, knowledge formalisation

1 Introduction

With automatic composition and improvisation of music expressing the individual style of a human composer as well as the automatic expressive performance of music, the two main steps of music production are by now quite well researched areas [19, 20]. Both areas include the need to formalise emotions to a certain degree. A third area where emotions are of importance to music production is the mastering of a recording into a final audio product.

The formalisation of affective, emotional statements or descriptive adjectives of an emotion is still a problem [12, 7]. This problem is often encountered by applications dealing with art, as art is deeply linked to emotions and perception of such. In the case of music, a variety of approaches already exist to formalise emotional annotations of music, see for example [23].

There are already a variety of approaches to automated composition of expressive music [19, 27] and the expressive performance of music [4]. Both tasks include the problem of formalising emotions to a certain degree. This formalisation is needed as Information relating to the intended emotional effect of a composition and/or performance must be encoded and then integrated in the systems knowledge. However, formalising emotions is not easy due to the different perception and emotional links individual humans might have to a percept.

Next to composition and performance, a third very important task in professional music production is the mastering of a recording of a sound or song. Robert A.Katz describes mastering as, quote:*"Mastering is the last creative step in the audio production process, the bridge between mixing and replication-your last chance to enhance sound or repair problems in an acoustically-designed room-an audio microscope. Mastering Engineers lend an objective experienced ear to your work; we are familiar with what can go wrong technically and aesthetically. Sometimes all we do is-nothing! The simple act of approval means the mix is ready for pressing. Other times we may help you work on that problem song you just couldn't get right in the mix, or add the final touch that makes a record finished and playable on a wide variety of systems."* [16].

The main process of mastering is described by applying a set of spectral modifications to the sounds spectrum to achieve a change in the timbre or more specifically the emotional effect the perception of the sound triggers in a listener. As the process is goal oriented—with the goal being the desired change in the emotional effect a sound has—the vocabulary used to describe this effect-change is given by terms that describe the emotion desired to be triggered or altered, most common increasing or decreasing an emotional effect. Thus we find terms like 'make it sound more warm' or 'make it sound less harsh' and the use of onomatopoeia common in the language of audio engineers. The experience of audio engineers now is applied in the linkage between these descriptive adjectives or emotional descriptors and the choice and application of spectral modifications used to achieve the desired change of the sound. This paper introduces our work of implementing a system that allows an audio engineer during the mastering stage of a music production to apply descriptive adjectives for the automatic selection of workflows, using presets of spectral modifications that deliver the intended alteration of the sounds emotional effect. A preset can be described as a selection of frequency descriptors assigned with definite values for said frequencies. A preset can further contain information on defined effects such as reverb or delay and the values to be applied to these effects. We provide a dynamic recommendation of workflow steps, consisting of the appliance of presets to the sound being mastered, during mastering sessions. To achieve this goal we map workflow descriptions such as the temporal ordering of preset use, next to the selection of presets, to the descriptive adjectives and descriptors of amounts of an effect contained in the vocabulary used by audio engineers. This approach seems worthwhile as it happens to be a fact that amongst audio engineers almost any changes applied to an audio product during the mastering stage of its production is labelled with descriptive adjectives. These adjectives are almost always based upon the change of a sounds Timbre that is to be achieved by the application of said changes. It is not uncommon to encounter a description like ‘make it sound a bit more warm and punchy’. To give some examples from the literature on mastering, quote:’ Gentle - Opposite of edgy. The harmonics-high and upper midst-are not exaggerated, or may be even weak. Grungy - Lots of harmonic I.M. distortion. Harsh - To much upper midrange, usually around 3 kHz. Or, good transient response, as if the sound is hitting you hard’[5]. As the interlinking of such descriptive terms, which we call Timbre descriptors, to defined settings of frequency reshaping is mainly based upon years of experience of a sound engineer, it seemed a worthwhile goal to establish an approach to reuse this experience. Thus we introduced our way of making this experience available in a music production software.

