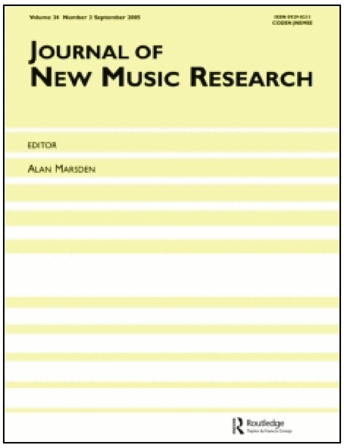

Journal of New Music Research
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713817838

TempEst: Harpsichord Temperament Estimation in a Semantic Web Environment

Dan Tidhar[a]; György Fazekas[a]; Matthias Mauch[b]; Simon Dixon[a]
[a] Queen Mary University of London, UK [b] National Institute for Advanced Industrial Science and Technology, Japan

Online publication date: 19 February 2011

To cite this Article Tidhar, Dan , Fazekas, György , Mauch, Matthias and Dixon, Simon(2010) 'TempEst: Harpsichord Temperament Estimation in a Semantic Web Environment', Journal of New Music Research, 39: 4, 327 — 336
To link to this Article: DOI: 10.1080/09298215.2010.520720
URL: http://dx.doi.org/10.1080/09298215.2010.520720

Journal of New Music Research
2010, Vol. 39, No. 4, pp. 327–335

TempEst: Harpsichord Temperament Estimation in a Semantic Web Environment

Dan Tidhar[1], György Fazekas[1], Matthias Mauch[2], and Simon Dixon[1]

[1]Queen Mary University of London, UK; [2]National Institute for Advanced Industrial Science and Technology, Japan

Abstract

Issues concerning tuning and temperament bear relevance to music research in areas such as historical musicology, performance and recording studies, and music perception. We have recently demonstrated that it is possible to classify keyboard temperament automatically from audio recordings of standard musical works to the extent of accurately distinguishing between six different temperaments often used in harpsichord recordings.

The current paper extends this work by combining digital signal processing with semantic computing and demonstrates the use of the temperament classifier in a Semantic Web environment. We present the Temperament Ontology which models the main concepts, relationships, and parameters of musical temperament, and facilitates the description and inference of various characteristics of specific temperaments. We then describe TempEst, a web application for temperament estimation. TempEst integrates the classifier with ontology-based information processing in order to provide an extensible online service, which reports the class and properties of both known and unknown temperaments. TempEst allows users to upload harpsichord recordings, and provides them with an estimated temperament as well as other inferred characteristics of the instrument's tuning.

1. Introduction

In recent years, historical performance practice of early music on period instruments has become well-established as part of mainstream scholarship, musicianship, as well as public music consumption. As part of this process, the hegemony of equal temperament as the main paradigm for tuning keyboard instruments has been weakened, and increasing attention is being directed towards historical, unequal temperaments (e.g. Lehman, 2005). Algorithms for music visualization and synthesis using arbitrary temperaments are available in software such as Scala,[1] however, the idea of analysing and extracting temperament from audio recordings has not yet been explored in such applications.

There are many ways in which automatic temperament estimation can be useful to musicians, musicologists, and listeners. For educational purposes such as ear training and studying different temperaments in the context of real recordings, it would be useful to have a music database which can perform retrieval according to temperament. Professional users such as keyboard tuners and performers would be assisted by a system providing feedback about tuning accuracy and stability, as well as helping to classify creative temperaments (i.e. those which do not strictly adhere to any known recipe) and determine some of their properties. Furthermore, musicologists are likely to find a temperament estimator to be a useful tool when studying performance practice from recordings. Advances in music signal processing have enabled the automation of many aspects of the analysis of music recordings. In particular, our recent work on keyboard temperament estimation (Tidhar, Mauch, & Dixon, 2010) shows that it is possible to design an automatic system to recognize temperament directly from audio recordings of unknown works. We presented a classifier which is capable of distinguishing, with high accuracy, between six different temperaments

[1]See Scala home page: http://www.huygens-fokker.org/scala

Correspondence: Dan Tidhar, Centre for Digital Music, Queen Mary University of London, London, E1 4NS, UK.
E-mail: dan.tidhar@eecs.qmul.ac.uk

DOI: 10.1080/09298215.2010.520720

commonly used in harpsichord recordings. Although the classification task did not rely on any prior knowledge of the score, it did assume that the harpsichord was tuned to one of a pre-defined set of six temperaments. This limitation is reasonable in an experimental setting which is, to the best of our knowledge, the first of its kind. However, in order to obtain an open-ended system which can be used for a public temperament estimation web service, a more general classifier is required, one which is able to cope with both standard and non-standard (i.e. unknown to the system) temperaments. The OMRAS2 project[2] provides a framework for describing and interlinking music related concepts and resources on the Semantic Web, including applications for web-based music signal processing. This framework offers an ideal context for designing and implementing such a generalized, open-ended classifier. In the OMRAS2 framework, we employ Semantic Web technologies. We represent information about music using standardized languages such as Resource Description Framework (RDF)[3] and structures having standardized semantics. For a thorough review of these technologies, languages and their applications in the music domain see Fazekas, Raimond, Jacobson, and Sandler (2010).

