Post-processing	none	re-classify	add rule 1	add rule 2
Performance	75.7%	76.3%	76.6%	76.8%

Table 4: Evaluation of the post-processing strategies on the validation dataset using the MFCC descriptors. Performance is computed as the percentage of time that the classification and the manual annotation coincides.

Post-processing	none	all
Performance	77.6%	78.5%

Table 5: Performance of the system on the testing dataset with no post-processing and performing all the post-processing strategies proposed, computed as the percentage of time that the classification and the manual annotation coincides.

Finally the system was used to process the testing set achieving a performance similar to that obtained in the validation set. Table 5 shows the results with no post-processing and considering all the post-processing strategies.

5. Discussion and conclusions

Analysis of the results obtained indicate that those descriptors that model the spectral content of the audio signal were the most appropriate for the problem (MFCC, LFPC, Spectral, PLPC). It is interesting that such a simple descriptor as LFPC outperformed all other features sets but MFCC, and that the general purpose Spectral outperformed PLPC. The results confirm that HC is not able to discriminate singing voice sounds from other harmonic musical instruments. The poor performance obtained with the Pitch descriptors is due to the utilization of a monophonic fundamental frequency estimation algorithm. We plan to apply other pitch estimation methods in our future work that could deal with polyphonic audio and to develop pitch descriptors that exploit singing voice pitch contour distinctive features such as intonation. Regarding MFCC, the feature selection performed points out that considering delta coefficients can boost performance (2% on the training set). However, they are generally not used when applying MFCC to this type of problem [Li and Wang, 2007] [Tsai and Wang, 2006]. Classification performance decreased significantly in validation and testing compared to 10-fold CV on the training set, which is not surprising because of the different origins and data variances of the databases. Additionally, the developed system roughly divides the audio file in fragments, so given our validation approach, we can take these results as worst-case or lower-bound estimations of the system performance.

We have studied the singing voice detection problem in music audio files by a statistical classification approach and we have compared, under equivalent conditions, the performance of several types of acoustic descriptors reported to be used for the problem. The results obtained confirm the usefulness of MFCC for this problem. As an outcome of our study, an effective singing voice detection system to process popular music audio files with a reduced set of descriptors has been developed. It is difficult to compare the performance achieved with other research work because there is no standard dataset for evaluation, but results obtained are promising and similar to the ones reported. Although our primary intention was to compare already used descriptors for this task, we have attempted to combine different descriptors as well as different classifiers. The overall classification performance was not improved so the results are not included. Also some other descriptors were tested without success. Our future work will follow this direction as it is reasonable to expect better results by combining different sources of information.

6. Acknowledgments

This work was carried out during an internship at the Music Technology Group (MTG) financially supported by Comisión Sectorial de Investigación Científica, UdelaR. First author is very grateful to all the people at the MTG for their support and to Dr. Alvaro Pardo for his careful reading of this article and his suggestions.

References

A. de Cheveignè, H. Kawahara (2002). YIN, a fundamental frecuency estimator for speech and music. *Journal Acoustic Society of America*, 111:1917–1930.

Berenzweig, A. and Ellis, D. (2001). Locating singing voice segments within music signals. *Proc. IEEE Workshop on Apps. of Sig. Proc. to Acous. and Audio, Mohonk NY, October 2001*, page 4pp.

Berenzweig, A., Ellis, D., and Lawrence, S. (2002). Using voice segments to improve artist classification of music. *AES 22nd International Conference*.

Chilton, T. (1999). Speech analisys. School of Electronic and Physical Sciences, University of Surrey.

Chou, W. and Gu, L. (2001). Robust singing detection in speech/music discriminator design. *International Conference on Acoustics, Speech, and Signal Processing*.

Cook, P. R. (1990). *Identification of control parameters in an articulatory vocal tract model, with applications to the synthesis of singing*. PhD thesis, Stanford Univ., Stanford, CA.

Ellis, D. P. W. (2005). PLP and RASTA (and MFCC, and inversion) in Matlab. Online web resource: `http://www.ee.columbia.edu/~dpwe/resources/matlab/rastamat/`.

Gerhard, D. (2002). Pitch-based acoustic feature analysis for the discrimination of speech and monofonic singing. *Journal of the Canadian Acoustical Assosiation*, pages 152–153.

Herrera, P., Klapuri, A., and Davy, M. (2006). Automatic Classification of Pitched Musical Instrument Sounds. In Klapuri, A. and Davy, M., editors, *Signal Processing Methods for Music Transcription*, pages 163–200. Springer, New York.

Kim, Y. E. and Whitman, B. P. (2002). Singer identification in popular music recordings using voice coding features. *In Proceedings of Interational Conference on Music Information Retrieval*, pages 164–169. Paris, France.

Li, Y. and Wang, D. (2007). Separation of singing voice from music accompaniment for monaural recordings. *IEEE Transactions on Audio, Speech and Language Processing*.

Li, Y. and Wang, D. L. (2005). Separation of singing voice from music accompaniment for monaural recordings. Technical Report OSU-CISRC-9/05-TR61, Department of Computer Science and Engineering, The Ohio State University, Columbus, Ohio, USA.

Maddage, N. C., Wan, K., Xu, C., and Wang, Y. (June 2004). Singing voice detection using twice-iterated composite fourier transform. *Multimedia and Expo, 2004. ICME '04. 2004 IEEE International Conference*, pages Page(s):1347 – 1350 Vol.2.

Maddage, N. C., Xu, C., and Wang, Y. (2003). A svm-based classification approach to musical audio. *Proc. ISMIR*.

Martin, K. D. (1999). *Sound-Source Recognition: A Theory and Computational Model*. PhD thesis, MIT. Cambridge, MA.

New, T. L., Shenoy, A., and Wang, Y. (2004). Singing voice detection in popular music. Technical report, Department of Computer Science, University of Singapore, Singapore, October 2004.

Rabiner, L. R. and Schafer, R. W. (1978). *Digital Processing of Speech Signals*. Prentice Hall, New Jersey.

Scheirer, E. D. and Slaney, M. (1997). Construction and evaluation of a robust multifeature speech/music discriminator. *ICASSP, Munich, Germany*.