The rest of the paper is structured as follows. We interlink our approach with the current state-of-the-art of research in the field of artificial music composition and performance in the following section. Based upon the identified problems we introduce the challenges we expect for our approach in section 3. We then introduce our approaches to externalising and formalising the tacit knowledge experienced audio engineers possess. Also we introduce our use of CBR for the provision of the formalised knowledge and introduce our use of the different knowledge containers of CBR [28] to provide the captured knowledge in section 4. A summary and outlook on future aspects of our work then concludes the paper.

2 Related Work

For music, there already exist a variety of approaches to formalise emotional annotations and/or descriptive terms that either describe the mood of the music or the way it is to be played [20, 10]. Such approaches for now deal with either playing music in a certain defined way to convey an emotion [8] or to select songs or sounds that are associated with a mood or emotional state [29]. Also, for automated composing, the question of integrating a formal description of the mood the composed music should match is already well researched [20, 4]. The main problem of the formalisation of emotions or emotional perceptions is that emotions per se are not easy to be a) defined and b) quantised/formalised in a machine compatible way [15, 9, 12]. Another problem we were facing was that we tried to quantify and cluster descriptive adjectives based on very vague data given by the individual descriptions of the emotional effect a sound has on a person describing this effect. The difficulties of capturing a sounds timbre are also described by Sophie Donnadieu [11],as she describes it as follows, quote: “*Indeed, it is timbre’s ‘strangeness’ and, even more, its ‘multiplicity’ that make it impossible to measure timbre along a single continuum, in contrast to pitch (low to high), duration (short to long), or loudness (soft to loud). The vocabulary used to describe the timbres of musical instrument sounds indicates the multidimensional aspect of timbre. For example, ‘attack quality, ‘brightness,’ and* *'clarity' are terms frequently used to describe musical sounds.*" So the vagueness of the data is based on said variation in the individuals perceptions when they either should describe an emotional effect or perceive something that is annotated with a particular emotion but have a complete different idea of the actual emotion this percept triggers [24, 11, 13].The work of Graham Darke described in [9] provides a good insight in the above mentioned problems. Darke conveyed an experiment in which he tried to establish if there is a common understanding of how humans perceive and describe timbres.

A way to circumvent the described problems of a lack of quantifiable measures and vagueness is given by also allowing for vagueness and a certain amount of ambiguity within the techniques used for formalising and retrieving presets based on descriptive adjectives. One way to circumvent the vagueness accompanying the formalisation of emotional descriptors is offered by the use of CBR (CBR) [1, 17]. CBR allows for similarity-based retrieval which also allows for a vague query and, given for example a fuzzy case representation, does not require an exact match of a query to produce a result. [30, 14] As shown in various approaches CBR can be used to guide the emotional component of automatic composition as well as performance of music [8, 3, 25, 20].

3 Challenges of Formalising Descriptive Adjectives and Preset Mapping

When describing the timbre of a sound, there is tacit knowledge present in the implicit emotional-descriptions provided by the use of descriptive adjectives on the timbre. Externalising this tacit knowledge about a timbre or emotional effect a sound has was the main challenge of the knowledge formalisation task at hand. The first challenge in this task was it to find common ground to start with a basic selection of timbres and their emotional descriptors. The questions that were raised by this challenge were:

1. Do annotations of timbre with adjectives/terms vary between different people?
2. Are there significant clusters, distances, patterns in the classification of the adjectives/terms used to describe timbre?

As we have seen in the related work, the first question can almost certainly answered with a solid *Yes*. This is additionally true in the light of the fact that the vocabulary we try to establish is not yet free of redundancy and ambiguities. This is also true as terms often overlap and have different meanings in different contexts, which is common for emotions themselves and especially if one tries to formalise emotions as described in [2]. As Donnadieu, Porcello, Darke [9, 11, 26] and others have indicated so far, it is difficult to map descriptive adjectives, or Timbre descriptors, to certain definite timbres. We therefore postponed our own knowledge gathering and relied on using interviews with experts, in our case experienced audio engineers, as our initial means of knowledge gathering. We used the audio engineers expertise to establish a first set of timbre descriptors most frequently encountering in the domain of audio mastering. After establishing the initial, basic set of timbre descriptors we further needed to map the changes of timbre to the application of workflows consisting of a sequence of applications of presets to the sound. Therefore, we had to measure the effect/change of timbre the application of a preset has on a sound.

