Modeling Embellishment, Duration and Energy Expressive Transformations in Jazz Guitar

Sergio Iván Giraldo Méndez

MASTER THESIS UPF / 2012
Master in Sound and Music Computing

Master thesis supervisor:
Rafael Ramirez
Department of Information and Communication Technologies
Universitat Pompeu Fabra, Barcelona

To Camilo

Acknowledgements

I would like to express my gratitude to my advisor Dr. Rafael Ramirez, whose expertise, guidance and patience made possible this master thesis work. I'm truly thankful for how he shared his vast knowledge and experience, and for his trust, when involving me in some other projects developed during the year. Also, I deeply appreciate his encouragement and support in my application for perusing an PHD career.

It wouldn't have been possible to write this master thesis with out the support and help from Sankalp Gulati and Jose Zapata, among other people at the MTG, as well as the SMC colleagues who gave their support with their friendship, help, comments, team work, etc.

Finally I would like to thanks Vanessa and Joan for their help and support, to my family for their spiritual support, and specially to Camilo for his patience.

Abstract

Professional musicians manipulate sound properties such as timing, energy, pitch and timbre in order to add expression to their performances. However, there is little quantitative information about how and in which context this manipulation occurs. This is particularly true in Jazz music where learning to play expressively is mostly acquired intuitively. We propose to develop a machine learning approach to investigate expressive music performance in Jazz guitar music. We extract symbolic features from audio performances and apply machine learning techniques to induce expressive computational models for embellishment, timing, and energy transformations. Finally, we apply concatenative synthesis techniques in order to generate expressive performances of new scores using the learnt computational models.

Contents

List of Figures

List of Tables

Chapter 1

INTRODUCTION

Unarguably, music performance plays an important role in our culture. People clearly distinguish the manipulation of sound properties by different performers and create preferences based on these differences. However, there is little quantitative information about how and in which contexts expressive performance occurs. This is particularly true in Jazz music where most of the performance information is acquired intuitively.

In this thesis work we propose a machine learning approach to investigate and recreate expression in jazz guitar music. Our methodology is divided in to three stages:

1. Symbolic feature extraction from both audio performances and scores

2. Machine learning modeling to induce computational models for embellishments, timing (duration and onsets), and energy transformations

3. Synthesis of the predicted scores, by applying concatenative synthesis.

Quantitative evaluation is based on accuracy ratings of the models, such as correctly classified instances, and correlation coefficients. Also qualitative evaluation is obtained from synthesized audio evaluation by users.

A comprised overview of this investigation and its preliminary results can be found in Giraldo and Ramirez [Giraldo and Ramirez, 2012],which was presented at the 5th International Workshop on Machine Learning and Music, held in Conjunction with the International Conference on Machine Learning (ICML, June 2012).

In the following sections we will present the research problem towards jazz music expression. We will expose our motivation and present the thesis statement. In chapter two, we will review a state of the art in music expression performance research, focusing in machine learning applications. On Section 3 we will explain

the implementation and methodology. Results and evaluation will be presented in section 4, and finally, a discussion with conclusions and future work will be presented in section 5.

1.1 Research problem

People can often recognize a jazz player just by listening to the head of a song. Professional jazz performers have a personal and unique way to interpret musical pieces. According to Goebl [Goebl et al., 2008] expression is what makes music appealing to listeners, and because of it, music can even induce emotions [Juslin, 2001]. This phenomena can not be easily defined and explained by musicians, and makes it a difficult concept to teach and to be learned.

Certainly, in classical music, a musician performs a score as it is written, following expressive indications (eg. staccato,rubato, piano, forte, etc). On the contrary, in Jazz music, real/fake book's scores [Rea, 2004] lack of performance indications, and performers freely modify the original melody in order to embellish it. These embellishments usually consists of groups of notes that the performer adds to the original melody (up to four or more notes), based on melodic, harmonic and rhythmic density context. Although some authors have categorize these embellishments [Crook, 2002], there are no clear rules about where and how to apply these embellishments, and this has to be learned intuitively by coping and analyzing the performance of professional musicians.

This learning process is depicted in Figure 1.1. On the right side we have the professional musician performance, where sound properties like timing, energy, pitch an timber are manipulated to add expression. On the left side we have the inexpressive score with no performance indications. At the bottom, the learning process takes place when the student analyze (read) the score, and copy (listen) the performance.

1.2 Motivation

Most of the past research in music expressive performance modeling using machine learning has focused on classical piano music [Widmer, 2001]. Exceptions include the work by Lopez de Mantaras et al. [de Mantaras and Arcos, 2002] and [Ramirez and Hazan, 2006]. Lopez de Mantaras et al. describe a system able to infer Jazz saxophone expressive performances from non-expressive monophonic descriptions using Case Based Reasoning. Ramirez et al. applies inductive logic programming to obtain models capable of generating and explaining expressive Jazz saxophone performances. Although these systems are able to handle simple

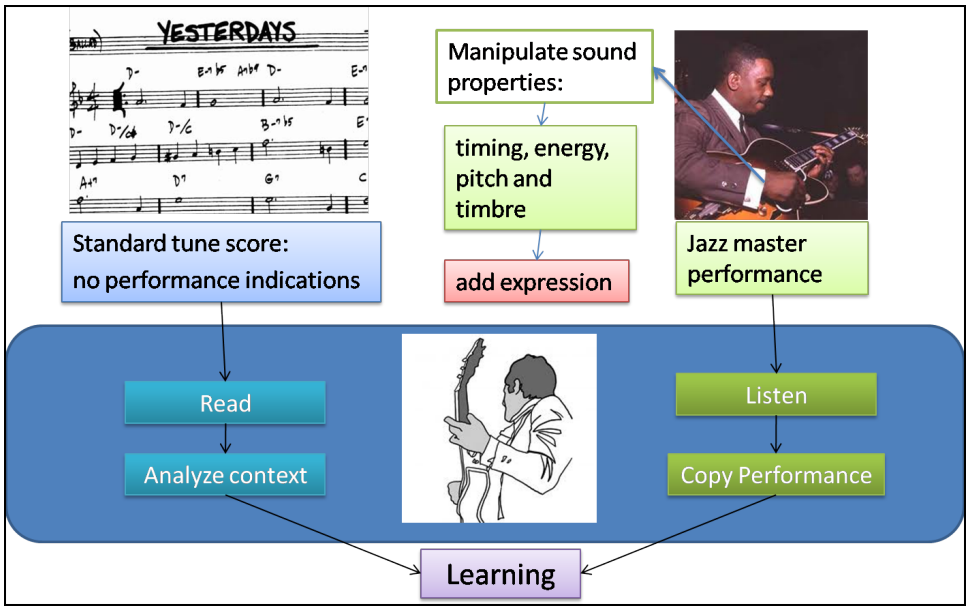


Figure 1.1: The learning process in popular music: On the left we have the standar tune score with no performance indications, on the right we have te professiona musician performance who adds expression. Learning process takes place by listening (imitation) and reading (analisis).

ornamentations (such as grace notes), they can't predict more complex melodic embellishments (improvised short note melodies between notes) where added notes may double the number of notes of the original score.

On the other hand, expressive synthesis has been studied in the past by Maestre et al [Maestre et al., 2009], who propose a concatenative synthesis approach to sax jazz music by generating a expressive performance of a score by reusing audio recordings and concatenating notes samples. Laurson et al. [Laurson et al., 1999] propose a expressive systems for acoustic classical guitar synthesis which can handle expressive input parameters (eg. vibrato, dynamics, and timing transformations). As far to our knowledge, any of the previous approaches have not been applied to jazz guitar music.

This expressive aspect of jazz music towards melodic embellishment is an unexplored area which is placed between written music and fully improvised music. The mentioned techniques can be used to obtain systems, which can be later applied to develop tools, such as educational or composition software, for understanding and recreating expression given a inexpressive score and a recorded performance.

1.3 Thesis statement

In this section we present our thesis statement, based on the following hypothesis:

> Its possible to capture expression by extracting information about the note itself as well as its musical context of both the performed and the score notes. Using this information models can be trained, by means of machine learning techniques, to accurately predict embellishments, as well as transformations in timing, and energy. Then, by applying concatenative synthesis, it will be possible to generate a final synthesized expressive performance of a certain score as it would have been performed by certain performer.

1.4 Goals

The specific goals of this thesis work are:

- To develop a system able to induce an expressive performance from an inexpressive score, taking as reference the performance of a real professional musician.
- To extract features both from the note itself, as well as from the musical context, and analyze which of these descriptors are the most relevant for each modeling stage.
- To compare different machine learning algorithms in order to check which of them give better results for each modeling experiment.
- To obtain a score with predicted embellishments, duration, and energy for each note, in order to synthesize it and compare it to a human performance.

Chapter 2

STATE OF THE ART

In this section we will review the state of the art in music expression, giving an overview of the past and present research in the field. We focus on specific aspects where expression modeling using machine learning has been applied to jazz performance. A second part is devoted to explain the relevant steps in music expression modeling such as data acquisition, melodic representation and description and Machine learning techniques.

2.1 What is music expression?

There have been different attempts to define expression in music performance. The variation in timing, dynamics, timbre and pitch that makes differentiable one performance from another is defined as music expression by Juslin [Juslin, 2001]. Goebl [Goebl et al., 2008] defines it as an integral part of music: ”without expression, music won’t be interesting for most part of listeners, (...) what listeners go to listen, when going to a concert, is the human expression that gives sense to mu-

sic". Ramirez [Ramirez and Hazan, 2006] defines it as the manipulation of pitch, timing, amplitude and timbre that performers do, and that are clearly distinguishable for listeners.

As depicted in Figure 2.1, and for the purpose of this study, we will focus on expressive variations in:

- Duration: a note may be enlarged or shortened from its score value
- Onset: A note may be delayed or anticipated with respect to the music score.
- Energy: A note may be played louder or softer
- Embellishment: A note may be replaced by a set of notes.

In previous studies the embellishments studied were grace notes, or a long note that turns into two notes or vice-versa. In this study we will focus on more elaborate embellishments, where the sets of notes used to embellish one single note range from two to five notes.