Shenoy, A., Wu, Y., and Wang, Y. (2005). Singing voice detection for karaoke application. *Visual Communications and Image Processing 2005, Proc. of SPIE*, page Vol. 5960.

Sundberg, J. (1987). *The science of the singing voice*. De Kalb, Il., Northern Illinois University Press.

Tsai, W. H. and Wang, H. M. (2006). Automatic singer recognition of popular music recordings via estimation and modeling of solo vocal signal. *IEEE Transactions on Speech and Audio Processing, January 2006.*, pages Vol. 14, No 1.

Tzanetakis, G. (2004). Song-specific bootstrapping of singing voice structure. Technical report, Department of computer science, University of Victoria.

Witten, I. H. and Frank, E. (2005). *Data Mining: Practical machine learning tools and techniques*. 2nd Edition, Morgan Kaufmann, San Francisco.

Zhang, T. (2002). System and method for automatic singer identification. HP Labs Technical Report.

Analyzing Harmonic Progressions with HarmIn: the Music of Antônio Carlos Jobim

Giordano Cabral[1], Robert Willey[2]

[1]Laboratoire d'Informatique de Paris 6
Université Pierre et Marie Curie

[2]School of Music
University of Louisiana at Lafayette

Giordano.Cabral@lip6.fr, willey@louisiana.edu

Abstract. *This paper describes a tool designed for the analysis, retrieval, and visualization of harmonic progressions which aim to provide the user with valuable statistics and graphs. This output is intended to help the users to better understand the harmonic content of a music dataset, hopefully causing the emergence of interesting findings. The analysis of the music of Antônio Carlos Jobim motivated the creation of our tool, and for this reason is presented as our case study. We also describe how this tool can be integrated into an accompaniment system, going beyond analysis into the fields of interactive music and composition.*

1. Preamble

Last year, Robert Willey contacted me. He was looking for IT tools to assist in analyzing the music of Antônio Carlos ("Tom") Jobim (Willey 2005). As I was working on a bossa-nova accompaniment system, which needed exactly the kind of information he was able to provide, I became immediately interested in his project. He is an American music researcher seeking hints that might point out the way Jobim composed, especially how his chords were arranged on the piano, and ultimately how to play the piano in Jobim's style.

We then started to build a set of tools using some techniques inherited from my previous work (Cabral et al 2006 and Cabral 2005). We began by testing the tools with the large corpus containing the transcription of a set of authoritative Jobim's songbooks: the Cancioneiro Jobim (Jobim 2005). As we were not actually sure about what those tools were intended to find, we developed one functionality at a time, reevaluating the utility and abilities of the tools iteratively. In the end, the tools were found to be flexible enough to work with different corpuses.

The purpose of this paper is to present such tools in their current stage. We hope it will be useful for other researchers, from whom we look forward to getting feedback and suggestions about new features to be implemented.

2. Introduction

One of the first applications of computer programs to the making of music was the automatic composition by Lejaren Hiller of the *Illiac Suite* in 1957 (Hiller and Isaacson

1993). Since that time computer programs have been used to analyze the compositional style of a number of composers. What makes our tool special is the way it is adjusted to respond to specific queries. We not only developed a tool that extracts statistical information from the data, but also invented a new way of plotting harmony that facilitates the understanding of the harmonic path traced by a song, a set of songs, a style or a composer's work (in our case study, Jobim's work). Beyond that, we integrated our tool to an accompaniment system, which allowed us to musically interact with the dataset, as well as to compose new Jobim-like harmonies.

This paper describes this tool, and illustrates its applications in the study of Jobim's music. It is divided into two main sections: the first presents a description of the tool's architecture, its main functionalities, and data formats; the second lays out the experiments and the results which were obtained.

3. HarmIn 1.0

The ensemble of tools developed is named HarmIn. It is made of a number of features, as mentioned above. The system takes a harmonic sequence as input and provides several outputs, intended to help the discovery of the rules, patterns, and harmonic shapes enclosed in the sequence. The following subsections explain each part of the HarmIn.

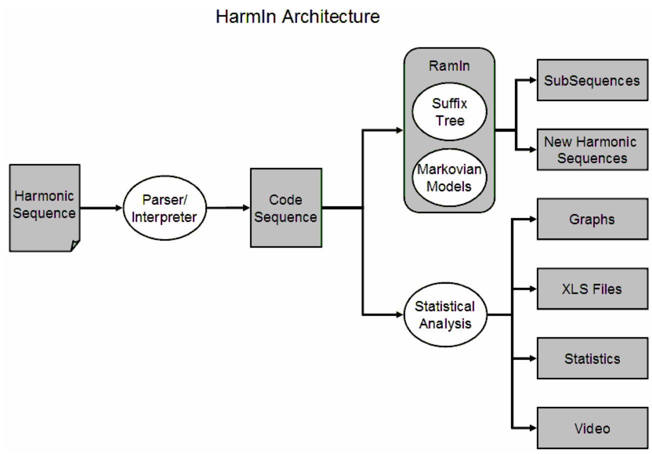


Figure 1. HarmIn architecture.

3.1 The Input: Harmonic Sequences and the Parser

The input of the system is a harmonic sequence. This harmonic sequence is a text file, containing a sequence of chord names. Symbols in the database that are not relevant for this study, such as repetition of a chord "/", or a rest "x" are ignored. The parser is responsible for reading and interpreting the input file. It was originally conceived to read files in the format defined by (Willey and Cabral 2007), but in fact it is flexible to allow other notations. The expected data format is formed by one or more songs, each one composed of a header and a body. The header contains general information (e.g. title, tempo, key, composer), and the body contains the list of chords. The header can be omitted, so a simple file with only a sequence of chords is also accepted by the system. Even tablatures may be correctly read, if the user chooses to ignore the chords not recognized by the system (so it will ignore the lyrics part). Additionally, the user can redefine the chord names understood by the system. Normally he or she will only add to the list of known chords, but he/she can also specify not to accept certain names. This can be done in two different ways: either the user edits a text file (chords.ini), or he/she answers a question when the system finds a chord it does not recognize.