4 Knowledge Formalisation Approach using CBR

As outlined we used Case-based Reasoning (CBR) in our approach to formalise and reuse the elicited tacit knowledge of audio engineers employed during task of mastering an audio product. We chose CBR as a suitable methodology [30] for our task as it is able to handle the described vagueness of the domain of our approach. This suitability is backed up by a variety of uses of CBR in the domain of music composition and expressive performance to date as well as its use for handling otherwise difficult to formalise knowledge [3, 25]. The manner in which we employed CBR for knowledge elicitation and formalisation are the subjects of this section.

For the purpose of modelling and testing the knowledge of our system we employed our SDK myCBR[1] in its latest version 3.0. Creating our CBR engine with myCBR provided us with an application-independent way to provide our CBR engine and its modelled knowledge to a wide variety of Java-based and Android-based applications via the API of myCBR 3.0.

4.1 Knowledge and Data present in our Domain

In our domain of audio mastering we face three sets of artefacts. The first set is given by the already described presets. Knowledge present with regard to this second set of artefacts is given by descriptive adjectives present in audio engineering literature and day to day practise of audio engineers. The third set of artefacts our approach is built upon is given by workflow descriptions. These descriptions are present in step by step or best practise descriptions of the application of one or more presets. The application of these workflows aims at achieving the reshaping of an audio product, resulting in a shift of the specific emotion being evoked by the audio product. Again this knowledge is partly available from literature but mainly only present as tacit knowledge of experienced audio engineers.

4.2 Advantages of CBR in our Domain of Interest

CBR is able to match the customer's language, in our case descriptive adjectives and likely vague terms describing the amount of an effect desired. In our case this means that we could use fuzzy descriptive adjectives like 'muddy' or 'bright'

[1] http://mycbr-project.net

to define queries or problem descriptions, to retrieve cases holding the workflow descriptions to achieve this effect as their solution part. CBR is able to retrieve cases even based on only sparse problem descriptions. This is very useful while trying to retrieve workflows that are only partly specified by timbre descriptors and amounts of effects. Also, CBR heavily relies on similarities which are, as introduced in section 3, comparatively easy to elicit within our domain of interest. Additionally CBR allows for queries that combine retrieval and filtering in the way of queries like: ‘This should sound really airy, but not harsh’.

4.3 Case Structure and Attributes

Our case structure is following the approach of structured CBR. Our cases consist of a problem description part, specifying the present timbre of a sound and the desired change in the timbre and an indicator for the amount of this change. The solution part of our cases is, for now, given by a workflow description of one or more presets to be applied to the sound and their order of application. The basic mapping of workflows to the three descriptive adjectives, given by the timbre descriptor describing the present sound, the timbre descriptor describing the timbre the sound should change to and the amount descriptor specifying the amount of change that should effect the timbre, suits the approach of structural CBR [6] very well. For now we rely on a percentage for the amount descriptor, ranging from 0 to 100, whilst 0 means no change to the timbre of the input sound at all and 100 translates to the total conversion of the timbre of the input sound to the timbre specified by the timbre descriptor of the cases problem description. We have not yet mapped these percentages of effect to verbal amount descriptors like ’a bit’, ’a lot less’, ’much more’. As this is not a task to complicated, we aim soon to reach a case format like the following:

Attribute	Value	Solution
Input Timbre:	Bassy	Apply preset 4
Target Timbre:	Hollow	then use filter 7 with
Amount of change:	+50	with 80 percent treble.