2.2 Music Expression Modeling

Music Expression has been empirically studied since the beginning of XIX century and beyond. Gabrieslsson [Gabrielsson, 1999] presents a review of nearly 600 paper in music performance research, making his review one of the most cited in music performance. Goebl et al. [Goebl et al., 2008] presents a general overview of Music expression performance focusing on data acquisition, computational studies, and Models. They explain from a more technological point of view the most relevant steps in when researching on music performance. From

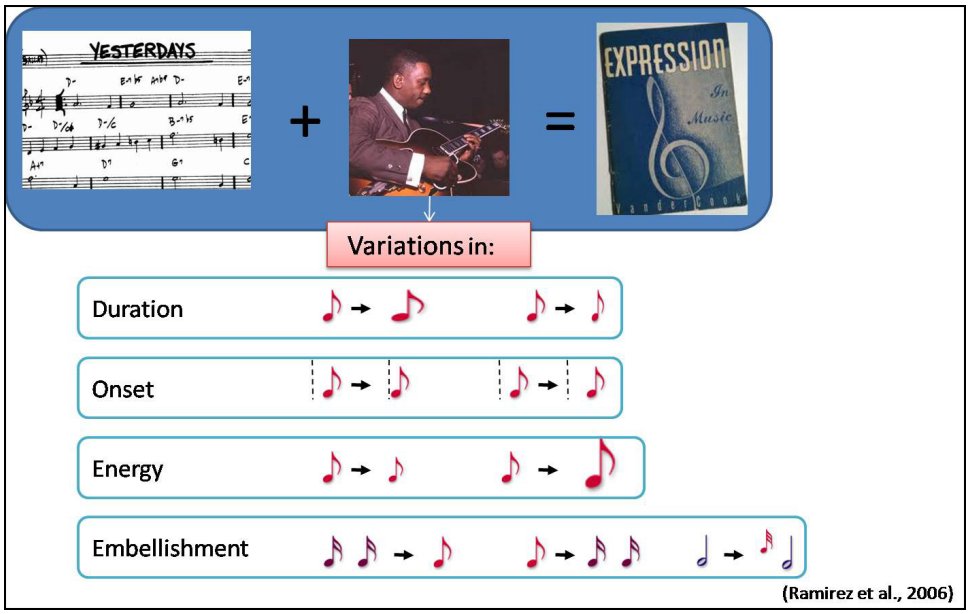


Figure 2.1: Expressive perfomance definition. Expression is defined as the manipulation a performer does in duration (enlarge shorten notes), onset (delay or anticipate notes), energy (play notes louder or softer) and embellishment(add ornament notes).

Goebl's overview its clear that most of music expression research has been done in the context of classical music and in particular in piano music, mainly because the piano keys can be easily treated as a on/off switch devices. A few studies have been conducted in jazz saxophone music, e.g. [Ramirez and Hazan, 2006], [de Mantaras and Arcos, 2002]. Guitar expression has been studied by Laurson et al. [Laurson et al., 2010], who propose a model of the *rasgueado* technique in the classical guitar. However the purpose of this model is mainly to improve guitar synthesis.

2.2.1 Expressive Jazz Music Performance

Most of the research in jazz expressive performance modeling has mainly considered the saxophone. The resulting expressive models aim to generate a synthesized performance based on real human performances taken from recordings, and also to explain the expressive transformations applied to a score in terms of tempo, note energy.

Tempo transformations of monophonic recordings of jazz saxophones are modeled by Lopez de Mantaras et al.[de Mantaras et al., 2007] to automatically render a performance played at a particular tempo in to a different target tempo preserving its expressiveness. Their model, TempoExpress, uses case-based reasoning, which uses the solution of similar previously resolved problems to solve the current problem, by adapting the corresponding solution. The model uses as inputs a MIDI score of a jazz standard, the melodic description of a monophonic recording of a saxophone performance, and the target tempo. The output is the transformed melodic description of the performance in the new tempo, which is used as a ba-

sis for the audio synthesis. The melodic content description and synthesis relies on the system developed by Gomez [Gómez et al., 2003]. Similarities between the case base (human performances at different tempos) and the input performance are assessed by tempo filtering, and performing a melodic similarity measure based on abstract representations. Melodic segmentation [Temperley, 2001] is used to generate matching segments as partial solutions of the input problem, and then apply constructive adaptation [Plaza and Arcos, 2002] to construct the final solution.

Another system for jazz modeling is the one proposed by Lopez de Mantaras et al. [Lopez de Mantaras and Arcos, 2002], the SaxEx. They report a system able to infer an expressive performance form a flat, non monophonic input, by using Case Based Reasoning. The musical context role of each note of the input is analyzed, and the system retrieves from a case memory of human performances, notes with similar roles, and using its properties transform the input notes. A first step of the note analysis is performed by spectral modeling techniques [Serra et al., 1997], by comparing qualitative values (a.e. dynamics in dB) over a single note, and compare this value to the average value over the whole preformed piece. The second step of the note analysis uses Narmour's model [Narmour, 1990] and Generative Theory of Tonal Music [Lerdahl and Jackendoff, 1983] to determine the role of the note within the musical phrase: place on the melodic progression, metrical strength, duration, harmonic stability and relative importance in the bar. However this model can not explain the transformations realized to the inexpressive performance.

Ramirez et al. [Ramirez and Hazan, 2006] compares different machine learning techniques to obtain a jazz saxophone model capable of generating synthe-

sized expressive performances and to explain the expressive transformations. They implement a system which uses inductive logic programming (ILP) which creates a logic set of rules that model expression from both intra-note and inter-note level. Intra-note level refers to a set of features that are relevant to the note itself, like pitch, attack, onset (inflections), while inter-note level refer to the musical context at which the note appears, like interval with the previous and the next note, duration and loudness. Thus, considering a set of inflections (intra-note level) and the musical context (inter note level), the system could predict the type of inflection to be used in a particular context. Using this information, they model various performers' styles of playing, so when a new performance is presented, the system is able to identify the performer by analyzing the performance style. The intra-note set of features are represented by: attack level, sustain duration, legato (left-right), energy mean, spectral centroid and spectral tilt. For inter-note features set are Pitch, Duration, Previous note pitch, Next note Pitch, Next duration, and 3 set of Narmour structures. Depending on the type of instrument, the aspects of the performance to be taken in account for the performer identification task may vary. For example, in piano, dynamics and timing are relevant aspects for performer identification, but not timber, conversely to saxophone or singing voice, where timber is the most relevant attribute for this task. Two classifiers are used: one to map melodic fragments to the different possible performers, and a note classifier that maps each note of the performer to be identified into an Alphabet symbol, which is the set of clusters generated by all the notes performed by all performers. These classifiers are obtained by the use of different machine learning techniques: K-means clustering, Decision trees, Support Vector Machines, Artificial Neural Networks, Lazy Methods and Ensemble Methods.

The training recordings are segmented and the intra-note descriptors are computed for each note. Fuzzy k-means clustering is applied using the intra-note information to group similar notes. Then for each performer, the training recordings of that performer are collected, and inter-note descriptors are computed for each segmented note in the performer's recording. A classifier (Decision Tree) is build using inter-note features as attributes and the cluster previously calculated as a class. By clustering inter-note features they obtain sets of similar notes for all performers, and by building decision trees with intra-note features, they predict the notes a performer will play given a musical context.

As far as we know none of the methodologies mentioned above have been applied to the guitar, in particular to electric jazz guitar.

2.2.2 Expression Modeling Process

In the following subsections we are going to review the overall process of expression performance analysis in music, the standard procedures, and the relevant work on each stage of the process.

Data Acquisition

The starting point is the data acquisition of music performance which can be done in two ways: monitoring the performance, using special devices (a.e. MIDI pianos, sensors, video capturing, etc) and extracting relevant data from the recorded signal. Although there are some performance aspects that can't be accessed from the recorded signal, it has the advantage that there is more than one century of recordings available for scientific research. In the first data acquisition way dif-

ferent systems are explained, most of them mainly focused on measuring piano performance by placing measurement devices to the keys. Such is the case of Piano rolls, the Iowa Piano, Shaffers photocells. Digital pianos, and synthesizers were widely used, and after the Yamaha Disklavier: a computer controlled grand piano involving an extended MIDI utility. See [Goebl et al., 2008] for an overview.

On the other hand, extracting information from audio recordings was initially done by hand, a time consuming process. This means getting the audio wave form in a computer software and manually marking onsets, of musical events. Dynamics are measured by reading the peak energy values of the root mean square averaged over a window. Initially, computational extraction of audio was developed for supporting the process of manual annotation [Dixon 2000 - 2004], but these systems required an important amount of interactive correction. An alternative approach is to use previous annotations from other performances of the same music, by aligning the recordings [Dixon and Widmer, 2005]. A step further is to avoid the initial annotation by having the score in symbolic format [Cano et al., 1999; Soulez et al., 2003; Turetsky and Ellis, 2003; Shalev-Shwartz et al., 2004]. Other systems have been developed to extract expression from performances in MIDI format [Cambouropolos, 2000]. Also the body movement of the performer has been used to extract expression by connecting devices to the performer's body or to the playing apparatus of the performer, as well as using less intrusive systems as video capture. [Goebl et al., 2008]

Melodic representation, description and extraction

Melodic representation is one of the first steps in modeling expression. The aim is to get a representation of the melody of a song in terms of numerical data, in order to be used in information retrieval systems, departing from the fact, that melody plays a major role in the process of recognizing a song. This representation should have compactness to be easily stored, keep the expressiveness of the melody and have portability to be adaptable to different types of inputs. Some of these descriptions are score oriented. This means that the melody is described by pitch information (mainly) and pitch contour. Gomez et al. [Gómez et al., 2003] review different techniques for melodic description and extraction, in particular the MPEG7 description scheme, and some of the intermediate steps to overcome this process such as fundamental frequency extraction, and melodic patterns induction.

Ramirez et al. [Ramirez et al., 2010] makes use of two different set of descriptors: one set for characterizing the internal description of a note (attack level, sustain duration, sustain slope, amount of legato with the previous note, mount of legato with the following note, mean energy, spectral centroid and spectral tilt) and another one to characterize the musical context of the note (relative pitch and duration of the neighboring notes as well as the musical structures to which the note belongs).