Figure 2 shows an example of a harmonic sequence input. The part going from “title” to “chord_symbol_misprint_m88” is the header. The body is enclosed by the “Page” and “Done” tags.

title	time	42	Cmaj7	G7(b9)	C
Imagina	3/4	bars	/	/	/
tempo	composer	100	/	/	/
moderato	TomJobim&ChicoBuarque	comment	G7(b9)	C	E7(b9)
comment	arranged	chord_symbol_misprint_m88	/	/	Done
FIX_KEY	PauloJobim	Page	/	/	
key	volume	G7(b9)	Cmaj7	G7(b13b9)	
C	1	/	/	/	
major	page	/	/	/	

Figure 2. Example of an input file containing a harmonic sequence.

3.2 Interpreting Rules

To perform the statistical analysis, the system must have a set of interpreting rules, in order to convert the chord names into actual chord classes. This is defined by the user in the same “chords.ini” file described above. In fact, when defining an accepted chord name, the user must also indicate the chord class it refers to. Figure 3 shows an example of such a “chords.ini” file. In the case the system does not find the interpreting rule for a chord in the input harmonic sequence, the user is requested to classify it, continuously increasing the number of chords defined in the chords.ini file.

```
case {'', '7M', '7M(9)', '6', '6/9', '6(9)', '7M(#11)',
'7M(#5/9)', '7M(6/9)', 'maj7', 'maj7(9)', 'maj7(6)', '96',
'maj7(#9)', '96(#11)', '(#5)', '(omit3)', '9', '9(#11)',
'(add9)', '6(#11)', 'maj7(#119)', 'maj7(#11)', '(#119)',
'(13)', '(#9)', '(#11)', '(maj7)', '(b13)', 'maj7(96)',
'maj7(13#119)', 'maj7(#5)', 'maj7(#9#5)', '(b9)', '(13#9)',
'maj7(#5', 'maj7(b5)', 'maj7(#116)', '6(#9)', '(#9#5)',
'maj7(69)', 'add9(#5)', '(b5)', '69', '69(#11)', '9(no3)',
'(add9omit3)', 'maj7(139)', '6(11)', '(b6)'}

type = 0;


case {'7', '7(b9)', '7(9)', '7(#9)', '7(13)', '7(b13)',
'7(b9', '7(b5)', '7(#11)', '7(#5)', '7(9', '7(b9/b13)',
'7(9/#11)', '7(b9/#11)', '7(b5/b9)', '7(#119)', '7(13b9)',
'7(b13b9)', '7(#11b9)', '7(b9b5)', '7(#11#5)', '7(9#5)',
'7(b9#5)', '7(#9#5omit3)', '7(b13#9)', '7(13#9)',
'7(13#119)', '7(13#9)', '7(13#11)', '7(#11#9)', '7(139)',
'7(#9#5)', '7(b139)', '7(9#)', '7(11)', '7(#9b9)',
'7(13#11b9)', '7(omit3)', '(b13b9)', '7(13#11#9)',
'7(#9b5)', '7(11#13b9)', '(b13#9)', '7(13omit3)'}

type = 1;


case {'m', 'm6', 'm(b6)', 'min', 'm6(9)', 'm7M', 'm7M(9)',
'm(7M)', 'm(9)', 'm(maj7)', 'm96', 'm(11maj7)', 'm(add9)',
'm(9maj7)', 'm(11)', 'm9(maj7)', 'm(#5)', 'm9', 'm7',
'm7(b5)', 'm7(9)', 'm7(11)', 'm7(b5', 'm7(b5/9)', 'min7',
'm74', 'm7(119)', 'min7(9)', 'm7(6)', 'm7(9b5)', 'm6(11)',
'm7(b9)', 'm(96)', 'm7(b13)', 'm7(96)', 'm7(69)',
'm6(119)', 'm7(b6)', 'm(7)', 'm(6maj7)', 'm7(13)',
'm7(139)', 'm7(11b5)', 'm7(#5)', 'm7(9#5)', 'm(b6add9)',
'm(9b6)', 'm7(omit5)', 'm7(13119)', 'm(119maj7)',
'm(maj76)'}

type = 2;


case{'°', '°(b13)', 'dim7', 'dim', 'dim7(b13)',
'dim(maj7)', 'dim(maj7)_Fm', 'dim7(9)'}

type = 3;


case {'7/4', '4/7(9)', '4/7', '4/7(13)', '7/4(b9)',
'4/7(b9)', 'aug', '74', '74(9)', '74(b9)', '74(b13)', '4',
'(134)', '4(add9)', '74(139)', '7(sus)', '4(96)', '4(b9)',
'74(1393)', '4(b9#5)'}

type = 4;


case{'7_C', '_C', '_Cb', '_C#', '_D', '_Db', '_D#', '_E',
'_Eb', '_E#', '_F', '_Fb', '_F#', '_G', '_Gb', '_G#', '_A',
'_Ab', '_A#', '_B', '_Bb', '_B#', '_Eb(add9)',
'_Ebm(add9)', '_Eb6', '_C6', 'm_Ab6', '_F6'}

type = 5;
```

Figure3. Examples of interpreting rules defined in the chords.ini file.

3.3 RamIn and the Code Sequence

The parser/interpreter generates a file which is readable by the RamIn system. RamIn is an accompaniment system designed to find and play chords to accompany a melody input via microphone. RamIn generates the chord lines by simulating a Markovian model which encapsulates the probability distribution of chord transitions learned from an input harmonic sequence. Since our tool transforms the initial input harmonic sequence into the RamIn input sequence, we can use all of its capabilities, which include finding the most frequent subsequences, and creating new harmonies coherent with the input harmony. These functionalities are possible thanks to the use of suffix trees, for details see (Cabral 2006). Section 4 presents some illustrative examples.

3.4 Statistical Analysis

The main objective of this study was to analyze the harmonic content of Jobim's songs. So, the most important aspect of the tools was to support the analysis by extracting data and statistics from the input, and constructing graphs. The output of this module of the system is simple statistical values, such as the number of chord changes, the standard deviation of the chords, or their distribution. Tabbed text files (XLS files), readable by sheet programs (such as Microsoft Excel) as well as by scientific ones (such as Matlab) are also exported, so that the user can perform his own queries and generate graphs.

3.5 Graphs and Videos

The system can also generate a set of graphs and videos. The graphs include a histogram of the single chords in the input harmony or the subsequences found by RamIn (Figure 4), a plot of the chord transitions as a two-dimensional matrix, where the two dimensions represent the two chords in the transitions (Figure 5), and a specially designed graph, where the chords follow the circle of fifths and are connected by lines representing the transitions (Figure 6). This graph gives a visual signature to the input harmony. This signature may refer to a single song or a group of songs from a composer or a style. It provides a new way of comparing them.