Our case structures thus reflects the problem description, consisting of an Input timbre as the timbre descriptor characterising the input sound on which the modulating workflow is to be used. The problem is further specified by a Target timbre which is a timbre descriptor characterising the way the sounds timbre should change (make it sound more/less ’timbre descriptor’). The third part of our problem description is the Amount of change desired to take effect, ranging from 0 (no effect) to 100 per cent (total conversion of the timbre), that can hold negative’.

The second part of our case structure is the Solution description, which is providing a workflow description of how to obtain the Timbre change to the extend desired. This workflow description can be stored as an URL, a text or a sequence of presets selected from a database.

Additionally to the timbre descriptor and amount descriptor as problem description and a workflow for the desired timbre change, possible further attributes for a case can, as described in [9], be seen in: Onomatopoeia describing the sound itself and a description of the sounding situation, e.g. 'opera house', 'marching' or 'club'.

4.4 Vocabulary

The vocabulary we use was identified to consist of timbre descriptors and the names or id's of pre-sets that provide this timbre after being applied to an audio product. The vocabulary further consists of a set of amount descriptors as for example 'a bit', 'much', 'a touch'. Additionally we also incorporate workflow descriptions into the vocabulary to provide the workflow suggestions as the solution parts of our case structure.

Our initial approach of establishing a vocabulary was limited to general music settings, with regard to the domain complexity and the described problems known with the formalisation of timbre describing adjectives or even onomatopoeia. By omitting specific genres, such as rock or jazz, we aimed at keeping the vocabulary as 'flat' or simple as possible. We did so to prevent our effort from being too specific (to a genre) and to be reusable in a more general way for audio mastering. As already mentioned we elicitated the basic timbre descriptors from the Literature as well as from audio engineers tacit knowledge. We ended up with a total number of 34 timbre descriptors. We, for now, haven't established amount descriptors but aim to fracture the interval of 0 to 100 per cent of timbre change to at least 20 amount descriptors to provide a 5 per cent granularity for the desired impact of a timbre changes described in the problem description of our cases. Thus we aim to discretise the interval to a number of amount descriptors ranging from 'Not' 'None' or 'Should not sound' to 'Totally' 'Convert to' 'Fully'.

4.5 Similarity Measures

To establish the similarity of attributes and between cases in our system, we needed to establish local as well as global similarity measures. For the formalisation of the similarity between Timbre descriptors we considered two options. The first option was given by employing Multidimensional scaling (MDS). MDS was intended to achieve a dissimilarity matrix describing the dissimilarity of descriptive adjectives. The computation of a dissimilarity matrix is identical to one of the main approaches used in CBR to formalise similarities and thus offers a way to capture the similarities of timbres, if there are any to discover as patterns. By using MDS we hoped to establish if there are patterns within the terms regarding their similarity of use when describing a timbre.

The second option was to establish a taxonomy of descriptive adjectives. Such a taxonomy can be seen as a comparable to a taxonomy of colours, classifying the emotion a sound triggers by the use of timbre descriptors. An example of a parent and two child notes in such a taxonomy would look like the following: [Treble (parent) - [Toppy (child)] [Bright (child)]]. The taxonomy would be used to establish the similarity of two emotional-descriptors by their position within the taxonomy and their distance and also store adaptation knowledge as we detail in the following subsection. Both options are very close to the common data structures used within CBR to formalise similarities and thus offer easy approaches to formalise the similarity of descriptive adjectives.

We decided to apply the second option of building a taxonomy of timbre descriptors based upon the elicitation of their similarity from the tacit knowledge of experienced audio engineers. The taxonomy consists now of 32 nodes beginning with the most abstract 'Timbre descriptor' and expanding down to its leafs with the most concrete descriptors of timbres. See 1 for a part of the initial taxonomy describing timbres in the higher frequency ranges.