Central to our framework is the Music Ontology (Raimond, Abdallah, Sandler, & Frederick, 2007), which provides a basis for communicating musical information on the Web. It is designed to be extensible in order to describe specific sub-domains of musical knowledge such as acoustical features of audio recordings or musicological terms. As part of the work reported here, we provide one such extension, namely, the Temperament Ontology.[4] This ontology provides means for describing the properties of temperaments and the relationships between temperaments and temperament families. The ontology is accompanied by a web service which provides descriptions of an increasing set of individual temperaments. For example, the RDF description of the Vallotti temperament may be accessed at the following URL: http://purl.org/temperament/symbol/Vallotti.

Based on our previous work on temperament estimation, the Temperament Ontology, and various other OMRAS2 components, we developed a web application called TempEst. The application enables users to upload a recording for classification, and receive results in RDF format describing the identified temperament according to the Temperament Ontology. The scope of the classifier is dynamically determined by the set of temperament descriptions provided by the data service using the ontology at any given moment. TempEst includes a simple rule-based inference module which enables it to accommodate temperaments not yet described using the ontology, and to produce useful statements about temperaments even if results are inconclusive due to partial or missing data. A set of inference rules which are described in the section *TempEst inference* below are currently in use.[5] The temperament classification algorithm in TempEst is implemented using Vampy,[6] a Python[7] wrapper for the Vamp audio analysis plugin system (Cannam, 2009), and exposed on the Web as part of the Sonic Annotator Web Application (SAWA). A detailed description of the SAWA system is provided in Fazekas, Cannam, and Sandler (2009).

The remainder of this article is organized as follows. The rest of this section provides a brief introduction to the subject matter of temperament and temperament estimation. The following section presents the Temperament Ontology, describing its design and giving examples of how specific temperaments may be expressed according to the ontology. We then describe TempEst and its various components, including the methods used to infer the properties of arbitrary temperaments. The final section contains conclusions from the work and describes directions for future work.

1.1 Temperament

Theoretical and practical aspects of temperament are covered thoroughly elsewhere (Barbour, 2004/1951; Rasch, 2002; Di Veroli, 2009), so we shall provide only a brief formulation of the problem here. Musical consonance has been explained in different theoretical frameworks throughout the years (Helmholtz, 1863; Lundin, 1947; Terhardt, 1977; Sethares, 2004; Palisca & Moore, 2010), but at least since Pythagoras it has been generally accepted that for musical sounds with harmonic spectra the sensation of consonance correlates to small integer frequency ratios, and specifically superparticular ratios of the form $(n+1)/n$ where $n \leq 5$. Continuous-pitch instruments as well as singers can dynamically adapt their intonation to form perfectly consonant intervals if required. Fixed-pitch instruments such as keyboard, some fretted, and some percussion instruments, need to commit to a certain tuning scheme for the duration of a piece, if not an entire concert. At least in the Western musical tradition, this gives rise to the need for temperament, because one cannot

[2]OMRAS2 web site: http://www.omras2.org

[3]RDF is a data model primarily used on the Semantic Web for expressing simple statements in the form of *subject-predicate-object*. It has several different serialization formats such as XML/RDF and N3. See http://www.w3.org/TR/rdf-primer/ for an introduction.

[4]http://purl.org/ontology/temperament

[5]Due to the dynamic nature of the system, further rules may be added to reflect accumulated experience.

[6]VamPy is available from http://vamp-plugins.org/vampy. html

[7]See www.python.org for details on Python, an increasingly popular language in the scientific community.

accommodate all pure intervals within an integer number of octaves (see e.g. Schroeder, 2009).

Consider 12 fifth steps (each being seven semitones; a pure fifth has a frequency ratio of 3/2) and seven octave steps (each 12 semitones; frequency ratio 2/1). On a keyboard instrument both of these sequences of intervals lead to the same key, despite the fact that $(3/2)^{12} > 2^7$. The ratio between the two sides of this inequality is referred to as the Pythagorean comma, and is roughly equal to 53/52. (Another relevant comma is the ratio between four consecutive fifths and two octaves and a major third, i.e. between $(3/2)^4$ and $(2^2 \times (5/4))$, which is called the Syntonic comma, and is equal to 81/80. For practical tuning purposes, these commas are often conflated and are just referred to as 'the comma'.)

One way of defining particular temperaments is according to the distribution of the comma along the circle of fifths. Equal temperament, for example, can be defined as such where each fifth on the circle is diminished by exactly the same amount, i.e. 1/12 of a (Pythagorean) comma. Determining a temperament can be regarded as an optimization problem, whereby keeping the octaves pure (i.e. 'closing' the circle of fifths) is a major constraint, and various considerations lead to different compromises between pureness of fifths and pureness of major thirds. Among these considerations, are the key, or set of keys which should work well in the given temperament; a temperament is considered to 'work well' for a given key if the most frequent harmonic intervals in the key (major thirds and to some extent fifths, most notably in tonic and dominant positions) are not too far from their pure underlying frequency ratios, and are therefore perceived as consonant. Temperaments which work well for most keys, and are bearable in *all* keys, are normally referred to as 'well' temperaments. Temperaments in which all fifths but one or two are equal to each other, are called regular temperaments. Equal temperament is the only temperament which is both regular and well.