The extraction of inter-note features follows the following general scheme: first the note boundaries are extracted from both multiband energy extraction based on psycho-acoustic knowledge [Klapuri, 1999] and fundamental frequency transitions. Then note descriptors are computed from the boundaries and low level

descriptors values, by averaging the frame values within the note segment. Also pitch histograms are used to compute pitch, in order to avoid wrong fundamental pitch frame values to be taken in to account in the mean computation. Note transitions are computed by the implementation of two descriptors: Energy minimum position over transition time, and legato. These characterize the detachment between two consecutive notes and are calculated by measuring ratios between areas and the distances between the decay line of the first note and the attack line of the second note. Finally Narmour theory of melodic perception is used to parse the performance and identify the Narmour structures over it.

The intra-note feature extraction is performed by first performing intra-note segmentation, in order to find the attack, sustain and release segments of each note. For doing so the energy envelope is treated as a derivative function over time where the second derivate maximums and minimums represent the points of inflection of the function. Low pass filtering is performed before derivation, and the cut off frequency is fitted by minimizing an error function defined in terms of the smoothed envelope and the frequency. A set of characteristic set of points is found and from the maximums, and slopes they form, each segment of the note is defined. The absolute and relative duration of each segment is calculated, and for the stable part of the note (sustain) the averaged spectral centroid and spectral tilt descriptors are computed in order to characterize the brightness of the note executed.

Computational modeling of music performance

The goal of modeling music performance is to understand and find the relationship between the various aspects of performance and the expression. Models serve

both as a descriptive and predictive tool to infer output data given the input data. This simulation process is used to predict the behavior of a certain phenomenon in different circumstances. To develop the parameter of a model two strategies are used: Analysis by measurement, which is based on the analysis of the deviations of the recorded performance from the score. The goal is to recognize regularities in the deviation patterns by the use of statistical data techniques such as regression analysis, linear vector space theory, neural networks, or fuzzy logic. Analysis by synthesis takes into account the human perception and subjective factors. Departing from the analysis of a recorded performance, a set of rules is derived by music experts, taking in to account musical knowledge, in order to synthesize an expressive performance. This second type of approach is used by the KTH group [Bresin, 2002; Friberg et al., 2000, 2006].

Machine learning in music performance modeling

Arthur Samuel (1959) defines Machine Learning as the ”field of study that gives computers the ability to learn with out being explicitly programmed”. A more technical definition is done by Tom Mitchell (1998) by a well posed learning problem as follows: ”A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E. For our study the task T will be to predict duration, energy and embellishments for a certain note based on intra-note and inter-note features, the experience E would be the performance of a professional musician meaning how he manipulates embellishments, energy and duration of notes, and finally the the performance P would be the measure of how well the model predict each note’s duration, energy and embellishments on a

given example.

In the preceding techniques the different sets of rules for expression model were taken from a preliminary hypothesis of the researcher, or the music expert. The other approach is to extract potentially interesting regularities in a large data set, without a preliminary hypothesis by means of machine learning algorithms. The advantage is that this approach avoids musical expectations or assumptions, and even more, some algorithms can induce models from intelligible rules with a musicological meaning. Widmer [Widmer, 2002] uses a machine learning system to describe regularities in a large data set of Mozart piano sonatas. Some of the rules obtained describe regularities already incorporated in the KTH model, but a few contradicted some common hypothesis.

Widmer and Toudic [Widmer and Tobudic, 2002] describe a system able to predict tempo and dynamic shapes at different levels of hierarchical music phrases. They implement a model of expressive phrasing and articulation at different levels by combining a K nearest neighbour algorithm to predict expressive shapes from identified shapes in similar pieces, and a rule learning algorithm that predict rules for note levels not attributed to the phrase structure. Basically the system recognises similar phrases on the training set and applies their expressive patterns to a new phrase. The accuracy of the model was demonstrated by a Mozart piano sonata generated by the model that won the second prize in the International Performance Rendering Contest (RENCON) in Tokyo, 2002.

Dovey [Dovey, 1995] uses inductive logic programing to obtain interpretative rules of the piano performances of Rachmaninoff recorded during 1920's on a Ampico Recording Piano. This device was able to record information about the duration and tempo of the notes, and also about the dynamics of the key pres-

sure and pedalling on a piano roll, which can be easily converted into a machine readable format. Vab Belen et al. [Baelen and Raedt, 1997] departs from Dovey's work and extend it to predict MIDI files from the musical analysis of a score using inductive logic programming.

Chapter 3

IMPLEMENTATION AND METHODOLOGY

A general overview of the basic frame work is shown in Figure 3.1. We obtain a symbolic description from high quality monophonic recordings. From the score we will encode a machine representation, and the structure of the pieces. Comparing this two sets of information, we will extract the expressive aspects of the recordings, and use machine learning to generate models to predict this expressive aspects. Later, a test score not considered in the training phase is parsed and used to generate an expressive performance, using the model learned by the machine.

3.1 Software

The implementation was done using Matlab [MATLAB, 2011]. Some Python [Jones et al., 01] scripts were used to implement Essentia [1] libraries for onset

[1] Essentia & Gaia: audio analysis and music matching C++ libraries developed by the MTG (Resp.: N. Wack),`http://mtg.upf.edu/technologies/essentia`.

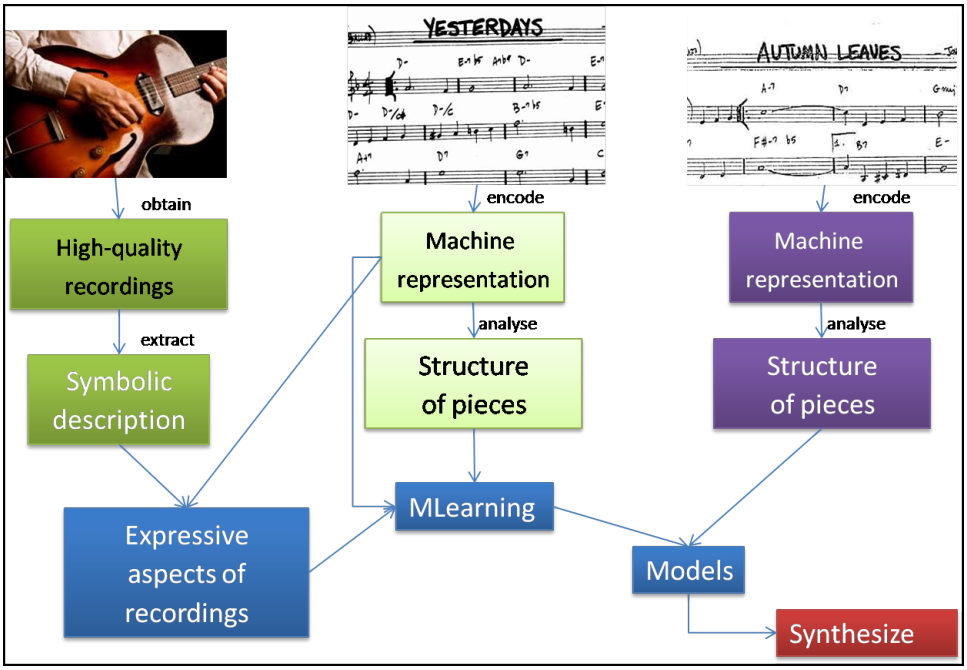


Figure 3.1: Basic research framework. On the left side we extract a symbolid description from audio recordings. On the center we obtain a machine representation of the score. From the previous two we extract the expressive aspects of recordings and generate models. On the right a test score is parsed through the model, to make performance predictions.

detection, and Weka [Hall et al., 2009] software was used to run the machine learning experiments. However all the implementation is integrated in Matlab, so Python scripts as well as Weka commands are run from the main Matlab script. A sumary of the script architecture and brief explanations of the algorithms and functions developed can be found in Appendix A.

Other software used include:

- Band in a Box: To handle chord and key information of the tunes.
- Finale: To produce a MIDI transcription of the scores.
- Sound forge: To perform manually note segmentation correction.
- Audacity + Guitar Rig: To add guitar effects to the final music excerpt.

3.2 Data acquisition

The training data used in this study are guitar monophonic recordings of standard Jazz pieces performed by a professional musician. Initially the pieces were obtained from a jazz guitar book's accompanying CD [Marshall, 2000] in which the melodies were recorded on a separate channel, apart from the accompaniment. The CD performances closely imitate original recordings of several jazz tunes by famous jazz guitar players. As the melody track on the CD, was recorded with reverb effect, this induced errors in the onset detection process. The performance was then re recorded by a professional musician, following the transcription score, and copying as close as possible the recording performance. An electric guitar was used, passing the signal directly to the recorder. This showed to improve the onset detection results.

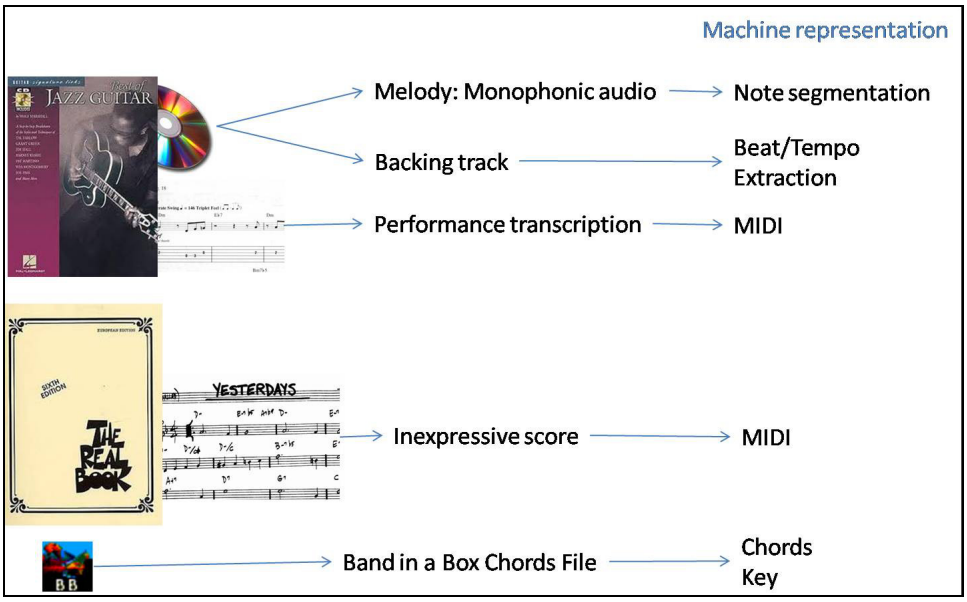


Figure 3.2: Data Acquisition. Data is aquire from audio recording of the melody and the backing track, the performance transcription, the inexpressive score, and the chord chart file.

