

# PAPER TEMPLATE FOR ISMIR 2014

**First author**

Affiliation1

author1@ismir.edu

**Second author**

**Retain these fake authors in  
submission to preserve the formatting**

**Third author**

Affiliation3

author3@ismir.edu

## ABSTRACT

The abstract should be placed at the top left column and should contain about 150-200 words.

## 1. INTRODUCTION

This template includes all the information about formatting manuscripts for the ISMIR 2014. Please follow these guidelines to give the final proceedings a uniform look. If you have any questions, please contact the Conference Management. This template can be downloaded from the ISMIR 2014 web site (<http://ismir2014.ismir.net>).

## 2. PAGE SIZE

The proceedings will be printed on portrait A4-size paper (21.0cm x 29.7cm). All material on each page should fit within a rectangle of 17.2cm x 25.2cm, centered on the page, beginning 2.0cm from the top of the page and ending with 2.5cm from the bottom. The left and right margins should be 1.9cm. The text should be in two 8.2cm columns with a 0.8cm gutter. All text must be in a two-column format. Text must be fully justified.

## 3. TYPESET TEXT

### 3.1 Normal or Body Text

Please use a 10pt (point) Times font. Sans-serif or non-proportional fonts can be used only for special purposes, such as distinguishing source code text.

The first paragraph in each section should not be indented, but all other paragraphs should be.

### 3.2 Title and Authors

The title is 14pt Times, bold, caps, upper case, centered. Authors' names are omitted when submitting for double-blind reviewing. The following is for making a camera-ready version. Authors' names are centered. The lead author's name is to be listed first (left-most), and the co-authors' names after. If the addresses for all authors are the same, include the address only once, centered. If the

authors have different addresses, put the addresses, evenly spaced, under each authors' name.

### 3.3 First Page Copyright Notice

Please include the copyright notice exactly as it appears here in the lower left-hand corner of the page. It is set in 8pt Times.

### 3.4 Page Numbering, Headers and Footers

Do not include headers, footers or page numbers in your submission. These will be added when the publications are assembled.

## 4. FIRST LEVEL HEADINGS

First level headings are in Times 10pt bold, centered with 1 line of space above the section head, and 1/2 space below it. For a section header immediately followed by a subsection header, the space should be merged.

### 4.1 Chord Recognition

Despite being one of the oldest MIREX tasks, evaluation methodology and metrics for automatic chord recognition is an ongoing topic of discussion. Several recent articles address issues and concerns with vocabularies, comparison semantics, and other lexicographical challenges unique to chord recognition []. Ultimately, the source of this difficulty stems from the inherent subjectivity in "spelling" a chord name and the level of detail a human observer can provide in a reference annotation [?]. As a result, a consensus has yet to be reached regarding the single best approach to comparing two sequences of chord labels, and instead are often compared over a set of rules, e.g Major-Minor, Sevenths, with or without inversions, and so on.

Thanks to the previous efforts of Harte [], text representations of chord labels adhere to a standardized format, consisting of a root, quality, extensions, and a bass note; of these, only the root is strictly required. However, in order to efficiently compare chords in a variety of different ways, it is helpful to first translate a given chord label  $C$  into a numerical representation, shown in Figure ?? . In this example, a  $G : 7(9)/5$  is mapped to split into 4 pieces of information: one, the root is mapped to an absolute pitch class  $\mathcal{R}$ , in  $[0, 11]$ , where  $C \rightarrow 0$ ,  $C\sharp/D\flat \rightarrow 1$ , etc; two, the quality is mapped to a root-invariant 12-dimensional bit vector  $\mathcal{Q}$  by setting the scale degrees of the quality; three, any extensions are applied (via addition or omission) to

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page.

© 2014 International Society for Music Information Retrieval.

the quality bit vector as scale degrees in a single octave, resulting in pitch vector  $\mathcal{P}$ ; and four, the bass interval (5) is translated to the relative scale degree in semitones  $\mathcal{B}$ . Note that the add-9 is rolled into a single octave as an add-2. This is a matter of convenience as extended chords (9's, 11's or 13's) are traditionally resolved to a single-octave equivalent, but the bit-vector representation could be easily expanded to represent such information.