Next, we briefly describe the six temperaments specified in Table 1, namely, twelve-tone equal temperament, Vallotti, fifth-comma, quarter-comma meantone, sixth-comma meantone, and just intonation. In equal temperament, all 12 fifths are equally diminished by 1/12 of a (Pythagorean) comma, resulting in 12 equal semitones, each of frequency ratio $\sqrt[12]{2}$. In a Vallotti temperament, six of the fifths are diminished by 1/6 of a comma each, and the other six fifths are left pure. Equal temperament and Vallotti are graphically described in Figure 1. In the fifth-comma temperament we used, five of the fifths are diminished by a 1/5 comma each, and the remaining seven are pure. In a quarter-comma meantone temperament, 11 of the fifths are shrunk by 1/4 of a comma, and the one remaining fifth is 7/4 of a comma larger than pure. In sixth-comma meantone, 11 fifths are shrunk by 1/6 of a comma, and the one remaining fifth is

Table 1. Deviations (in cents) from equal temperament for the six target temperaments we used in Tidhar et al. (2010). The abbreviated temperament names below stand for: equal temperament, Vallotti, fifth comma, quarter-comma meantone, sixth-comma meantone, and just intonation.

Note	Equal	Vallotti	1/5 C	1/4 CMT	1/6 CMT	Just
C	0	5.9	8.2	10.3	4.9	15.6
C#/D♭	0	0	−1.6	27.4	13.0	−13.7
D	0	2	2.7	3.4	1.6	−2
D#/E♭	0	3.9	2.3	20.5	9.8	−9.8
E	0	−2.0	2.0	−3.4	−1.6	2.0
F	0	7.8	6.3	13.7	6.5	13.7
F#/G♭	0	−2.0	−3.5	−10.3	−4.9	−15.6
G	0	3.9	5.5	6.8	3.3	17.6
G#/A♭	0	2.0	0.4	24.0	11.4	−11.7
A	0	0	0	0	0	0
A#/B♭	0	5.9	4.3	17.1	8.1	11.7
B	0	−3.9	−0.8	−6.8	−3.3	3.9

5/6 comma larger than pure. In the just-intonation temperament we used all tones are calculated as integer ratios, given by the following vector which represents twelve chromatic tones above the reference A: (16/15, 9/8, 6/5, 5/4, 4/3, 45/32, 3/2, 8/5, 5/3, 9/5, 15/8, 2/1). The deviations in cents[8] from equal temperament of the five other temperaments we use are given in Table 1. Apart from being relatively common, this set of six temperaments represents different categories: equal temperament is both well and regular, just intonation is neither well nor regular, Vallotti and fifth-comma are well and irregular, and the two variants of meantone (quarter and sixth comma) are regular but not well.

1.2 Temperament estimation

Temperament estimation was first formulated as a music signal processing task by Tidhar et al. (2010). Building a system to classify musical recordings by temperament presents particular signal processing challenges. First, the differences between temperaments are small, of the order of a few cents. For example, if A = 415 Hz is used as the reference pitch, then middle C might have a frequency of 246.76, 247.46, 247.60, 247.93, 248.23 or 248.99 Hz, based on the six example temperaments described above (see Table 1). To resolve these frequencies in a spectrum, a window of several seconds duration would be required, but this introduces other problems, since musical notes are not stationary and generally do not last this long.

[8]A cent is one hundredth of a semitone, i.e. one twelve-hundredth of an octave.

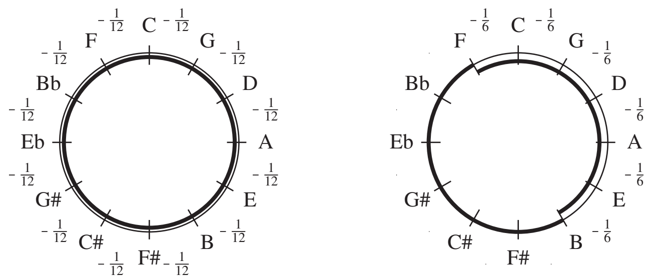

Fig. 1. Circle of fifths representations for Equal Temperament (left) and Vallotti (right). The fractions specify the distribution of the Pythagorean comma between the fifths (fifths without a fraction are pure).

The second problem is that in musical recordings, notes rarely occur in isolation. There are almost always multiple notes sounding simultaneously, and this has the potential to bias any frequency estimates. To make matters worse, the intervals which are often favoured in music are those where the partials coincide. In particular, it can be difficult to discern whether a spectral peak is a fundamental frequency or a partial of another fundamental. The ability to distinguish between these cases is crucial to accurate pitch estimation and thus also to successful temperament classification.

For example, if an A_2 at 110 Hz is played, the spectrum will also have a peak at 330 Hz, which is the third harmonic of the note A_2, but also (if the fifth is pure) the fundamental frequency of an E_4. In many temperaments however, the actual note E_4 will have a frequency different from 330 Hz (e.g. 329.6 Hz in equal temperament), so the estimation of E_4 would be biased if this partial was used in its estimation. We therefore require a method for distinguishing peaks corresponding to fundamental frequencies from those which are caused by higher harmonics. This would not be the case if we did assume knowledge of the score, but this was deliberately avoided in our original formulation of the problem, in order to increase generality, which is important for practical applications. In the context of a web service such as TempEst, the benefits of not requiring any knowledge of the score are apparent.