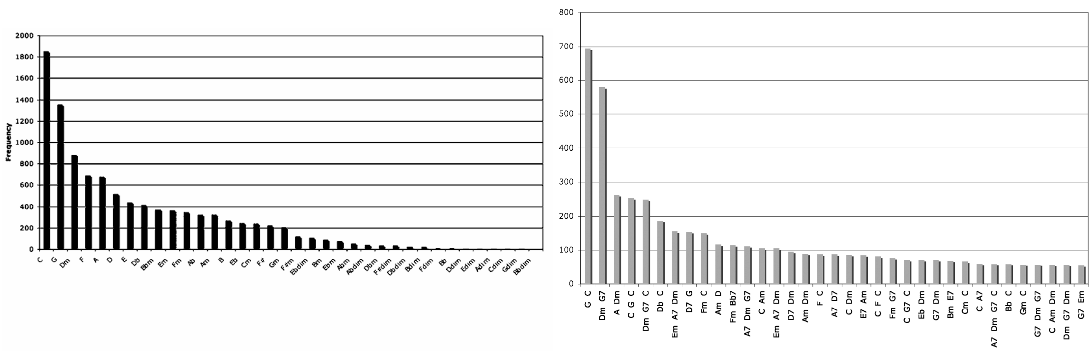
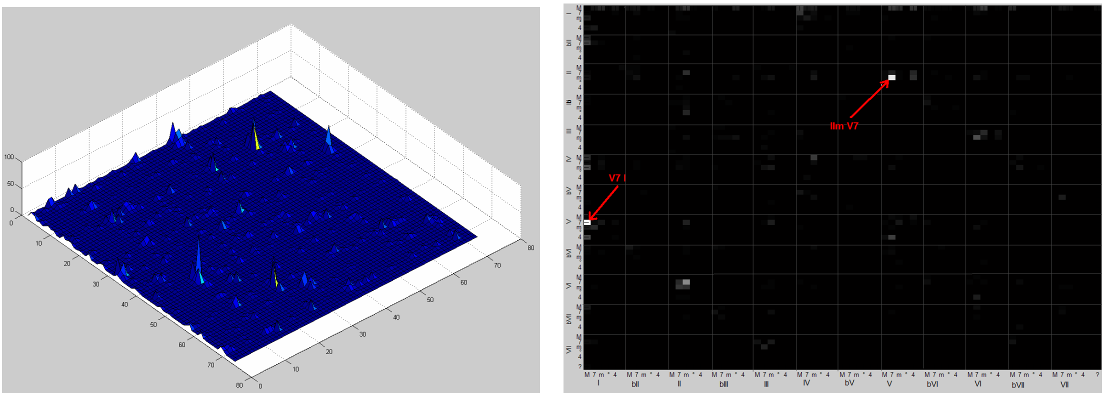


Figure 4. Histogram of the chords (at the left) and of the subsequences of chords (at the right).

Figure 5 shows the frequency in which the chord transitions occur. A chord transition is a change from a first chord to a second chord. The graphs represent the first chord as the row and the second chord as the column. Higher points (in Figure 5a) and brighter pixels (in Figure 5b) indicate a higher incidence of that transition. For example, the pixels representing the transitions from *ii* to *V7* and from *V7* to *I* are clearly brighter, and are annotated in figure 5b.



Figures 5a and 5b. Distribution of chord transitions, plotted as a 3D graph (on the left) or a 2D graph (on the right).

Figure 6 shows the same information but in a more legible way, as one can also follow the sequence of transitions. The little circles represent the chords, and the lines connecting them represent the transitions. The chord roots are displayed following the circle of fifths clockwise, so that the final graph generally becomes simpler. The quality of the chord varies according to its distance from the center, and has a different color (in the case illustrated in Figure 6, major=blue, dominant=green, minor=yellow, diminished=salmon, suspended=purple). The thickness and transparency of the lines indicate the frequency in which each chord transition occurred. Progressions that occurred less than 1% of the time were omitted to make the graph more readable. The direction of the transition is indicated by the coordinate of the line vertex inside the circle. Each line starts in the upper part of a circle and points to the lower part. For example, in Figure 6, the thick line between *Vdom* and *Imaj* indicates a frequent transition from chord *Vdom* to *Imaj*.

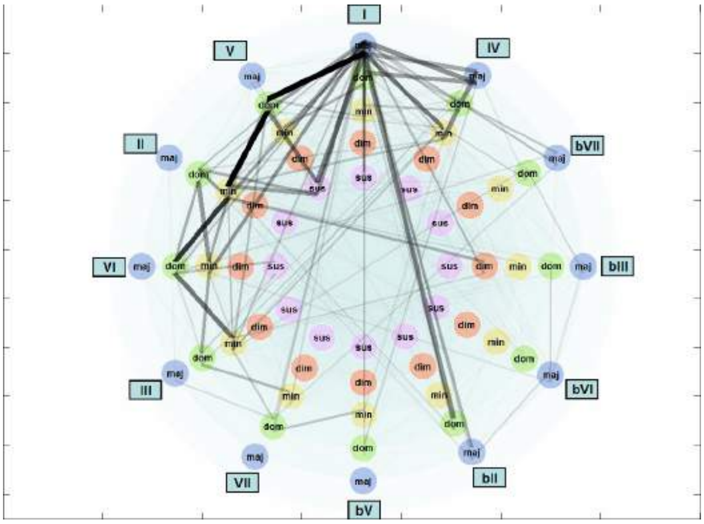


Figure 6. HarmIn chord transition graph.

Besides being used for music analysis, this graph can be used in composing music, since a chord sequence can be produced by following a path of chord connections. The more one chooses thicker and more opaque lines, the more the resulting sequence will be coherent with the input harmony.

Video animations can also be created showing the evolution of these graphs over the songs, the time, the key, or any other characteristic the user judges to be pertinent. Some examples of these videos are available at http://willshare.com/jobim.