Having gone through this bit of effort, it is now straightforward to compare chords along the five rules used in MIREX 2013:

1. Root:
  - (a)  $\mathcal{R}_{est} == \mathcal{R}_{ref}$
  - (b)  $\forall \mathcal{Q}_{ref}$
2. Major-Minor: Rule 1.a, plus
  - (a)  $\mathcal{Q}_{est} == \mathcal{Q}_{ref}$
  - (b)  $\mathcal{Q}_{ref} \in \{Maj, min\}$
3. Major-Minor w/Inversions: Rule 2, plus
  - (a)  $\mathcal{B}_{ref} \in \mathcal{Q}_{ref}$
4. Sevenths: Rule 1.1, plus
  - (a)  $\mathcal{Q}_{est} == \mathcal{Q}_{ref}$
  - (b)  $\mathcal{Q}_{ref} \in \{Maj, min, Maj7, min7, 7\}$
5. Sevenths w/Inversions: Rule 4, plus
  - (a)  $\mathcal{B}_{ref} \in \mathcal{Q}_{ref}$

Following recent trends in MIREX, an overall score is computed by weighting each comparison by the duration of its interval, over all intervals; stated another way, this is the piecewise continuous-time integral of the intersection of two chord sequences,  $(\mathcal{C}_{ref}, \mathcal{C}_{est})$ , expressed as follows:

$$S(\mathcal{C}_{ref}, \mathcal{C}_{est}) = \frac{1}{T} \int_{t=0}^T \mathcal{C}_{ref}(t) == \mathcal{C}_{est}(t) \quad (1)$$

Here, this is achieved here by forming the union of the boundaries in each sequence, and summing the time intervals of the correct ranges. Note that equivalence is subject to one of the rules defined previously.

Finally, the total score over a set of  $N$  items is given by a discrete summation, where the importance of each score,  $S_n$ , is weighted by the duration,  $T_n$ , of each annotation:

$$S_{total} = \frac{\sum_{n=0}^N T_n * S_n}{\sum_{n=0}^N T_n} \quad (2)$$

## 4.2 Melody Extraction

Melody extraction, or predominant fundamental frequency ( $f_0$ ) estimation, involves the estimation of the pitch ( $f_0$ ) sequence of the melody from a polyphonic music recording [1]

String value	Numeric value
Hello ISMIR	2014

**Table 1.** Table captions should be placed below the table.

### 4.2.1 Third and Further Level Headings

Third level headings are in Times 10pt italic, flush left, with 1/2 line of space above the section head, and 1/2 space below it. The first letter of each significant word is capitalized.

Using more than three levels of headings is highly discouraged.

## 5. FOOTNOTES AND FIGURES

### 5.1 Footnotes

Indicate footnotes with a number in the text.<sup>1</sup> Use 8pt type for footnotes. Place the footnotes at the bottom of the page on which they appear. Precede the footnote with a 0.5pt horizontal rule.

### 5.2 Figures, Tables and Captions

All artwork must be centered, neat, clean, and legible. All lines should be very dark for purposes of reproduction and art work should not be hand-drawn. The proceedings are not in color, and therefore all figures must make sense in black-and-white form. Figure and table numbers and captions always appear below the figure. Leave 1 line space between the figure or table and the caption. Each figure or table is numbered consecutively. Captions should be Times 10pt. Place tables/figures in text as close to the reference as possible. References to tables and figures should be capitalized, for example: see Figure ?? and Table 1. Figures and tables may extend across both columns to a maximum width of 17.2cm.

## 6. EQUATIONS

Equations should be placed on separated lines and numbered. The number should be on the right side, in parentheses.

$$E = mc^2 \quad (3)$$

## 7. CITATIONS

All bibliographical references should be listed at the end, inside a section named "REFERENCES," numbered and in alphabetical order. Also, all references listed should be cited in the text. When referring to a document, type the numbering square brackets [?] or [?, ?, ?].

<sup>1</sup> This is a footnote.

## 8. REFERENCES

- [1] J. Salamon, E. Gómez, D. P. W. Ellis, and G. Richard. Melody extraction from polyphonic music signals: Approaches, applications and challenges. *IEEE Signal Processing Magazine*, 31(2):118–134, Mar. 2014.