Data Acquisition process is depicted in Figure 3.2, and can be resumed in the following steps:

- Note segmentation: From the melody track we perform note segmentation to obtain a data base of note samples. Is performed using frequency and energy frame descriptors. The process relies on the implementation of the Essentia[2] audio processing library. Manual correction is performed after the onset detection process.

- Beat and tempo extraction: The algorithm proposed by Dan Ellis [Ellis, 2007] is used for beat tracking, and quantization of beat onsets is performed by linear regression. This beat and tempo information is used to set the tempo of

[2]Essentia & Gaia: audio analysis and music matching C++ libraries developed by the MTG (Resp.: N. Wack),`http://mtg.upf.edu/technologies/essentia`.

the MIDI files: Score, performance and test files.

- Machine representation of the scores: A complete transcription of each performance is included in the volume. This performance was written into MIDI format using score writing software. The same process was applied to obtained the inexpressive scores in MIDI format. These scores were taken from known jazz standard books compilations. MIDI format was parsed using MIDI toolbox for Matlab [Eerola and Toiviainen, 2004].

- Chord and key information was obtained form Band in a Box type files available on the internet. This type of files was converted to text files using a chord chart converter by Choi [Choi, 2007], to make possible to read chord and key information in Matlab.

3.3 Note Description

Each note in the training data was annotated with a number of attributes representing both properties of the note itself (Intra-note features) and some aspects of the context in which the note appears (inter-note features). Information about the note included note duration, energy, pitch, and metrical position within a bar, while information about its melodic context included information on neighboring notes (i.e. relative pitch and duration), as well as, melodic analysis with respect to the key and harmonic analysis with respect to the ongoing chord.

Intra-note features

The intra-note features used for this study were:

- Pitch: Is the pitch of the note, given by the MIDI number from 1 to 127

- Pitch mod: Is the pitch in one octave, called also as chroma. Is measured by a number representing the 12 semitones starting in zero (0 to 11).

- Note to key: Is the interval between the current note and the key.

- Bar: Is the bar at which the note occurs. Bars are numbered starting from zero.

- Onset: Is measured in seconds as well as in beats, or beat fractions. Beats are numbered starting from zero.

- Beat onset mod: Is the beat at which the note occurs within a bar. Is measured in beat or beat fractions.

- Duration:Is measured in seconds as well as in beats, or beat fractions.

- Metrical Strength: This is the strength of the note in terms of the place in the bar: 1st beat: Very strong (ss) 2nd beat: Weak (w) 3rd beat: Strong (s) 4th beat: Weak (w)

  Up beats are calculated as beat fractions and their strength is set to very weak (ww) as follows:

  16th notes up beats 11/2, 21/2, 31/2 and 41/2 beats are label as very weak (ww)

  8th note triplets up beats 11/3, 12/3, 21/3, 22/3, 31/3, 32/3, 41/3, 42/3 are label as very weak (ww)

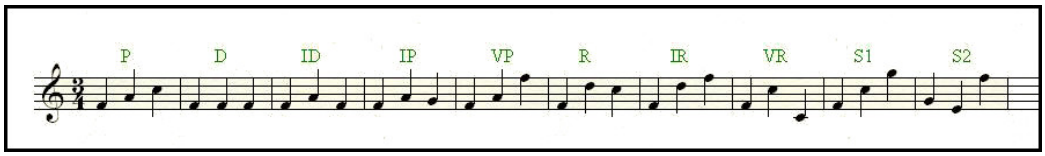


Figure 3.3: Narmour structures. Narmour implication realization model as used by Ramirez et al. 2006. We incorporate two more structures (S1 and S2) to complete all the possible melodic movements .

- Velocity: Is the loudness of the note, is given by a MIDI parameter between 1 to 127

Inter-note features

The inter-note features used for this study were:

- Previous duration: Refers to the duration of the previous note.
- Next duration: Refers to the duration of the next note.
- Previous interval: Refers to the interval measured from the previous note to the current note.
- Next interval: Refers to the interval measured from the current note to the next note.
- Narmour: Refers to the Narmour's implication - realization model of perception. A particular note can be at the initial, middle and third position of a 3 note Narmour structure. For this study we implemented in Matlab a system to analyze each note and its neighbors notes, to classify it into the corresponding three Narmour structures. This classification was performed according to the classification used by [Ramirez and Hazan, 2006]

Table 3.1: Narmour classification: The table shows the 8 possible combinations for three note intervalic motion.

NARMOUR STRUCTURE	INTERVALS SIZE CC	DIRECTION
P	SMALL SMALL "SS"	SAME
R	LONG SMALL "LS"	OPPOSITE
D	UNISON	-
ID	EQUAL SMALL INTERVALS	OPPOSITE
IP	SMALL SMALL "SS"	OPPOSITE
VP	SMALL LONNG "SL"	SAME
IR	LONG SMALL "LS"	SAME
VR	LONG LONG "LL"	OPOSITE
SA (S1)	LONG LONG "LL"	SAME
SB (S2)	SMALL LONG "SL"	OPOSITE

and shown in Figure 3.3. We added two more cases: two long intervals in the same direction, and, a short interval followed by a long interval in the opposite direction. This two cases complete all the possible intervalic movements. Long and short interval limit was set to 6 semitones. The process is performed creating a 3 note window and for each window we diferentiate the pitch to get the intervals. Each of the two interval are labeled small or large (e.g. both intervals small: "ss", large interval after a small interval: "sl", etc), and also we checked if both intervals go move in the same direction or in the opposite. This two conditions are evaluated simultaneusly and all the possible combinations are labeled as showed in Table 3.1

- Chord: Is the current chord in which each note occurs.

- Chord root: It refers to the root of the chord, in terms of chroma. Is given by a number between 0-11 (12 semitones)

- Chord type: It refers to the type of chord whether it is a triad, four-triad or

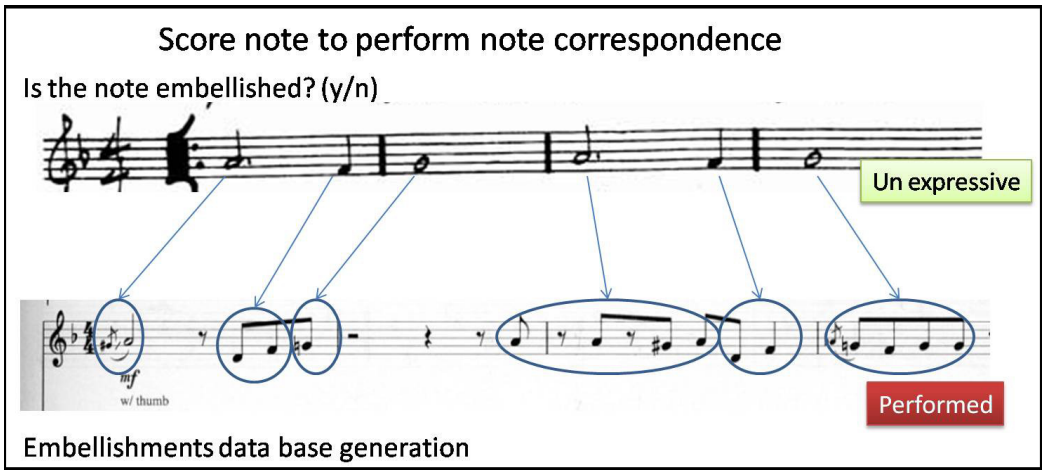


Figure 3.4: Embellishments correspondence. Above: a fragment of the inexpressive score. Bottom: same fragment as performed by a professional musician. Circles and arrows indicate the type of embellishment performed to a certain note

a chord with tensions (more than 4 notes). A data base of chord extensions was build with the interval note components of each chord, measured in semitones starting from zero, making also possible to know if the current note is a chord note. The chord list is presented in table 3.2

- Note to chord: Is the relation of the note and the current chord. Is measured with the interval note and the the root of the current chord. Is calculated by the distance between the module of the midi note over 12 (chroma) and the chord chroma value

- Is a chord note? : Is ”yes” when the note is a note of the current chord definition. ”No” otherwise.

Table 3.2: Chord description. A list of chords definitions. The numbers on the left indicate the index of the notes belonging to the chord, (zero indexed, in 12 semitones).

CHORD TYPE	INTERVALS
MAJOR	0 4 7
M (MINOR)	0 3 7
2 (SUS2)	0 2 7
SUS (4)	0 5 7
MAJ7	0 4 7 11
6TH	0 4 7 9
M7	0 3 7 10
M6	0 3 7 9
MMAJ7	0 3 7 11
M7B5	0 3 6 10
DIM	0 3 6 9
7TH	0 4 7 10
7#5	0 4 8 10
7B5	0 4 6 10
7SUS	0 5 7 10
MAJ9	0 2 4 7 11
6/9	0 2 4 7 9
M9	0 2 3 7 9
9TH	0 2 4 7 10
7B9	0 1 4 7 10
7#9	0 3 4 7 10
13	0 2 4 7 9 10
7B9B13	0 1 4 7 8 10
7ALT	0 1 3 4 6 8 10

Table 3.3: Embellish transformation. Example of an annotated embellishment transformation for the half dotted A of Figure 3.4. Quarter notes are set as 1. Semitones are set as 1.

ORNAMENT NOTE NUMBER	PITCH OFFSET	ONSET OFFSET	DURATION
1	0	-1/2	1/2
2	0	+1/2	1/2
3	-1	+3/2	1/2
4	0	-2	1/2

3.4 Embellishment data base generation

In Jazz, performers usually modify the original melody in order to embellish the melody. Thus, embellishments are characterized by the changes in duration, onset, and pitch with respect to the original note. For each note of the score we manually searched its corresponding performance note(s). In Figure 3.4 we present a fragment of the inexpressive score, and below it the same fragment as performed by a professional musician. The circles and arrows indicate the notes of the performance that correspond to an embellishment of a certain note.