4. Case Study: The Music of Antônio Carlos Jobim

For this study, the music of Antônio Carlos Jobim was used, initially to quantify the distribution of chords and the variety of progressions. Experiments were performed in order to determine which approaches to analysis would show the best results. Separating the analysis into groups of songs in major and songs in minor reduced the number and frequency of significant chord transition pairs, clarifying the text and graphic analysis output.

We also wished to investigate the effect of considering internal modulations in the songs or not. The data was analyzed and represented graphically, leading to observations concerning how his harmonic language changed over the course of his career, and we eventually developed graphing techniques that made the data easy to read.

4.1. Cancioneiro Jobim

The data for this study was entered into Microsoft Word files corresponding to the *Cancioneiro Jobim* songbook series, an authoritative five-volume set of approximately 250 songs covering Jobim's complete recorded works, written by Antônio Carlos Jobim, Paulo Jobim, and a team of arrangers under their supervision (Jobim 2001). Then, a VBA program converted it into the harmonic sequence serving as the input of the system.

The details of the data representation, as well as the process of converting the piano arrangements and accompanying chord symbols, and a preliminary comparison of Jobim's work with that of the Beatles is described in a separate paper submitted for the conference (Willey and Cabral 2007). The database can be obtained from the authors for research purposes.

4.2 Tonality and Modulation

All chords are considered relative to key of the song in which they occur. For example, a D major chord in a song in A major is considered to be the same as a B major chord in a song in F# major. This follows the method used in KG Johansson's study of the recorded works of the Beatles (Johansson 1999) of transposing all the songs to the key of C major or minor before beginning the analysis. This way, the data from songs in different keys can be combined.

Modulations in the music are indicated in the database embedded amid the chord data. The analysis can either include or ignore them. Figure 7 shows an example of this. If modulations are considered, the *Eb*, *Ab* and *Db7* would be respectively a *I*, a *IV*, and a *bVII7*. Otherwise, this sequence would be seen as a *bIII*, a *bVI*, and a *bII7*. However, if modulations are considered, the transitions from *C* to *Bb7*, and from *Bb7* to *Eb* would be respectively [*I→bVII7*] and [*V7→I*]. This means that the same chord (*Bb7*) would be considered differently in the two consecutive transitions (as a *bVII7* in the first, as a *V7* in the second). The current implementation ignores the modulation, in order to maintain the accurate representation of the input chord sequences.

```
                Key: C         Eb                  C
              Chord: C F   G C Bb7    Eb    Ab   Db7    C F   G
    With Modulation: I IV  V I bVII7  I     IV   bVII7  I IV  V
Ignoring Modulation: I IV  V I bVII7  bIII  bVI  bII7   I IV  V
```

Figure 7. How tonality and modulation may be considered in HarmIn.

4.3 Chords

One of the innovations of the bossa nova was its inclusion of extended and altered chord tones, increasing the variety of possible chord types. For example, there are twenty three varieties of minor chords present due to permutations of added *6*, *9*, *11*, and/or *13* and their alterations. In the analysis for the study these extensions were ignored, and the chords were grouped into five types: major, minor, 7 (dominant), diminished, and suspended. The reduction of music to a series of chord symbols acts as a filter to the complexity of compositions, creating a generalization that allows the statistical analysis to find significant trends. If every possible pitch combination were allowed as input there would be so many possibilities that the analysis would not be useful. Slash chords, such as *F#7/D* which did not resolve to simple inversions, unusual polychords, such as an E^b triad over a $C\#m^7$ chord (written as one chord symbol separated from another by a horizontal line), or any other unrecognized type were left out of the analysis.

The most common in songs in major keys, in order of decreasing frequency, were found to be *I, IIm, V7, VI7, IVm, IV, II7, IIIm, VIm, Vsus, III7, bVI, bII, bVII7, bIII, bIII7, VIsus, III, and VI*. An analysis of the transitions from one chord to another explained the presence of non-diatonic chords such as *bII (subV7), bVII7* (*V7* of *bIII*), and *bVI* (*subV7* of *V*). Yet, the order holds some surprises, specially the *IVm* befor the *IV*, and the low positions of the *III* and *VI*.

4.4 Chord progressions

The focus of the statistical analysis was to analyze the transitions between pairs of chords, and to see how the chord progressions changed over the course of the five periods in Jobim's compositional life. The arrangement of chords in songs in major keys display a preponderance of root movement by fourths and fifths, as seen in Figure 8a by the number of short lines connecting chords on adjacent spokes (i.e. between chords in the *I* or *IV* groups). The most dramatic exception to this appears to be the distant move from *bII* to *I* nearly opposite on the circle. This is misleading for the eye, however, since *bII* is actually just a substitute for *V,* the neighbor of *I*.

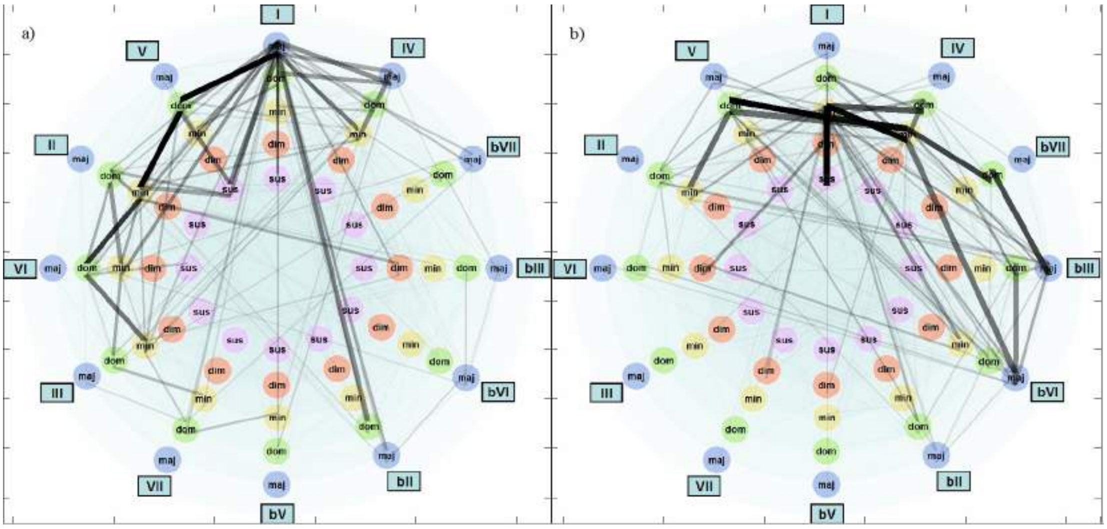


Figure 8. Chord transitions in all songs in major keys (a) and in minor keys (b).