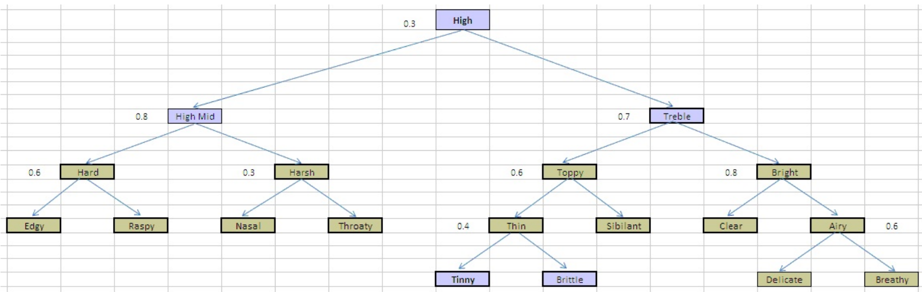

Fig. 1. A part of the similarity measure as graphic taxonomy, describing timbres in the high frequency range

The initial taxonomy then was refined and modelled using the Workbench component of the myCBR 3.0 application. Within myCBR 3.0 we modelled the thre local similarity meassures for the present timbre, the timbre a sound should change to and the amount of this change. Please see 2, 3 and 4 for an impression of these local the similarity measures.

Beyond the formalisation of the basic timbre descriptors we further grouped the timbre descriptors into more abstract groups describing families of timbres and even more abstract the frequency ranges were these families of timbre descriptors are most commonly used. This approach aims at being able to include additional information in our similarity measures. Next to the timbre descriptors we yet have to, as described, provide a similarity measure for the amount descriptors. A possible future addition could be seen in another taxonomy describing the context of an audio signal being manipulated. Such a context could be provided by the instrument that is used to generate the audio signal. So again we could establish a taxonomy of instruments which would begin with abstract families of instruments, like 'brass' or 'strings' and get more specific in the deeper levels of the taxonomy distinguishing individual instruments of a family, like for example: [Organ (parent) - [Hammond (child)] [Pipe (child)]]

For the global similarity measure we initially use a non-weighted sum of the local similarities. For more complex cases, that include a set of more then input

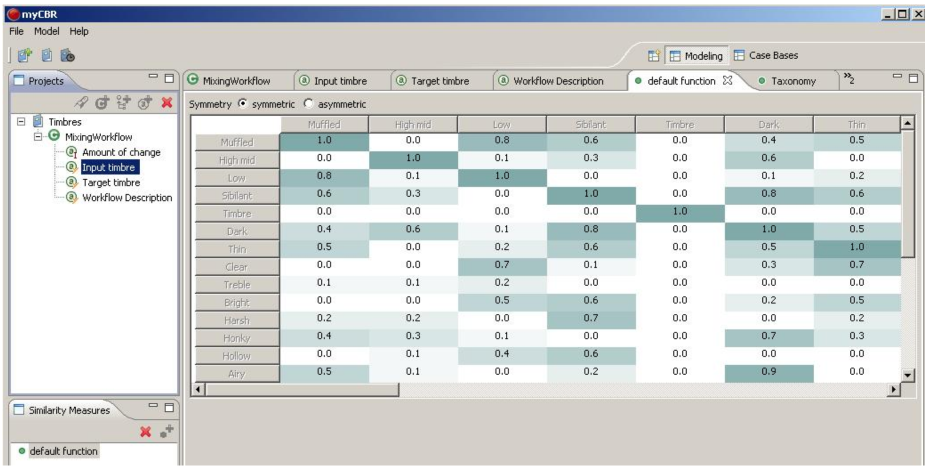

Fig. 2. The similarity measure for the timbre of the present (input) sound in table view

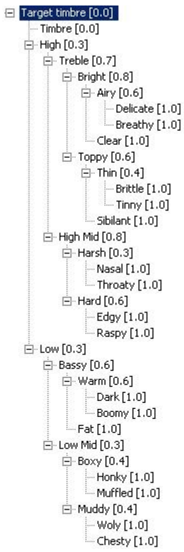

Fig. 3. The similarity measure for the target timbre modelled as taxonomy

and target timbre descriptors and their amount descriptors and may also be extended to include the context description of a sound, like the instrument, we plan to provide weighted sums of the respective local similarities.