A particular note can then remain the same or be transformed to another note or set of notes. This way we classify each note of the original score as *embellished or not embellished*. When a note is transformed (embellished), we annotate how many notes were used for the embellishment and the relative offset in pitch, onset, and duration with respect to the original note. In Table 3.3 is shown an example for the half dotted A of Figure 3.4. This way we generate a database of performed embellishments.

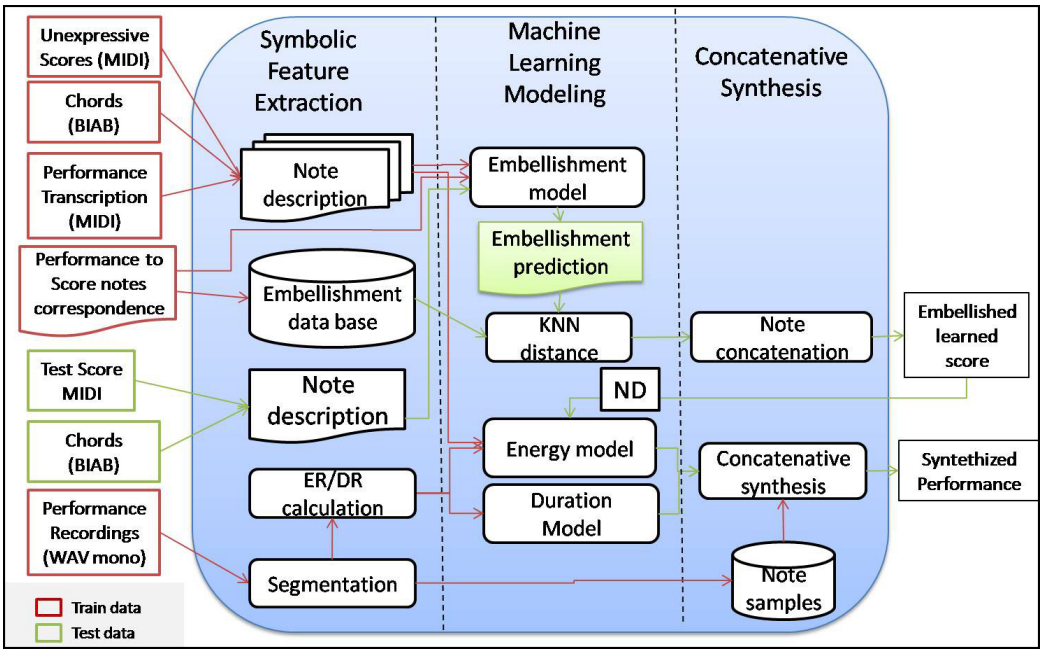


Figure 3.5: Modeling Jazz Guitar Performance. Machine Learning Core Process. training data is in red boxes, test data is in green boxes. The process follows three stages: Symbolic feature extraction, Machine learning modeling and concatenative synthesis. The output of the model is a embellished score and its corresponding synthesized audio excerpt which include transformations in energy and duration.

3.5 Machine learning

We applied several machine learning algorithms (i.e. k-NN, artificial neural networks and support vector machines) to induce duration, energy and embellishment transformation models. Once the embellishment model predicts a note to be embellished, we use k-NN to search for the closest embellishment note transformation in our database. The descriptor set for this search contains note duration, previous note duration, and metrical position within a bar. The core process is depicted in Figure 3.5, and each step will be explained in the following sections.

training data

In Figure 3.5 the training data is shown in red squares. It correspond to the information extracted from the recording of the performance of the professional musician, the score transcription of this performance and the data base of the note correspondence of embellished notes as explained in section 3.4. It consist of:

- Inexpressive Scores (MIDI)
- Chords (BIAB)
- Performance Transcription (MIDI)
- Performance to Score notes correspondence
- Performance Recordings (WAV mono)

Test Data

Also in Figure 3.5, the test data is depicted in green squares. It correspond to the MIDI notes and Chord information extracted from the test MIDI file, in other words the one in which we want to predict and apply embellishments.

3.5.1 Symbolic feature extraction

Note description

In this part we created our data sets for Embellishment, Energy and Duration, using our set of descriptors explained in section 3.3. The process was performed as follows:

- Embellishment Data Set. Having all the set of descriptors for both the inexpressive score, the performed score, we create a embellishment data set. Using the inexpressive score and its features we create a new field where we classify each note as embellished or not, using the embellishment approach explained in section 3.4.

  The same field was created for the test file, having the same set of descriptors. This field was filled with interrogation marks (?).

- Energy Data Set. The same set of descriptors was used. The new field in this case is the energy ratio. This is defined as the energy of the current note, compared to the overall energy of the melody excerpt. What we measure here is if a certain note was played louder or softer than the overall performance. This calculation was performed using the following equation:

$$EnergyRatio(ER) = \frac{OverallRMS}{NoteRMS}$$

- Duration data set. Again, the same set of descriptors was used but now the new field is the Duration ratio, which is defined as the difference between the score duration and the actual duration of the performance. What we measure here is in which amount the performer stretched or enlarged the duration of each note performed compared to the note in the score transcription. The calculation was performed using the following formula:

$$DurationRatio(DR) = \frac{ScoreNoteDuration}{PerformendNoteDuration}$$

Note segmentation

From the note segmentation process described in section 3.2 we use the audio information to calculate the Energy Ratio and the Duration Ratio as described in the previous section. We also create a data base of note samples that will be used later for the Concatenative Synthesis process.

3.5.2 Machine Learning Modeling

A first stage is the creation of the arff file from our data base. We implemented an algorithm to create the arff type files from a Matlab structure type data. The Field names of the structure are used as attribute names.

Initially the experiments were run using the Graphic User Interface of Weka. We perform several experiments using different algorithms, and 10 fold validation. The attribute selection was performed by removing one by one the features were less relevant taking into account musical knowledge, and checking which combination arose the highest Correctly Classified Instances (CCI) for the classification experiment (embellished or not embellished) and Correlation Coefficient (CC) for the regression experiment (Duration ratio and Energy ratio prediction). From the initial set of descriptors explained in section 3.3, the set of descriptors that arose better accuracy were the following:

Energy model

- Pitch
- Onset mod

- Pitch mod
- Previous Interval
- Next interval
- Note 2 Key
- Chord
- Chord Type
- Is Chord note

Duration Model

- Duration
- Prev Duration
- Next Duration
- Onset mod
- Previous Interval
- Next Interval
- Chord type
- Note to chord

The implementation of the exhaustive feature selection built in Weka will be left as future work. What it does is to perform the experiment using all the possible combinations of the features, and retrieves the one that has higher accuracy. Ideally we will perform the this exhaustive selection and compare it to the musical knowledge based selection we perform.

Scripts used

We run all the experiments from Matlab creating the command lines to run Weka from the command line prompt. To do so, a bat file was created form Matlab with the instructions to read the arff files, and also to set the type of machine learning algorithm, and its parameters. A phython script by Sankalp Gulati at the MTG group (UPF), was used to clean the output text file of Weka to make easy to read the predictions in Matlab. In Apendix 1 we present a complete resume of the model and the different algorithms created.

Cross validation considerations

As mentioned before, 10 cross fold validation was used during the machine learning experiments stage. However we found a problem because when performing the 10 fold process in Weka, the data is randomized before the process. Weka's output returns the predicition for each instance in this randomized order, and there is not a way to go back to the initial order. In other words, we have the predictions but we don't know exactly to which instance this prediction belongs. To solve this issue we implemented our own 10 cross fold experiment in Matlab, and performed for each fold the train - test experiment in Weka (with out 10 fold). In this case the predictions heavily depend on the seed we are using for randomization, so we

performed the 10 fold experiment 10 times, using a different seed each time, and finally choose the one with higher percentage of correctly classified instances.

Embellishment Model and Prediction

Two types of experiments were performed: using the same data set for test and train, and using different one as test. In the first case we wanted to know how close the model could predict a performance of a certain tune using the performance of a professional musician as training set. In the second case we wanted to predict a performance of a new tune, to test if our model could create a performance such as we could somehow listen this new performance as if it would have been played by the professional musician who recorded the training excerpt. To select the new tune we considered it to have a similar note density, similar key, and similar key progression to the tune used for training the model.

KNN similarity search for the embellishment model

In both cases after running the embellishment experiment we obtained a prediction of each note of the inexpressive score, to be embellished or not. For each note to be embellished we performed a similarity search on our data base of embellishments, where we previously annotated the transformations and the description of that particular note-transformation. After running several tests, and adding some musical knowledge, we found the most relevant descriptors for doing this search, and also we weighted each descriptor to get more accurate results. This set of descriptors and its weights are shown in Table 3.4 . For instance, duration and previous duration were the ones that were more relevant: if notes are too short the number of notes one can add to embellish it is less than if the note's duration is

Table 3.4: Descriptors and weights used for KNN search similarity for note embellishment. 1 >more weight, 1 <less weight.

DESCRIPTORS CHOSEN FOR KNN SEARCH	WEIGHTS
DURATION	3
PREVIOUS DURATION	2
NEXT DURATION	0.3
BEAT ONSET MOD	0.6
PREVIOUS INTERVAL	0.4
NEXT INTERVAL	0.6

long. The same occurs if the embellishment consist of passing notes: the number of passing notes one can add to embellish the target note, heavily depend on the duration of the previous note. This process was implemented using Matlab knnsearch in built function, using euclidean distance, and a scale factor to weight each feature.

3.5.3 Synthesis

In this section we explain the process of synthesis of the predicted score. Synthesis is a important stage of all the process, as we are working on the concept of Popular Music, in which listening is the most relevant part in the learning process. Notice that many modern professional and non professional musicians do not know how to read written music. This is why we consider that the score is not enough information as an output from our model.

Note concatenation

After finding the best embellisment for each predicted embellished note we perform note transformation using the data set of embellishments. Each note is trans-

formed to a group of notes following the transformations stored in our embellished database as explained in section 3.4. Each transformed note is concatenated into a new MIDI matrix. After this concatenation a correction process for duration is performed to avoid note overlapping.