Like those in major keys, the chord progressions used in songs in minor keys also have movement by fourths and fifths and their substitutions. However, there is more root in these songs by step and third, which shown in Figure 8b as connections with longer lines which cross neighboring spokes.

4.5 Subsequences

A list of the most common subsequences found in Jobim's songs in major keys is shown in Figure 9. As expected, transitions like *<VI7 ii V7 I>* were found among the most frequent. This group of frequently used elements is typical of popular music and jazz standards of from the second half of the 20th century, indicating that chord progressions alone are not sufficient to explain the unique characteristics of Jobim's style. The chord symbols in the database generally do not include the non-chord tones that Jobim often used in his melodies. Another important element of his style that is missing is the description of his smooth voice leading and countermelodies.

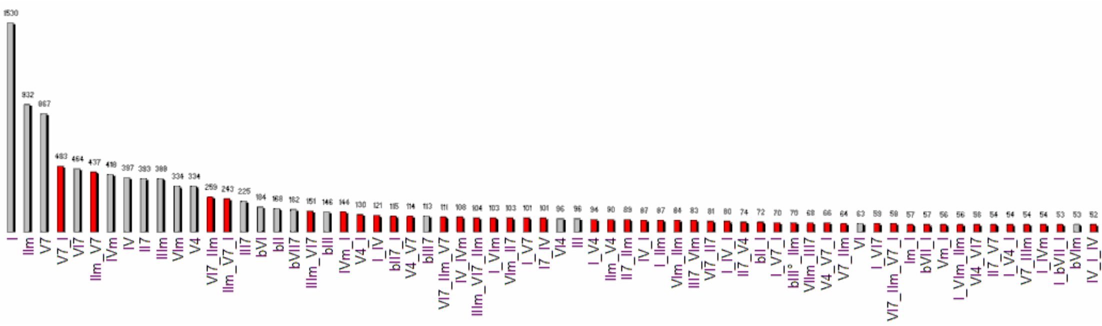


Figure 9. Subsequences in songs in major keys occurring more than fifty times The bars in gray indicate single chords, which can be omitted from the graph.

Still, some surprises emerged, like the progression *<II7 IIm>* among the most usual. Another important finding is that some single chords appear after entire sequences. For example, the *III* (35th position, 96 occurrences) has less occurrences than the sequence *<iii V7/ii ii>* (29th position, 104 occurrences), and the *VI* (52nd position, 63 occurrences) has less occurrences than sequences like *<bIII° ii>* (48th position, 70 occurrences) and *<vii III7>* (49th position, 68 occurrences).

4.6 Change in progressions over the years

Figure 10 shows the evolution of the harmony in the music of Jobim over the volumes of the Cancioneiro series. The last graph refers to the sum of all five volumes. The most apparent change in harmonic progressions over time was observed to occur between volumes 1 and 2, shown in Figures 10a and 10b. This marks a major change in Jobim's vocabulary, beginning in volume 2 which covered the bossa nova period, recognized for its rich harmony. The HarmIn tools also generated video animations showing the evolution volume by volume and song by song. These videos, as well as other material is available at the website http://willshare.com/jobim.

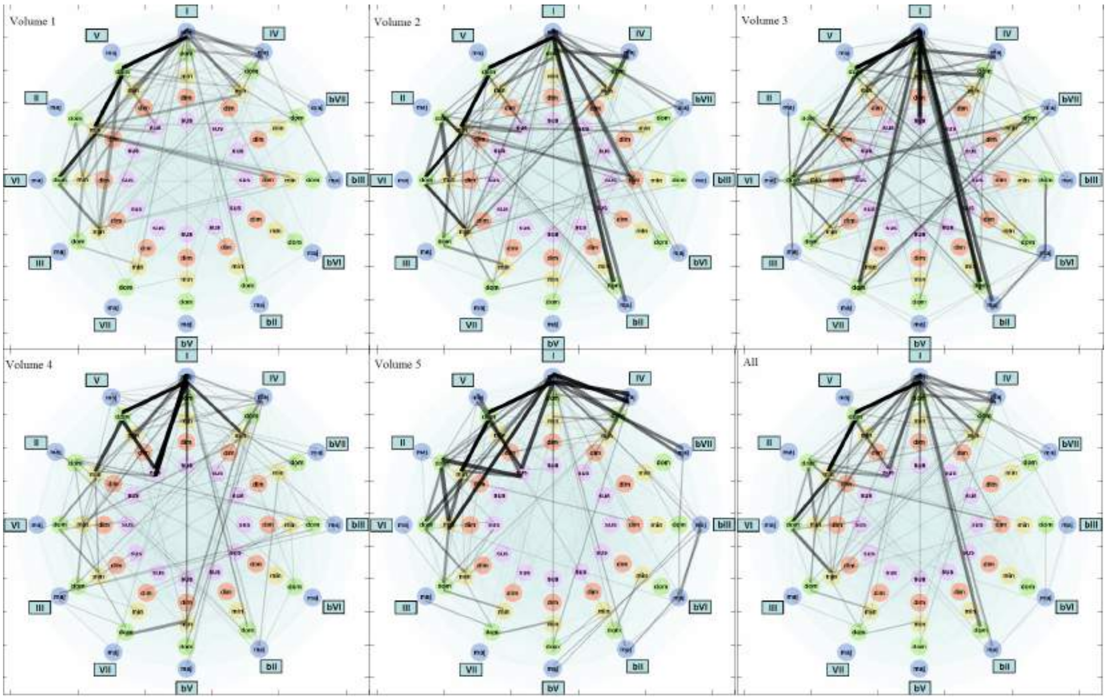


Figure 10. Chord transition graphs related to songs in major keys from the 5 Volumes of the Cancioneiro Jobim.