To avoid the bias caused by overlapping partials, while circumventing the problem of full polyphonic transcription,[9] we introduced the concept of *conservative transcription* (Tidhar et al., 2010), which entails ignoring frequencies which could potentially be harmonics of lower co-occurring frequencies. We showed that conservative transcription is applicable in practical situations, and that it can improve temperament estimation in recordings of typical harpsichord music. Figure 2 demonstrates the result of conservative transcription on the first two bars of J.S. Bach's prelude in C Major from the first volume of the *Well-Tempered Clavier*, BWV 846.

Using various pitch estimation algorithms, we obtained satisfactory results (23 out of 24 real harpsichord recordings and 24 out of 24 sythesized harpsichord recordings were classified correctly). When applied with suitable parameters and enhanced by methods such as conservative transcription, several different algorithms performed well with only minor differences. For our current implementation of TempEst we use the Quadratically Interpolated Fast Fourier Transform (QIFFT) (Smith, 2008).

2. The temperament ontology

In order to formally express individual temperaments, as well as temperament classes and properties, and in order to facilitate reasoning and inference, we developed the Temperament Ontology, which extends and complements the Music Ontology. The ontology is publicly available online.[10]

The Temperament Ontology is designed to accommodate extensions and modifications by the user community, so that additional individual temperaments as well as richer details and properties can be added as the ontology develops. The current ontology is admittedly biased toward historical temperaments, but this bias may change according to the user community's interests.

2.1 Ontology design

Temperaments can be classified according to different criteria. In order to account for all of these, and to accommodate additional criteria in the future, our model consists of a shallow hierarchy, in which all classes are subclasses of `Temperament` (see Figure 3). While hierarchical relations are kept to a minimum due to possible ambiguities and in order to facilitate extensibility, more complex relations can be expressed, for example, using multiple class memberships.

Since there is more than one way to associate tuning systems with their properties, we treat temperament descriptions as concepts as well, and use reification[11] which keeps the model open and extensible. The `Temperament Description` class currently allows two alternative temperament specification methods. Individual temperaments can be assigned with either or with both, and additional specification methods can be easily accommodated.

This design allows a high degree of expressive flexibility, and is easily extensible. However, this

[9]Polyphonic transcription is still regarded an unsolved problem (Casey et al., 2008; Klapuri, 2009).

[10]http://purl.org/ontology/temperament

[11]In this context: representing a relation by an object in order to reason about the relation itself.

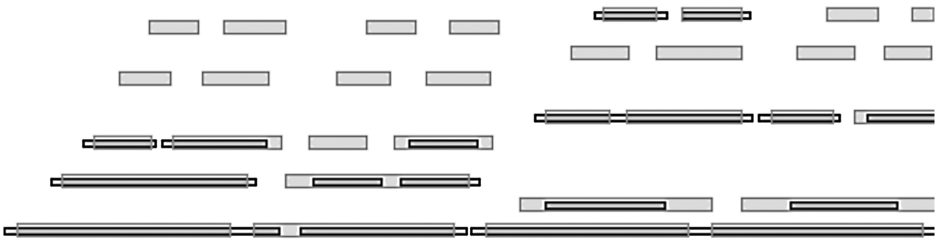

Fig. 2. The first two bars of J.S. Bach's *Prelude in C Major* are shown in piano-roll notation (the horizontal axis represents time and the vertical axis represents pitch). The actual notes are rendered in light grey, with the Conservative Transcription superimposed in dark borders.

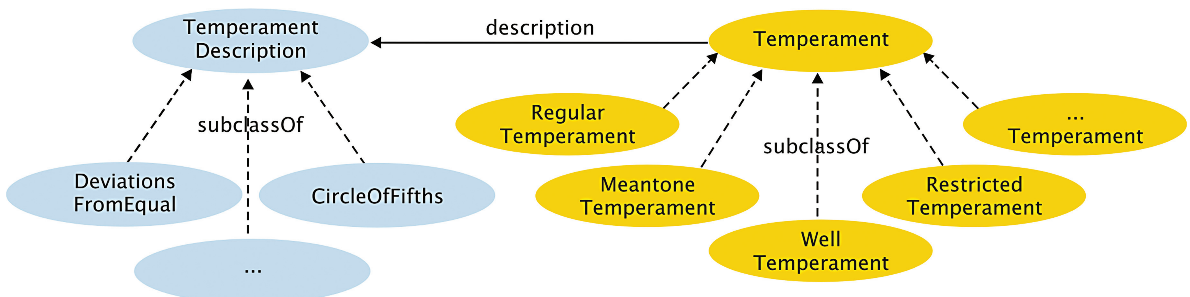

Fig. 3. The main classes of the Temperament Ontology, including some of the Temperament subclasses (right) and the parallel Temperament Description class (left).

structure does not enforce semantic coherence, and it is up to the modellers to avoid contradictions such as a single individual being defined as a Well Temperament (i.e. a temperament that works well for all keys) and as a Restricted Temperament (i.e. a temperament that does not) at the same time.