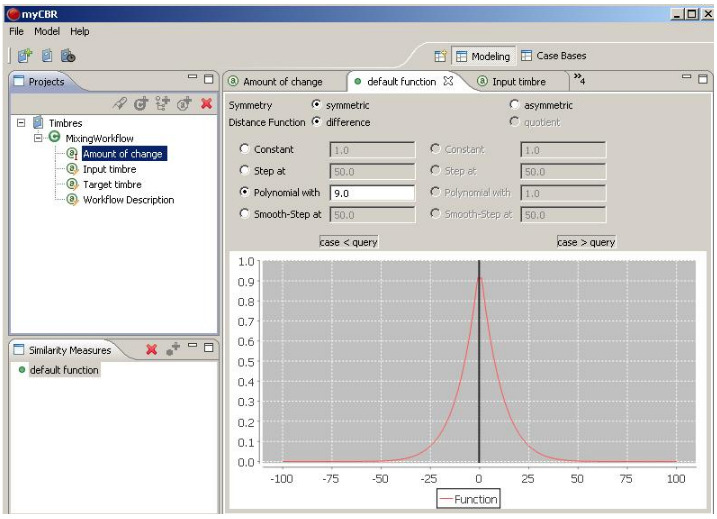

Fig. 4. The simple percentage based similarity measure for the amount of effect

4.6 Adaptation Knowledge

The purpose of adaptation knowledge in CBR is it to adapt the solution the most similar case provides to problem at hand, if needed. A basic example of this is adaptation by replacement. By storing adapted and tested (verified) cases the system also gains new knowledge. Such adaptations are also desirable for our system we therefore will integrate adaptation knowledge in a number of ways. An obvious, at least from a CBR perspective, way to obtain and formalise adaptation knowledge is given by making use of the taxonomies introduced as similarity measures. Our system uses taxonomies for the descriptive adjectives as well as in the future for the sound context, aka instrumental families. These taxonomies are used as adaptation knowledge in the way that adaptation knowledge is stored in the parent-child relations formalised in the taxonomies. For example the taxonomy of descriptive adjectives can be used to provide replacements for invalid or unwanted adjectives in the following way: Assume that within the taxonomy there are nodes of the following kind: [Treble - [Toppy] [Hard]], see 1. If 'Toppy' was defined within a query case's problem description but is not available within any case from the case base, the taxonomy could be used to select the most similar adjective, 'Hard' instead and thus use either 'Hard' or fall back to the parent node 'Treble' as the next more abstract timbre descriptor to replace 'Toppy'.

A more complex way to adapt a case in our domain is enabled by the fact that there do exist timbres that are 'opposing' each other. By 'opposing' we mean that there can be two timbres, like 'Airy' and 'Boxy' that cancel each other out if they are applied to the same sound. As we elicitated the knowledge from the audio engineers they pointed out that applying such opposing timbres is a common practise while mastering an audio product. We thus asked them to provide us, next to the similarity of two timbres, also with the 'oppositeness' of them. We formalised this oppositeness in a negative similarity measure, ranging from 0, not opposite at all, to -1 marking total opposition, thus describing two timbres that cancel each other out. The degrees in between 0 and -1 describe the ability of two timbres to soften the effect of the other one again roughly mapped to a 100 per cent interval. So, for example, if we look at the Timbre 'Nasal' the Timbre 'Dark' has an oppositeness of -0.2, so applying 'Dark' to a 'Nasal' sound reduces it 'Nasal' Timbre by roughly 20 per cent. The way we intend to use this oppositeness as adaptation knowledge is by providing rules off the following nature: Assume a query case asking for a shift from a Nasal Timbre to a Harsh Timbre with 40 per cent effect strength. The best case in the case base only provides the workflow for a change for a Nasal Timbre to a Harsh Timbre with 20 per cent effect strength. The remaining 20 per cent of shifting the Nasal Timbre to the Harsh Timbre could be accomplished by applying a 20 per cent opposite Timbre, like the Dark Timbre, thereby reducing the Nasal Timbre by another 20 per cent. So we cancel out 20 per cent of the Nasal Timbre by applying another workflow to add the -0.2 opposite Dark Timbre. The resulting rules thus will be of the form general form: If effect strength not reached / exceeded: Find a case Similar or opposite Timbre to apply with regard to the missing amount of effect strength. A particular example there would be: If shift from Nasal to Harsh with x per cent not reached: Apply shift from Nasal to Dark with x-best case applied effect strength.