Duration modeling for the embellished score

Having a predicted score with embellishments, we perform a similar process as described in section 3.5.2. In this case our experiment is a regression type, as we want to predict the duration ratio of each note of the predicted score. First we calculate all the set of descriptors for the embellished score. After, we run our experiment using Weka, and got a prediction of the duration ratio for the predicted score. Having the prediction, we apply this ratio to the corresponding note. After this process we end up with a MIDI matrix with a embellished score with the predicted duration for each note.

Energy modeling

We performed the same process to predict the energy ratio. We run a similar experiment using Weka, and obtained a predicted energy ratio for each note. However, we didnt apply the transformation to the score, because if we were going to synthesize using note samples extracted from the audio wave, this information is somehow contained in the note sample. This part of the process was left for future implementation.

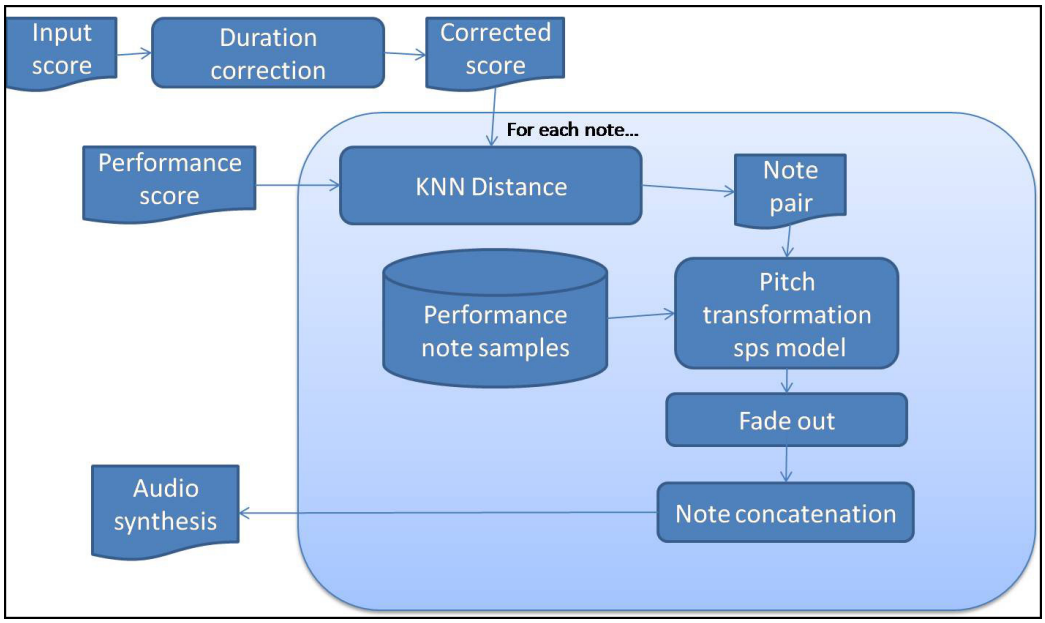


Figure 3.6: Concatenative Synthesis Approach. For each note we search the closest note in a note sample data base using KNN distance. After applying Pitch and correcting the duration, notes are concatenated one after the other.

Concatenative Synthesis

We have applied concatenative synthesis concept [Maestre et al., 2009], to collect notes samples from the audio signal, which are close in context to the predicted score learned by the system. After we have applied note transformations using the sinusoidal model of [Bonada et al., 2011] to adjust the pitch of the sample note to the score note.

The implementation is depicted in Figure 3.6. The input parameters are the predicted score, after performing the note duration correction process, and the performance score. For each note of the predicted score we searched the closest note using the same KNN distance approach of section 3.5.2. This way we found the closest note in context of in the performance score. A new attribute selection

Table 3.5: Descriptors and weights used for KNN search similarity for note concatenation. 1 >more weight, 1 <less weight.

DESCRIPTORS CHOSEN FOR KNN SEARCH	WEIGHTS FOR SYNTHESIS
DURATION	3
PITCH	0.5
PREVIOUS DURATION	0.2
NEXT DURATION	0.3
BEAT ONSET MOD	0.6
PREVIOUS INTERVAL	0.4
NEXT INTERVAL	0.6

and weighting was realized in this stage to get the most musical coherent results. After some tests, the best attribute - weight combination is the one found on table 3.5. The set of descriptors is similar, however, Pitch becomes a mayor attribute to take into account on the similarity search.

Each note sample is transformed using the frequency scaling approach presented on the sinusoidal plus stochastic model of [Bonada et al., 2011]. After several tests the best input parameters found to transform all the set of notes, taking into account the tessitura of the instrument, are presented in Table 3.6. Duration considerations were taken into account after the transformation: the notes transformed to a longer duration were left with the same length and filled out with silence to preserve the natural string decay. On the contrary, a fade out process of the 5% of the final note length is performed to the note samples that were to be shortened to avoid sound clipping. Then, the notes are concatenated one after the other in a audio string. This way we obtained the final audio wave output of the synthesis process. A tempo scaling approach to duration transformation may be consider as future work.

Table 3.6: Input parameters for sinusoidal plus stochastic model

PARAMETER	VALUE
WINDOW SIZE	1001
WINDOW TYPE	BLACKMAN
FFT SIZE	2024
THRESHOLD (NEG DB)	-113
MAXIMUM NUMBER OF SINUSOIDS	20
DECIMATION FACT. OF MAG. SPECTRUM FOR STOCHASTIC ANALYSIS	16

3.5.4 Tempo and onsets considerations

Onsets and beat deviation are relevant parts of the musical expression and were not included in this research work. Although they are left for future work, we implemented a function to take into account the original beat deviation found by the beat tracker. Initially we quantize all the MIDI files and perform all the modeling. After having the final predicted score, we impose the original beat tracking as a tempo grid to the predicted score. This way we preserve the tempo deviations of the human performance. However this temporal fact will be treated deeply in future work.

Chapter 4

RESULTS

We apply the learnt models for synthesizing new expressive interpretations of musical pieces from their (inexpressive) score. In this section we present the obtained synthesis results, as well as a quantitative evaluation based on the model's accuracy, and a qualitative evaluation based on an online survey, where subjects listened and compared human and machine generated performances.

4.1 Output information

In the following sections we will explain the different outputs obtained from our model.

4.1.1 MIDI representation of a embellished musical fragment

The initial output of our model is an expressive score induced by the machine. It consists of a quantized Midi score with some induced embellishments. Figure 4.1 shows the MIDI roll of the original inexpressive score (top), the performed score

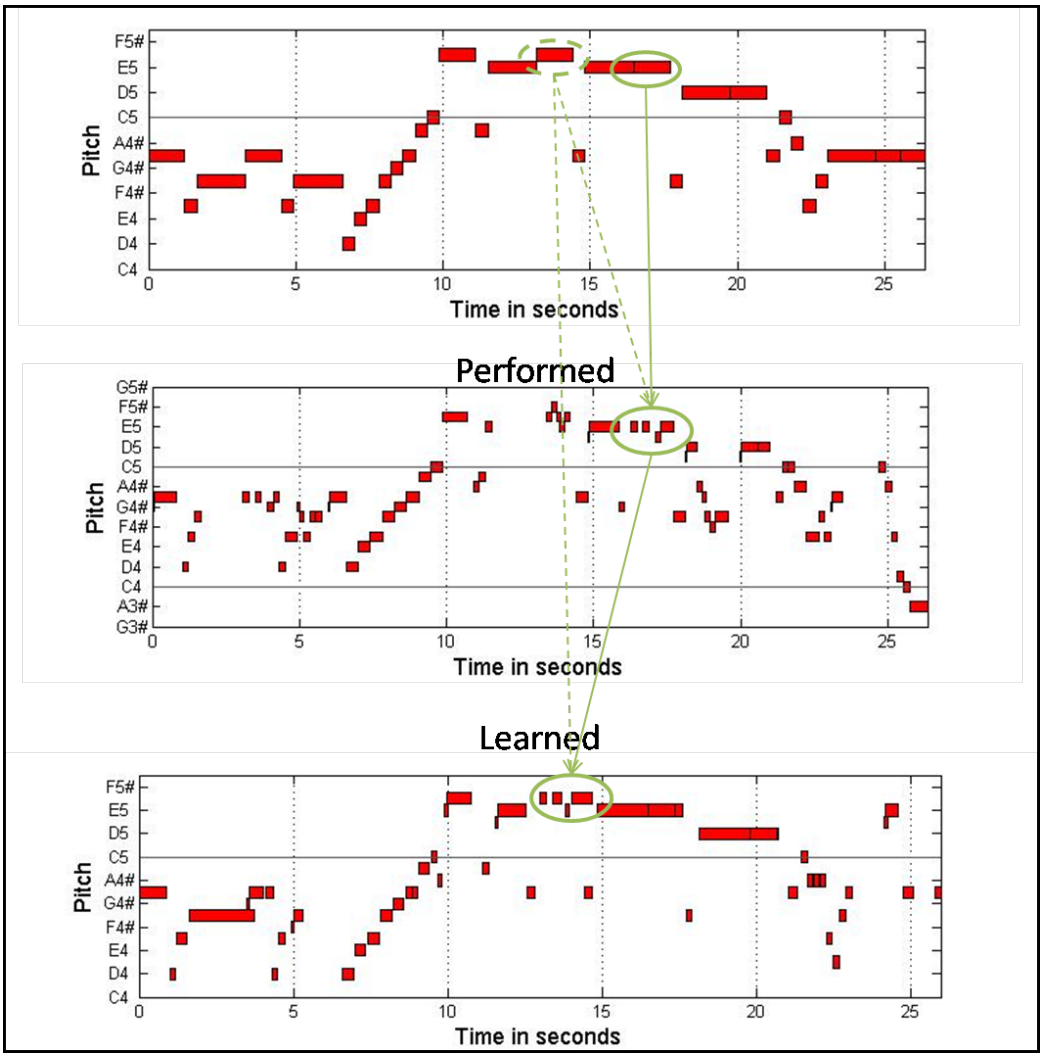


Figure 4.1: Piano roll of a melody fragment. Score (top), performed (center), and predicted (bottom). Circles show an embellishment done by the performer, and how the machine placed it in a similar note context. Test and train scores are the same (first experiment)

(center), and the predicted by the systems (bottom). Circles show the embellished notes, learned from the performed score and predicted in the output score. In Figure 4.2 a short melody excerpt with same key, similar harmony, and melodic

density is used as test data. The arrows and circles show an example of the embellished learned from the performed score, and predicted on the output score.