2.2 Temperament descriptions

We model two forms of temperament specifications which are subclasses of the Temperament Description class, reflecting two of the commonly used methods. Other methods can easily be accommodated as additional subclasses (see Figure 3). The two methods are deviations (in cents) from equal temperament (see Table 1), and deviations from pure fifths (frequency ratio 3:2) which are specified as signed fractions of a comma (see Figures 3 and 4). Various commas are currently modelled including the Pythagorean and the Syntonic ones, and others can be added.

Figure 4 shows a segment of an RDF graph describing Vallotti using the temperament description type `CircleOfFifths`. The graph expresses the fact that in this particular temperament, the fifth from C to G is diminished by one sixth of the Pythagorean comma. The description type `CircleOfFifths` encodes specific temperaments in terms of the deviations of each fifth from a pure fifth, expressed as signed fractions of the Pythagorean comma. In Listing 1 (in Section 3.3 below), we also provide a part of a serialized RDF document showing a more detailed example of describing a temperament using both supported methods.

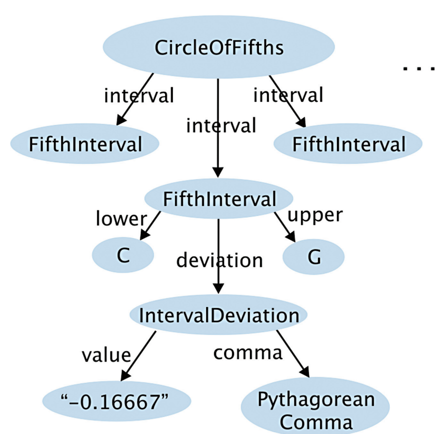

Fig. 4. A segment of a temperament description of type CircleOfFifths for Vallotti.

2.3 Individual temperaments

Temperaments are modelled as individual instances of at least one class of the Temperament class hierarchy (see Figure 3), and would normally inherit from more than one subclass according to their properties. The ontology currently includes individuals for many of the commonly-used historical temperaments and others can naturally be added as required.

3. TempEst

TempEst is a web application built on Semantic Web technologies. It enables users to upload harpsichord recordings and retrieve information about the instrument's tuning. TempEst integrates temperament estimation methods with temperament profiles expressed using the Temperament Ontology, and with some degree of inference over the measurement data expressed in RDF. The system will be developed further as the user community grows, as a result of further ontology development (e.g. additional temperaments) as well as further development of the inference rules.

3.1 System components

TempEst is constructed from a number of modular components, combining software developed in the OMRAS2 project with some third-party libraries. In this section, we briefly describe the most important parts of the system, which allow easy exposure of the music processing algorithms as a web application.

3.1.1 Vamp plugins and Vampy

Vamp is a plugin system for audio feature extraction using a dedicated API (Cannam, 2009) developed in OMRAS2 for this purpose. Vamp plugins accept audio input and produce structured data output as described in the C++ SDKs provided for plugin and host development. A large number of Vamp plugins are available[12] for audio analysis, ranging from low-level feature extractors to plugins computing high-level annotations such as structural segmentation.

Vampy provides an easy to use interface between the Python programming language and Vamp plugin hosts. Vampy itself is implemented as a Vamp plugin which may be installed in the usual manner.[13] When it is installed, any appropriately structured Python scripts found in its script directory will be presented as if they were individual Vamp plugins for any Vamp host to use. Vampy allows for taking advantage of Python's high-level and efficient software libraries such as NumPy and SciPy for numerical computation and signal processing.

[12]Vamp plugins: http://vamp-plugins.org/download.html

[13]See http://vamp-plugins.org/download.html#install

Vamp and Vampy plugins require a host environment to run, for example Sonic Visualiser[14] and Sonic Annotator.[15] In the TempEst system, we implemented the previously described temperament estimation algorithm as a Vampy plugin. Besides the ability to use the numerical libraries available for Python, this dynamic environment enables us to generate an RDF description from numerical results, and feed it into an inference engine such as Cwm[16] (Closed World Machine), which was developed as part of the Semantic Web Application Platform (Berners-Lee et al., 2006).

3.1.2 Web application

TempEst is implemented as part of the SAWA[17] framework. This framework uses Semantic Web technologies to offer services to audio researchers and interested end-users. SAWA can be used to upload a set of audio files and perform feature extraction in batch mode using Sonic Annotator, its efficient underlying Vamp plugin host. SAWA uses its own web server based on the Cherrypy[18] library. This allows for writing HTTP request handlers as ordinary methods defined within a web application class, making it straightforward to accept audio files or data entered into HTML forms, as well as serving dynamically generated web pages and publishing data received from other system components. A detailed description of the components of SAWA is available in Fazekas et al. (2009).

The main SAWA application was designed for general purpose web-based audio feature extraction. However, this has been extended to incorporate other applications with similar needs such as SAWA-Recommender,[19] a query-by-example music retrieval application. TempEst[20] is the latest addition to this framework. Using TempEst, one may upload harpsichord recordings for analysis, and receive a detailed description of the temperament(s) used in the recordings. This includes the temperament identified by the classifier, some known properties of the temperament resulting from inference on measurement data,[21] and the measurements themselves on which the classification and inference are based. These results are expressed using the Temperament Ontology and published as RDF documents.