5 Summary and Outlook

In this paper we presented our approach to externalising tacit knowledge encoded in a special vocabulary used by experienced audio engineers to effectively describe emotional effects and Timbres of audio products. We have described our approach to formalise the Timbre descriptors and to map them to frequency reshaping workflows of pre-sets application. We surveyed the known problems associated with the attempt to formalise and quantise emotions in general and adjectives describing timbre in music specifically. Upon this survey we introduced CBR as a methodology to amend the two main problems experienced while trying to formalise emotions and/or adjectives describing timbres. We identified these problems as the vagueness of terms and the variance of emotions invoked by the same sound in different humans. We then introduced our approach to use CBR's ability to process fuzzy and incomplete queries and the ability to choose between grades of similarity of retrieved results to emulate the vagueness. We detailed especially on the approaches we used to formalise the knowledge into the four knowledge containers of CBR.

For the future of our approach we aim at adding a more detailed way of user modelling into our system. We do so, as we have established that it is a major difference if our system will interact with artists from different genres and/or users of different level of experiences, aka novice to expert engineers. It is also of importance to establish what a user might have as a goal overall, because mixing, composing and mastering are three different contexts in which the retrieval of presets would differ significantly in a later version of our system. We also aim to research about the importance of the dialogue between audio engineers themselves and audio engineers and students as it is for example described by Porcello [26]. The tacit knowledge conveyed within these dialogues is also of concern to our approach, aside the basic approach of mapping timbre descriptors to workflow selections. This concern is introduced by the fact that workflow knowledge on how to change the timbre of a sound is often encoded within the dialogues occurring during a in a mastering session. As we are aiming for an extension of our system we have to consider the possibilities to extract workflow information from dialogues, which is a current research area in CBR [21, 18, 22].

References

1. Aamodt, A., Plaza, E.: Case-based reasoning: Foundational issues, methodological variations, and system approaches. AI Communications 1(7) (Mar 1994), ftp://ftp.ifi.ntnu.no/pub/Publikasjoner/vitenskaplige-artikler/aicom-94.pdf; letzte Verifikation 11. Juni 2007
2. Arbib, M., Fellous, J.: Emotions: from brain to robot. Trends in cognitive sciences 8(12), 554–561 (2004)
3. Arcos, J., Grachten, M., de Mántaras, R.: Extracting performersŠ behaviors to annotate cases in a cbr system for musical tempo transformations. Case-Based Reasoning Research and Development pp. 1066–1066 (2003)
4. Arcos, J., De Mantaras, R., Serra, X.: Saxex: A case-based reasoning system for generating expressive musical performances*. Journal of New Music Research 27(3), 194–210 (1998)
5. Bartlett, J.: Practical recording techniques. Focal Pr (2002)
6. Bergmann, R., Göker, M.: Developing industrial case-based reasoning applications using the inreca methodology. In: Workshop at the International Joint Conference on Artificial Intelligence, IJCAI-Automating the Construction of Case Based Reasoners, Stockholm (1999)
7. Broekens, J., DeGroot, D.: Emotional agents need formal models of emotion. In: Proc. of the 16th Belgian-Dutch Conference on Artificial Intelligence. pp. 195–202 (2004)
8. Canamero, D., Arcos, J., de Mántaras, R.: Imitating human performances to automatically generate expressive jazz ballads. In: Proceedings of the AISBŠ99 Symposium on Imitation in Animals and Artifacts. pp. 115–20. Citeseer (1999)
9. Darke, G.: Assessment of timbre using verbal attributes. In: Conference on Interdisciplinary Musicology. Montreal, Quebec (2005)
10. DE MANTARAS, R.: Towards artificial creativity: Examples of some applications of ai to music performance. 50 Anos de la Inteligencia Artificial p. 43