4.7 Analysis of most famous songs

Figure 11 shows HarmIn ability to compare datasets. It can display the graphs side by side, compute statistical data, and generate graphs showing the differences between the datasets. Figure 11 presents a selection of sixteen of Jobim's most famous songs in major keys (the Hits dataset, in Figure 11b) compared with the entire group of songs in major keys (Figure 11a). The Hits dataset appears to have more modulations to distantly related keys, and shows increased use of substitute dominants, resulting in smooth bass lines. Figure 11c shows the graph of the difference between them. The red lines represent a higher incidence of certain chord transitions in the first dataset, green lines represent a higher incidence in the second one.

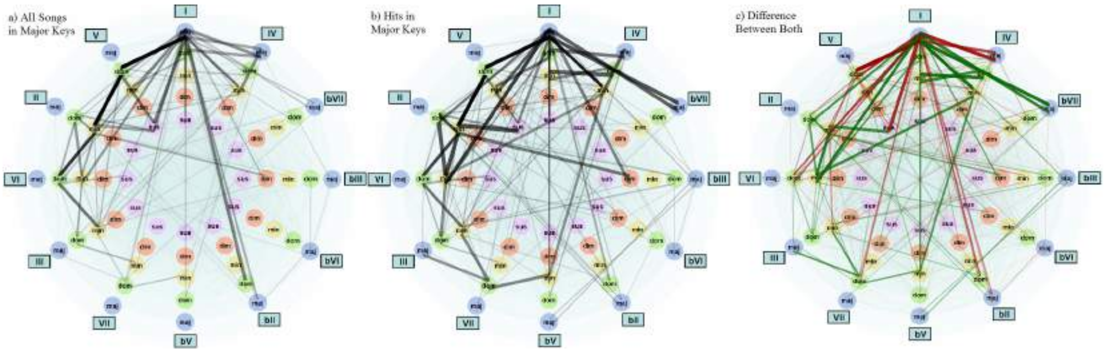


Figure 11. Comparison between the chord transitions in 16 of Jobim's most famous songs and the rest of his work. The 3rd graph plots the difference between them.

4.8 Comparing Datasets

Section 4.8 already exposed how HarmIn can be used to compare datasets. We performed another experiment in this sense, intended to compare the music of Antônio Carlos Jobim with other composers or styles. We gathered many tablatures from a website specially dedicated (Mvhp 2007). Then, we created six datasets, three for composers (The Beatles, Chico Buarque, Caetano Veloso), and three for music genres (respecting the classification given by the site: Axé, Forró, Brazilian Funk). The Beatles' dataset is composed of 107 songs, Chico Buarque's of 121, and Caetano Veloso's of 122. The Forró dataset contains 121 songs by 15 artists, the Axé dataset contains 396 songs by 16 artists, and the Funk 25 songs by 8 artists. Although these datasets are not authoritative and parsing them may lead to errors, introducing the results still seems worthwhile.

Figure 12 shows comparative graphs between his music and that of The Beatles. Figure 12a shows the difference in the number of times each chord has been used. This graph reveal that The Beatles used much more certain few types of chords (negative values in the histogram), while Jobim used a wider range of chords (positive values in the histogram). Figure 12b shows the difference between the number of times each chord transition has occurred. White pixels indicate a higher incidence of a particular chord transition in the music of The Beatles, black pixels indicate a higher incidence in the music of Jobim. Figure 12c shows the same difference in the chord transition graph. Green lines represent chord transitions used more often by The Beatles, red lines are the ones used more often by Jobim.

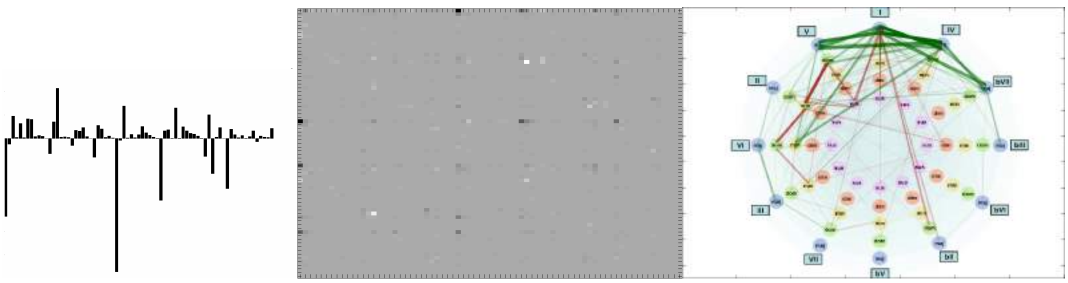


Figure 12. Comparison between the music of Jobim and the Beatles.

HarmIn can also give some statistics about the comparison. For example, the correlation or the covariance of two datasets, the standard deviation of a group of datasets, the sum of the absolute differences between two datasets, or the sum of their product. The table below presents the correlation coefficients between them. We can see how the Axé and the Brazilian Funk harmonies are highly correlated. Also, the data point Chico Buarque's harmonies being the most correlated to Jobim's, and interestingly a certain similarity between Caetano Veloso's harmony and that of The Beatles.

Major	Tom	Beatles	ChicoBuarque	Caetano	Forro	Axé	Funk
Tom	1,00000	0,47868	0,68656	0,61271	0,28819	0,57131	0,61775
Beatles	0,47868	1,00000	0,69512	0,85065	0,69558	0,83029	0,63504
Chico Buarque	0,68656	0,69512	1,00000	0,78441	0,51290	0,72072	0,62971
Caetano	0,61271	0,85065	0,78441	1,00000	0,69159	0,83953	0,68011
Forro	0,28819	0,69558	0,51290	0,69159	1,00000	0,84511	0,68511
Axé	0,57131	0,83029	0,72072	0,83953	0,84511	1,00000	0,87386
Funk	0,61775	0,63504	0,62971	0,68011	0,68511	0,87386	1,00000

Minor	Tom	Beatles	ChicoBuarque	Caetano	Forro	Axé	FunkCarioca
Tom	1,00000	0,44540	0,68752	0,59144	0,33371	0,34375	0,21177
Beatles	0,44540	1,00000	0,27671	0,65560	0,45967	0,39702	0,38435
Chico Buarque	0,68752	0,27671	1,00000	0,51849	0,38903	0,36033	0,23657
Caetano	0,59144	0,65560	0,51849	1,00000	0,48903	0,53617	0,37496
Forro	0,33371	0,45967	0,38903	0,48903	1,00000	0,62917	0,55129
Axé	0,34375	0,39702	0,36033	0,53617	0,62917	1,00000	0,52432
Funk Carioca	0,21177	0,38435	0,23657	0,37496	0,55129	0,52432	1,00000

Figure 13. Table with the correlation coefficients between each pair of datasets considered.