[14]Available from: http://sonicvisualiser.org/

[15]Available from: http://omras2.org/SonicAnnotator

[16]Available from: http://www.w3.org/2000/10/swap/doc/cwm.html

[17]Available at http://www.isophonics.net/sawa/

[18]Available at: http://www.cherrypy.org/

[19]Available at: http://www.isophonics.net/sawa/rec

[20]Available at: http://www.isophonics.net/sawa/tempest

[21]The signal processing component of TempEst produces *deviations from equal temperament*. This description format is converted to the *circle of fifths* before inferring additional properties of the temperament, such as regularity.

3.2 TempEst inference

In order to deal with a potentially unlimited number of different temperaments, the classification framework needs to be extended beyond a particular set of predefined temperaments. By linking the classifier directly to the ontology, its target set becomes extensible, since any adjustments or extensions to the ontology take immediate effect on the classification scope. At any given moment, however, this scope is inevitably limited by the coverage of the ontology. The TempEst inference module attempts to generate useful information about any given temperament, even if it is not covered by the ontology. This is done in the form of a set of implication rules which deduce information from heuristic measurements as described below. The input to the rules, as well as the rules themselves, are currently regarded as proof of concept. One of our research goals is to explore inference algorithms which closely reflect the open-ended structure of the ontology, such that for a given temperament, temperament classes are determined with each inference step as the identification becomes more specific. It remains an open question whether it is easier to identify a specific known temperament and then derive its properties, or to determine these properties first and use them as a basis for identification.

For example, in the former approach, upon encountering a Vallotti temperament, prior to identifying it as Vallotti, the inference system would deduce its irregularity, its wellness, and the fact that it is a sixth-comma temperament, and only then determine its identity. In the latter approach (as implemented in the current system), however, the identification is based on nearest neighbour classification, independently of any deduced characteristics. We currently implement two sets of rules, based on heuristically defined thresholds as described below.

3.2.1 Tolerance thresholds for pureness and equality

To facilitate the inference rules described below, we determine how close to 3:2 a frequency ratio needs to be to be regarded as a pure fifth, and how close two frequency ratios need to be in order for them to be considered ‘equal’. We derive these heuristically from our original classification experiment, by checking which are the most distant fifths still classified as equal to one another in any of the temperaments, and which fifths are farthest from 3:2 but still classified as pure. Although the classifier does not directly provide any information about fifths, this information is implied by the estimated temperament, e.g. in Vallotti, $C\# - G\#$ is pure, $F - C$ and $C - G$ are equal to one another, and so on. The thresholds determined this way are used as defaults (currently set to 0.02 Pythagorean Comma or approximately 0.47 cents), which can be overridden by the user issuing the query.

3.2.2 Rules based on pure fifths

The following rules deduce temperament characteristics from the number of fifths which are considered pure, i.e. fall within the vicinity of 3:2 defined by the above-mentioned threshold. The total number of pure fifths can theoretically vary between 0 and 11 (12 is obviously not possible within the circle of fifths). Note that in order for the output to be expressed in RDF, the consequents have to be positively formulated, e.g. rather than concluding that a temperament is not a well temperament (in the third rule below), we use the class of restricted temperaments modelled by the ontology.

1. *If* any fifth is pure, *then* the temperament is unequal
2. *If* more than two fifths are pure, *then* the temperament is irregular
3. *If* more than 8 fifths are pure, *then* the temperament is restricted
4. *If* 11 fifths are pure, *then* the temperament is Pythagorean
5. *If* exactly 6 fifths are pure, *then* the temperament is 6th-comma-like

3.2.3 Rules based on sets of equal fifths

The following rules deduce temperament characteristics from the largest set of fifths that are equal to each other, where equality is defined by the above-mentioned threshold. The number of equal fifths can vary between 0 and 12.

1. *If* all fifths are equal, *then* the temperament is equal temperament
2. *If* at least 10 fifths are equal, *then* the temperament is regular
3. *If* less then 10 fifths are equal, *then* the temperament is irregular

3.3 Query interface and system output

TempEst queries are based on audio material which is uploaded via the SAWA interface. The user is required to specify an approximate tuning reference frequency.[22]

[22]The system is capable of fine-tuning the reference frequency within a vicinity of ± 40 cents, but requires a reference to avoid ambiguities between, for example, a $B\flat$ in baroque pitch and an A in modern pitch, as both can potentially be associated with a frequency in the vicinity of 440 Hz.

```
@prefix : </> .
@prefix pc: <http://purl.org/ontology/temperament/pitchclass/> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix tm: <http://purl.org/ontology/temperament/> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .

:temperament_0 a tm:Temperament ,
        tm:Vallotti ;
    tm:description :description_00 ,
        :description_01 .

:description_00 a tm:DeviationsFromEqual ;
    rdfs:label "Deviations from equal temperament in cents." ;
    tm:deviation_from_equal
        [ a tm:PitchClassDeviation ;
            tm:pitch_class pc:G ;
            tm:value "3.90984723"^^xsd:float ] ,
        [ a tm:PitchClassDeviation ;
            tm:pitch_class pc:Gs ;
            tm:value "1.95488264"^^xsd:float ] ,
            . . .

:description_01 a tm:CircleOfFifths ;
    rdfs:label """Deviations from pure fifth given by a fraction of
        Pythagorean Comma.""" ;
    tm:interval
        [ a tm:FifthInterval ;
            tm:deviation [ a tm:IntervalDeviation ;
                    rdfs:label "-1/6" ;
                    tm:comma tm:PythagoreanComma ;
                    tm:value "-0.166666666667"^^xsd:float ] ;
            tm:lower pc:D ;
            tm:upper pc:A ] ,
        [ a tm:FifthInterval ;
            tm:deviation [ a tm:IntervalDeviation ;
                    rdfs:label "0" ;
                    tm:comma tm:PythagoreanComma ;
                    tm:value "0.0"^^xsd:float ] ;
            tm:lower pc:B ;
            tm:upper pc:Fs ] ,
            . . .
```