11. Donnadieu, S.: Mental representation of the timbre of complex sounds. Analysis, Synthesis, and Perception of Musical Sounds pp. 272–319 (2007)
12. Fellous, J.: From human emotions to robot emotions. Architectures for Modeling Emotion: Cross-Disciplinary Foundations, American Association for Artificial Intelligence pp. 39–46 (2004)
13. Halpern, A., Zatorre, R., Bouffard, M., Johnson, J.: Behavioral and neural correlates of perceived and imagined musical timbre. Neuropsychologia 42(9), 1281–1292 (2004)
14. Hirota, K., Yoshino, H., Xu, M., Zhu, Y.: An application of fuzzy theory to the case-based reasoning of the cisg. Journal of Advanced Computational Intelligence Vol 1(2) (1997)
15. Hudlicka, E.: What are we modeling when we model emotion. In: Proceedings of the AAAI spring symposium–Emotion, personality, and social behavior (2008)
16. Katz, B., Katz, R.: Mastering audio: the art and the science. Focal Press (2007)
17. Lenz, M., Bartsch-Spörl, B., Burkhard, H.D., Wess, S. (eds.): Case-Based Reasoning Technology: From Foundations to Applications, Lecture Notes in Artificial Intelligence, vol. LNAI 1400. Springer-Verlag, Berlin (1998)
18. Madhusudan, T., Zhao, J., Marshall, B.: A case-based reasoning framework for workflow model management. Data & Knowledge Engineering 50(1), 87–115 (2004)
19. de Mantaras, R.: Making music with ai: Some examples. In: Proceeding of the 2006 conference on Rob Milne: A Tribute to a Pioneering AI Scientist, Entrepreneur and Mountaineer. pp. 90–100
20. de Mantaras, R., Arcos, J.: Ai and music: From composition to expressive performance. AI magazine 23(3), 43 (2002)
21. Minor, M., Bergmann, R., Görg, S., Walter, K.: Towards case-based adaptation of workflows. Case-Based Reasoning. Research and Development pp. 421–435 (2010)
22. Minor, M., Tartakovski, A., Bergmann, R.: Representation and structure-based similarity assessment for agile workflows. Case-Based Reasoning Research and Development pp. 224–238 (2007)
23. Oliveira, A., Cardoso, A.: A computer system to control affective content in music production
24. Pitt, M.: Evidence for a central representation of instrument timbre. Attention, Perception, & Psychophysics 57(1), 43–55 (1995)
25. Plaza, E., Arcos, J.: Constructive adaptation. Advances in Case-Based Reasoning pp. 306–320 (2002)
26. Porcello, T.: Speaking of sound. Social Studies of Science 34(5), 733–758 (2004)
27. Ribeiro, P., Pereira, F., Ferrand, M., Cardoso, A., Polo, I., de Marrocos, P.: Case-based melody generation with muzacazuza
28. Richter, M.M.: Introduction. In: Lenz, M., Bartsch-Spörl, B., Burkhard, H.D., Wess, S. (eds.) Case-Based Reasoning Technology – From Foundations to Applications. LNAI 1400, Springer-Verlag, Berlin (1998)
29. Typke, R., Wiering, F., Veltkamp, R.: A survey of music information retrieval systems (2005)
30. Watson, I.: Case-based reasoning is a methodology not a technology. Knowledge-Based Systems 12(5), 303–308 (1999)