The correlation between chord transition matrices, however, does not always explain harmonic similarities or divergences. Even if all matrices are normalized by the overall number of chords in each database, the scaling may affect the results. Sometimes a big difference in a certain chord transition hides many similarities. The main harmonic characteristic of a composer may be the use of certain chords in specific situations, and as these chords may seldon appear, they won't impact the final result. Thus, the analysis must comprise all graphs and statistics: the number of chords used, the variety and shape of the chord transition paths, etc. HarmIn does not give an answer to questions like "Which composer is most similar to Jobim?", but it provides the user with the tools to analyze and reach its own conclusions. To demonstrate that, the relation of the three genres we studied is shown in Figure 14. In fact the graphs seem quite similar while very different from the others. Check our website for the full graphics and data.

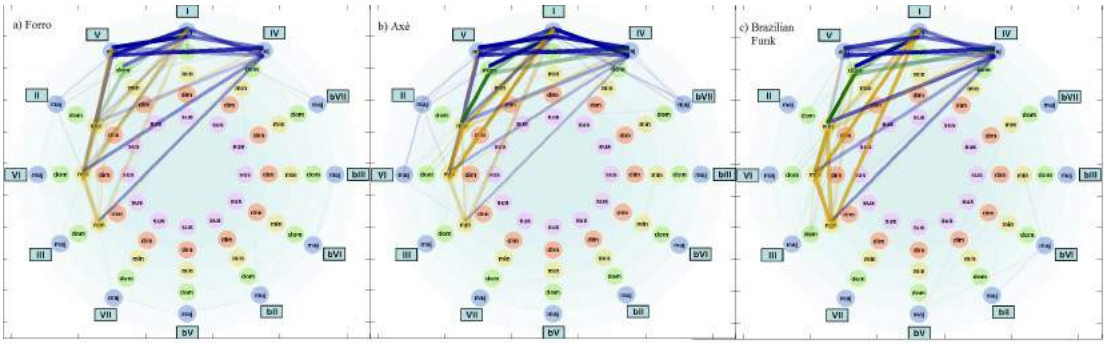


Figure 14. Chord transition graphs for Forro, Axé, Brazilian Funk genres.

4.9 Generation of new progressions

A final feature of HarmIn we experimented was its capability to integrate with RamIn, then to use its power to create new chord progressions, coherent with the harmony used as input. RamIn allows an iterative and interactive composition while the user repeatedly accepts or rejects continuations for a given harmonic input sequence. Figure 15 below shows a progression made with the help of RamIn.

```
Dm7(9) / G7(13) / Dm7(9) /
G7(13) / Dm7(9) / G7(13) /
Dm7(9) / G7(13) /

C7M(9) / G(add9) / G/A Em7
F#m7 Em6/G G#m7(11) / C#7(9)
/ G#m7 / C#7 /

C#7(b13) / Bm7 / Bm / Bm7 /
E7(13) / Em7 / A7 A6(9) Dm7
/
/ / Gm7(9) / C7 / Gm7 /
A7/E / Dm7(9) / / /

Dm7(9) / / /
G7/B / Gm7/Bb / Am7 / / /
Am7 D7 Am7 D7 Am7 D7 Am7 D7

Dm7 G7(13) Dm7 G7(13) C / D
/ D / C#7 / D / Bm7 /
Em7 / A7 A6(9) D6(9) / / /
```

Figure 15. A new chord progression based on the analysis of Jobim's, made with the help of RamIn.

5. Conclusion

This paper presented a set of tools grouped under the name of HarmIn, intended to help the analysis of harmony, especially chord progressions. The main features of HarmIn were presented, along with a case study with the music of Antônio Carlos Jobim, as well as some comparisons between him and other composers. A website accompanying this paper (http://willshare.com/jobim) includes sample song data, color graphics and videos, and audio examples of generated sequences.

6. References

Cabral, G. and Briot, J.-P. and Pachet, F. *Incremental parsing for real time accompaniment systems*. In Special Track on Artificial Intelligence in Music and Art (AIMA'2006) Florida Artificial Intelligence Research Society Conference 19th International FLAIRS'2006 Conference, Florida Artificial Intelligence Research Society, Melbourne Beach, Florida, USA, 2006.

Cabral, G. *Comping by Humming : Adaptation x Continuité dans l'Interaction Musicale*. Rapport de Pré-Soutenance de Thèse, Paris, 2005.

Cabral, G., Briot, J.-P., and Pachet, F. *Impact of Distance in Pitch Class Profile Computation*. In Proceedings of the 10th Brazilian Symposium on Computer Music (SBCM 2005). Belo Horizonte, Brazil, 2005.

Fujita, T. *The Beatles: Complete Scores*. Milwaukee: Hal Leonard Publishing Corporation, 1993.

Hiller, L. and Isaacson, L. "Musical Composition with a High-Speed Digital Computer", in Machines Models of Music, Schwanauer, M. And Levitt, D. Eds, MIT Press, pp. 9-21, 1993. Reprint of original article in Journal of Audio Engineering Society, 1958.

Jobim, A. C. *Cancioneiro Jobim: complete works, arrangements for piano*, 5 volumes, Jobim, P., coordination. Rio de Janeiro: Jobim Music, 2001.

Johansson, KG. *The Harmonic Language of the Beatles*, STM-Online Vol. 2, 1999. http://www.musik.uu.se/ssm/stmonline/vol 2 1/KGJO/index.html

Mvhp. *MV Portal de Cifras*, Website: http://www.mvhp.com.br, 2007.

Willey, R. *"Antônio Carlos ('Tom') Jobim"*, Encyclopedia of Recorded Sound in the United States, edited by Frank Hoffmann. New York: Routledge, 2005.

Willey, R., and Cabral, G. *Representing Antônio Carlos Jobim's Harmonic Progressions*. Submitted to the 11th Brazilian Symposium on Computer Music (SBCM 2007), São Paulo, 2007.