Listing 1. RDF description of Vallotti temperament (abbreviated).

Users may optionally also specify alternative tolerances for determining which fifths are considered pure and which are considered equal by the system. These parameters then override the system's defaults. The user interface for configuring TempEst's parameters is shown in Figure 5. The output, in RDF format, consists of an estimated temperament, the estimated pitch-class frequencies, as well as any output produced by the inference module.

As the output of the system, we may obtain the RDF document using N3 notation[23] shown in Listing 1. This describes Vallotti temperament using both temperament

[23]For an explanation of RDF syntax and serialization formats including N3 see http://www.w3.org/TeamSubmission/turtle/ and http://www.w3.org/DesignIssues/Notation3.html. See also Fazekas et al. (2010) in the present issue for examples and applications in the music domain.

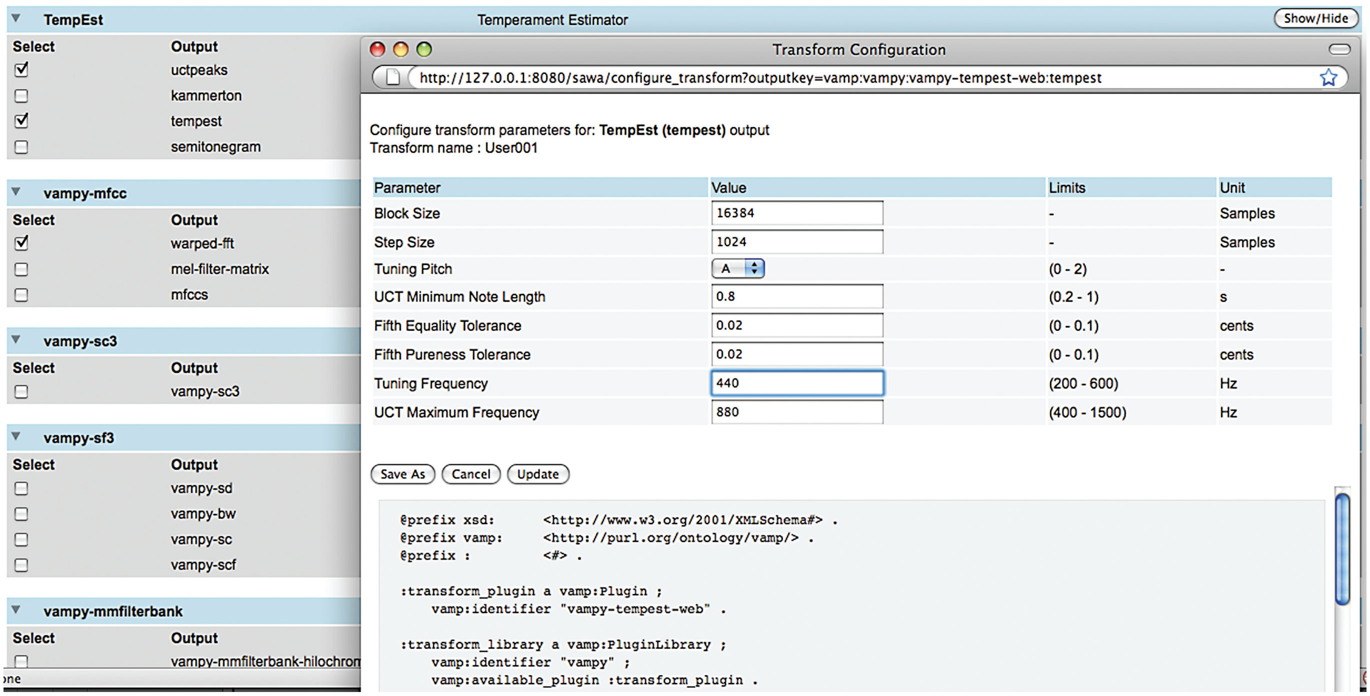

Fig. 5. Temperament Estimator Configuration Interface in SAWA.

description methods. Note that the actual document includes 12 elements of type `PitchClassDeviation` for each description type `DeviationsFromEqual`, and 12 elements of type `FifthInterval` for each description type `CircleOfFifths`; due to space limitations we only show two of each.