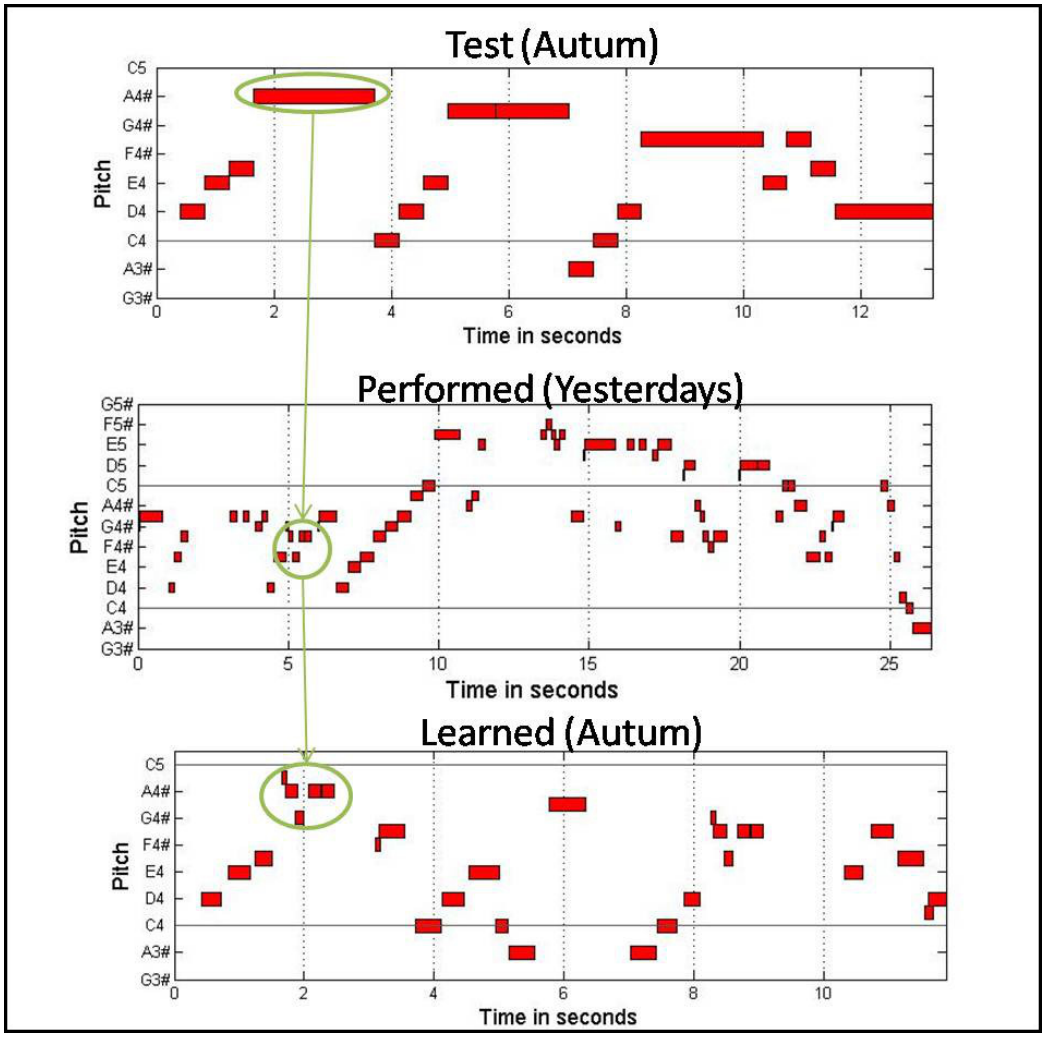


Figure 4.2: New test file, same training set. New Score (top), Performed (center), and predicted (bottom). Circles show an embellishment done by the performer, and how the machine placed it in a similar note context. Test and train scores are different (Second experiment)

4.1.2 Duration modeling

Once we have obtained a embellished score as described above we apply the duration model in order to predict the variations in duration for each note of the embellished score. The duration model takes as input the predicted embellished score, and produces as output an embellished score in which the duration of each note is transformed to the predicted value.

4.1.3 Timing grid

Once we have a embellished score, and each note's duration is set to a predicted value, we apply a beat tempo grid which is the one obtained from the beat tracking of the musical excerpt used as input for our model. Each note onset is moved to this new tempo grid where beats are not quantized. Duration of each note is scaled in proportion to the new duration between two consecutive beats. In other words, our output is an unquantized MIDI score.

4.1.4 Synthesis

All the MIDI excerpts were synthesized, using concatenative synthesis. The performed score was also synthesized using the same approach. This was done in order to give the same conditions to all the musical excerpts at the qualitative evaluation stage, to avoid any bias tendency in the listening. The learned excerpts were synthesized from the midi matrix after applying the duration model. This musical excerpts (Sample1, sample2, etc) were allocated temporally at the following url, to be available for listening:

`http://sergiomusicresearch.wordpress.com/samples/`

The music excerpts are organized as follows:

- Train_1.mp3: Yesterdays as performed by Wes Montgomery, synthesized form the transcription score.
- Test_1.mp3: Yesterdays score with no embellishments, as written in Real-Fake Books.
- Learned_1.mp3: Yesterdays learned by the model from same train same test set experiment.
- Test_2.mp3: Autum score with no embellishments, as written in Real-Fake Books.
- Learned_2.mp3: Autum learned by the model form same train set different test set experiment.

4.2 Evaluation

In order to evaluate our model we conducted a quantitative evaluation based on the accuracy measures obtained from the machine learning experiments, and a qualitative evaluation based on a listening questionnaire involving the music excerpts obtained after modelling and synthesis process.

4.2.1 Quantitative Evaluation

Quantitative evaluation was performed using Correctly Classified Instances (CCI) in the classification task (embellish or not embellish), and Correlation coefficient for the regression experiment (Energy and duration ratio).

Table 4.1: Accuracy measures obtained. The table shows Correlation coefficient for the duration and Energy experiment. For the embellishment experiments it shows correctly classified instances. Experiments were run for different algorithms, using 10 cross fold validation.

MODEL	DUR. CC	ENER. CC	EMBEL. CCI
K*	0.4904	0.7094	81,25
K-NN(K=1)	0.7431	0.3920	78,13
ANN	0.5636	0.4330	75.00
SVM	0.5915	0.3605	65.13

Lazy methods (k-NN and K*) were found to be the most consistent. Table 4.1 shows the correlation coefficients (CC) and the correctly classified instances percentage (CCI%) obtained with the different methods.

Figure 4.3 shows the energy and duration ratio predictions obtained with k-NN and K*, respectively. The curves show duration deviation with respect to the score, and energy deviation with respect to the mean loudness. The darker line correspond to the deviations performed by the musician, and the gray line correspond to the deviation predicted by the model. In both figures it is shown how the model follows in a consistent way the energy and duration deviations done by the performer.

4.2.2 Qualitative Evaluation

For the qualitative evaluation an on-line survey was performed. Subjects were asked to to listen to the synthesized musical fragments and compare how human the modelled music fragments with respect to the ones obtained from the performance of the professional musician, as well as, with respect to the ones synthesized from the plain score with no embellishments. Ideally, if the learnt

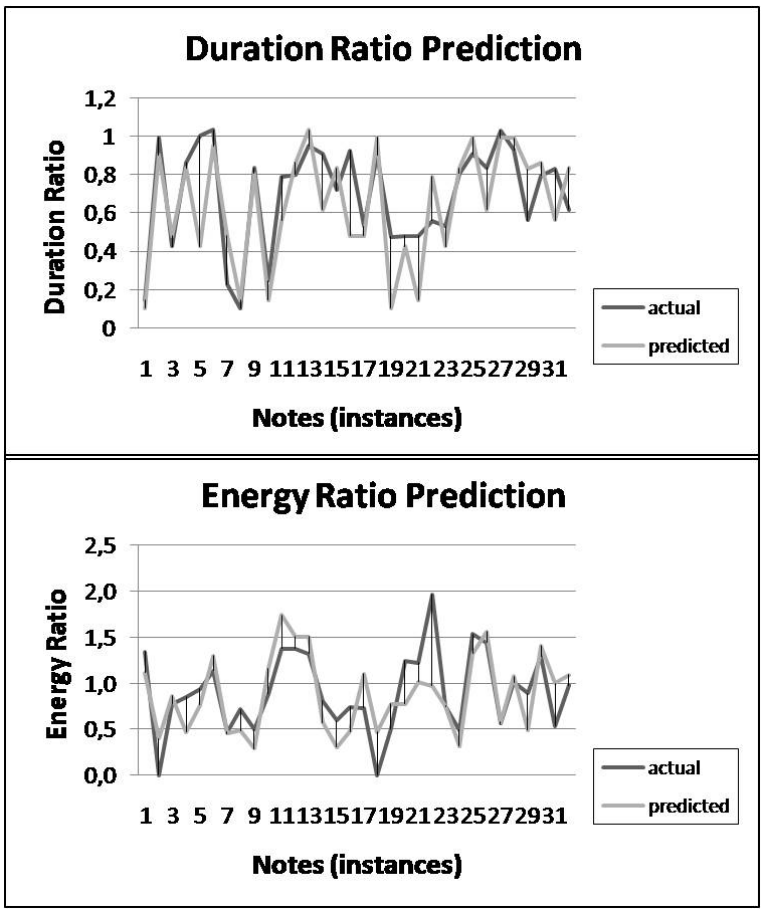


Figure 4.3: Duration and energy ratio predictions: The darker line correspond to the actual deviations in both energy and duration ratio. The gray line correspond to the predicted deviations in energy and duration. It can be seen how both lines follows almost the same path, meaning the model predict the deviations accurately. The predictions were obtained using K* algorithm.

models indeed capture the musician expression, it would be expected that the listeners would have difficulties differentiating the human-performed audio and the machine-generated one. Additionally, it would be expected that the listeners would easily distinguish the audio corresponding to the inexpressive original score from the audio corresponding to the machine-generated performance. This

way we wanted to know how close the model can imitate a human performance, when listened by people.