4. Conclusion and future directions

TempEst combines signal processing techniques with semantic computing and Semantic Web technology, and is motivated by musicological research goals. It illustrates the use of OMRAS2 technology in the type of end-user application which was targeted by the project. TempEst makes extensive use of OMRAS2 components such as the Music Ontology and SAWA. In this paper, we reviewed our previous work in temperament classification, and have extended it with the Temperament Ontology, an inference module, and a web application. The resulting system incorporates many of the OMRAS2 goals, most notably full integration of digital signal processing and semantic information processing to perform an MIR task in a Semantic Web environment.

We plan to facilitate collection of usage data, and although copyright issues dictate that we cannot maintain the uploaded recordings, we will be able to collect and publish some statistics describing the recognized temperaments. These may include, for example, information about the most popular temperaments used with the system, as well as co-occurrence analysis of individual temperaments. At a later stage, we plan to implement an optional log-in mechanism, which would enable users to maintain their data across sessions, and would enable us to collect more detailed information about the users and their queries.

The music information retrieval task of temperament estimation can be extended to include keyboard instruments other than the harpsichord, as well as recordings involving additional instruments. Temperament, rather than intonation, is a potentially relevant characteristic of any recording which involves at least one fixed-pitch instrument, such as recordings of early music ensembles with keyboard continuo.

Further research will focus on the development of the semantic inference module based on analysis of accumulated data. In particular, we are interested in inference rules which reflect the structure of the ontology, by corresponding to particular temperament classes. Since TempEst is directly linked to the Temperament Ontology, it has the potential to expand as more data is gathered and analysed. The ontology and the temperament data service are designed to grow and to accommodate further individual temperaments as well as temperament classes, and will thus be adjusted to reflect the interests of its user community. As TempEst matures, with richer ontology, data service, and inference components, it will be suitable for analysing large audio collections, and facilitate in-depth studies of temperament including, for example, diachronic analysis of recordings.

Acknowledgements

The authors acknowledge the support of the School of Electronic Engineering and Computer Science, Queen Mary University of London, and the EPSRC-funded ICT project OMRAS2 (EP/E017614/1).

References

Barbour, J.M. (2004/1951). *Tuning and Temperament, A Historical Survey*. New York: Dover.

Berners-Lee, T., Kolovski, V., Connolly, D., Hendler, J., Kolovski, V., & Scharf, Y. (2006). A reasoner for the web. Retrieved January 2010 from http://www.w3.org/2000/10/swap/doc/paper/

Cannam, C. (2009). The Vamp audio analysis plugin API: A programmer's guide. Retrieved from http://vamp-plugins.org/guide.pdf

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

Di Veroli, C. (2009). *Unequal Temperaments: Theory, History, and Practice*. Bray, Ireland: Bray Baroque.

Fazekas, G., Cannam, C., & Sandler, M. (2009). Reusable metadata and software components for automatic audio analysis. In *Proceedings of the IEEE/ACM Joint Conference on Digital Libraries (JCDL'09) Workshop on Integrating Digital Library Content with Computational Tools and Services*, Austin, TX, USA.

Fazekas, G., Raimond, Y., Jacobson, K., & Sandler, M. (2010). An overview of Semantic Web activities in the OMRAS2 project. *Journal of New Music Research 39*(4), 295–311.

Helmholtz, H. v. (1863). *Die lehre von den tonempfindungen als physiologische grundlage für die theorie der musik (On the Sensations of Tone as a Physiological Basis for the Theory of Music)*. Richmond, VA: Brunswick.

Klapuri, A.P. (2009). A method for visualizing the pitch content of polyphonic music signals. In *Proceedings of the 10th International Society for Music Information Retrieval Conference (ISMIR 2009)*, Kobe, Japan, pp. 615–620.

Lehman, B. (2005). Bach's extraordinary temperament: Our Rosetta Stone—1. *Early Music*, *33*(1), 3–23.

Lundin, R.W. (1947). Towards a cultural theory of consonance. *Journal of Psychology*, *23*, 45–49.

Palisca, C.V., & Moore, B.C.J. (2010). Consonance. Retrieved from http://www.grovemusic.com/

Raimond, Y., Abdallah, S., Sandler, M., & Frederick, G. (2007). The music ontology. In *Proceedings of the 7th International Conference on Music Information Retrieval (ISMIR 2007)*, Vienna, Austria, pp. 417–422.

Rasch, R. (2002). Tuning and temperament. In T. Christensen (Ed.), *The Cambridge History of Western Music Theory* (pp. 193–222). Cambridge: Cambridge University Press.

Schroeder, M.R. (2009). *Number Theory in Science and Communication: With Applications in Cryptography, Physics, Digital Information, Computing, and Self-similarity*. Berlin: Springer.

Sethares, W. (2004). *Tuning, Timbre, Spectrum, Scale*. Berlin: Springer.

Smith, J.O. (2008). *Spectral Audio Signal Processing, October 2008 Draft* [Online book]. Retrieved from https://ccrma. stanford.edu/sasp/

Terhardt, E. (1977). The two-component theory of musical consonance. In E. Evans & J. Wilson (Eds.), *Psychophysics and Physiology of Hearing*, pp. 381–390. London: Academic.

Tidhar, D., Mauch, M., & Dixon, S. (2010). High precision frequency estimation for harpsichord tuning classification. In *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP2010)*, Dallas, TX, USA, pp. 61–64.