Other relevant data we included in the survey was the musical background of the subjects. This is important fact as the listening criteria may differ considerably between subjects. Other less relevant data included were the age and gender of the subjects.

In order to obtain more data and make the experiment more accurate, we decide to cut the musical excerpt into short musical phrases. This way the listening would not be biased by previous knowledge of the song melody which may create some expectations on the listener. Also this segmentation gave us more data to be analysed, and this way we may be able to infer more strong claims as results from the data. The musical excerpts were then segmented and organized for comparison in the following way:

- Question 1: Yesterdays segment Y 3, performed vs induced.

  Induced: Sample_1.

  Performed: Sample_2

- Question 2: Yesterdays segment Y 3, score vs induced.

  Score: Sample_3.

  Induced: Sample_4.

- Question 3: Autum segment A 1, score vs induced.

  Induced: Sample_5.

  Score: Sample_6.

- Question 4: Yesterdays segment Y 4, performed vs induced.

  Induced: Sample_7.

  Performed: Sample_8.

- Question 5: Yesterdays segment Y 2, score vs induced.

  Induced: Sample_9.

  Score: Sample_10.

- Question 6: Yesterdays segment Y 4, score vs induced.

  Induced: Sample_11.

  score: Sample_12.

- Question 7: Yesterdays segment Y 2, performed vs induced.

  Performed: Sample_13.

  Induced: Sample_14.

- Question 8: Yesterdays segment Y 1, score vs induced.

  Score: Sample_15.

  Induced: Sample_16.

- Question 9: Yesterdays segment Y 1, performed vs induced.

  Performed: Sample_17.

  Induced: Sample_18.

- Question 10: Autum segment A 2: score vs induced.

  Score: Sample_19.

Induced: Sample_20.

Internet Links

The survey form can be found at:

`https://docs.google.com/spreadsheet/viewform?fromEmail=true&formkey=dE9sZ0JRTm01MHJtMkVNN0pjR2NaUXc6MQ`

The samples can be heard at:

`http://sergiomusicresearch.wordpress.com/`

The results of the survey are presented in Appendix B. A table with all the collected responses of 54 subjects is presented in Apendix B.1. and a summary of the recorded and its graphic representation can be found on Appendix B.2.

The data was collected from the answers from 54 subjects, 63% male, 33% female, with an average age of 33 years old (+- 6 years). Only 33% of the subjects didn't have any musical knowledge, and 22% didn't played any instrument. All the rest had musical training at different degrees in a more or less homogeneous way (See first 3 figures of appendix B.2).

4.3 Discussion

A summary of the responses obtained for all the musical fragments is presented in Table **??**. In general the fragments induced by the machine got less punctuation, than the performed and the score. This means that in general users were bias by how natural the fragments sound, which is linked with how smooth notes transitions were, as well as the notes envelope (e.g. if notes decay too quickly). In this sense, synthesis become a relevant aspect, when evaluating how human a

Table 4.2: Quantitative Evaluation Summary. The question in the form was: Which sample sounds more human?. Table show the number of responses each segment received by listeners. The percentage of responses over total responses for each segment is shown in bold text.

SEGMENT	INDUCED	SCORE	BOTH	NONE
A1	14	27	9	4
A2	14	22	17	2
TOTAL (%)	**26**	**45**	**24**	**6**
SEGMENT	INDUCED	SCORE	BOTH	NONE
Y1	17	21	12	4
Y2	9	24	20	3
Y3	11	29	10	4
Y4	11	26	12	3
TOTAL (%)	**22**	**46**	**25**	**6**
SEGMENT	INDUCED	PERFORMED	BOTH	NONE
Y1	10	20	18	6
Y2	7	22	23	2
Y3	7	22	11	9
Y4	4	33	14	9
TOTAL (%)	**13**	**47**	**31**	**9**

”performance” sound. This issue may be fixed when synthesizing by sampling the complete embellishment instead of doing it note by note, as it was performed in this study. However the option ”none of the samples is human” is the one with less punctuation. As all the samples were generated by synthesis, this confirms how synthesis already induce the deviations that characterize a human performance, buy it also confirms how the embellishments generated show to have good musical sense.

The size of the music excerpts was also an important factor when subjected to evaluation, as some samples were not long enough stablish a musical context. This lead that for some listeners, some samples that were generate by the machine, and even some belonging to the real performance sounded like randomly generated.

Also some ornamentations ended abruptly making them to sound less natural.

The ”human” term could be some how ambiguous as it may refer to the smoothness of the performance or to the compositional aspect. For future work, this aspect may be considered to be able to properly evaluate the embellishments incorporated to the score.

In general we found that the qualitative evaluation is subjected to different perceptual aspects that are out of the main focus of this thesis work. This give more relevance to the quantitative result that show that the features selection for each model, and the algorithms chosen for each of the learning tasks could follow in a consistent way the deviations of the real performance.

Chapter 5

CONCLUSIONS

In this study we have applied machine learning techniques in an attempt to recreate musical expression in jazz guitar music, by training models for duration (timing), energy and embellishment transformations. We have applied different Machine Learning Algorithms and found that Lazy methods (K) were the ones that show the best accuracy in the experiments. We have selected the most relevant features for each experiment based on musical knowledge as well as on performance indicators. The results of the quantitative evaluation seem to indicate that the features extracted contain sufficient information allowing learning models to accurately capture the considered transformations.

We developed a computational model in Matlab that takes as inputs an inexpressive MIDI score, a monophonic audio performance of that score (performed by a professional musician) with its corresponding backing track and MIDI transcription and a test MIDI score; and outputs a MIDI score with predicted embellishments, and an audio excerpt synthesized using note samples from the training audio performance. From this model it will be possible to scope the future investi-

gation towards an implementation of a software or plugin to aid education and/or composition, able to learn form performed audio and retrieve an audio excerpt of the learnt performance aspects of a certain performer.

By following our methodology to apply the transformations in to new scores, we were able to obtain some musical excerpts with overall good musical sense. Qualitative evaluation indicates that synthesis is a relevant aspect in the way that people perceive music as played by a human or by a machine. Relevant facts include the way that notes are sustained and transitions between notes, and also the linearity or non linearity of the final melody after embellishments transformation (e.g. large intervals in between one single melodic phrase). These aspects will be taken into account on the PHD investigation that will be carried out from this study.

5.1 Contributions

- Most of the research has been carried out in classical piano music. As far to our knowledge it has not been done in jazz guitar

- Our study is focused in a research area somewhere between Classical music where ornaments are written, and fully improvised music where there are no melodic information at all (Figure 5.1 . As far to our knowledge there are not any previous studies of melodic embellishments of this complexity.

- Our work has shown to have relevance in the research community as it was accepted and presented at at the 5th International Workshop on Machine Learning and Music, held in Conjunction with the International Conference

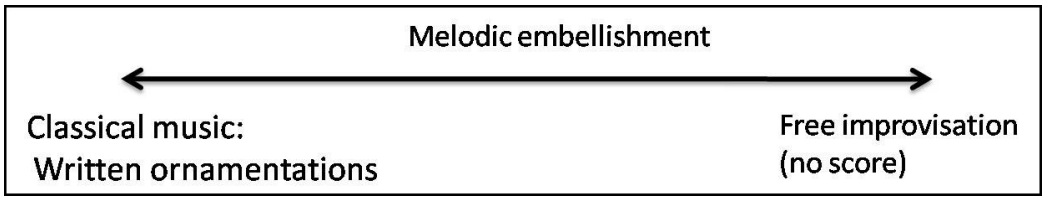


Figure 5.1: Contributions. Our study focuses on a place between written ornamentations and improvisation

on Machine Learning (ICML, June 2012).

- Along with the main model implementation, we have implemented some interesting functions for music analysis such as a Narmour structure finder, and a intra and inter note feature descriptor data base creator. This work can be useful for future work in similar investigations.

5.2 Future Work

Future work have been proposed an accepted as PHD investigation at the Popeu Fabra University. The main objectives that will be covered are:

- To improve the methodology developed for the master thesis project, as well as to obtain a more robust model, by automatizing the note description extraction process, and also improving the machine learning techniques used in the preliminary tests by re-implementing and tuning the parameters for the algorithms used.

- To improve the segmentation stage of the process by automatizing the onset and offset detection on the sound wave form, by finding the envelop of the

wave form and defining filters based on the tesitura of the musical instrument and tuning the decibel thresholds to avoid noise problems.

- To enlarge our data set by gathering data from different performers, playing different tunes styles, in order to have data with varied tempo, note density and harmonic progressions.

- To improve the previous concatenative synthesis approach to obtain a synthesized version of the learned score, by taking into account parameters such as the note envelope (attack, sustain and decay), and also the transition between notes (legato).

- To implement the model into a software or plugin able to predict a expressive performance of a given inexpressive score, using the type of embellishment and transformations that a certain performer would use, based on a audio recordings of that particular performer.

Bibliography

[Rea, 2004] (2004). *The Real Book*. Hall Leonard, Milwaukee, W, USA.

[Baelen and Raedt, 1997] Baelen, E. V. and Raedt, L. D. (1997). Analysis and prediction of piano performances using inductive logic programming. In *Selected Papers from the 6th International Workshop on Inductive Logic Programming*, ILP '96, pages 55–71, London, UK, UK. Springer-Verlag.

[Berendt, 2010] Berendt, J. (2010). *El Jaz: de Nueva Orleands a los Aos 80*. Fondo de Cultura Ecommica de Espaa, Madrid, Spain.

[Bonada et al., 2011] Bonada, J., Serra, X., Amatriain, X., and Loscos, A. (2011). Spectral processing. In Zolzer, U., editor, *DAFX Digital Audio Effects*, pages 393–444. John Wiley and Sons Ltd.

[Bresin,] Bresin, R. Articulation Rules for Automatic Music Performance. In *Proceedings of the 2001 International Computer Music Conference (ICMC 01)*, San Francisco. International Computer Music Association.

[Cambouropoulosr, 2000] Cambouropoulosr, E. (2000). Score extraction from midi files. In *In In Proceedings of the 13th Colloquium on Musical Informatics (CIM 2000)l*, pages 381–386, LAquila, Italy.

[Cano et al., 1999] Cano, P., Loscos, A., and Bonada, J. (1999). Score-performance matching using hmms. In *In Proceedings of the ICMC*, pages 441–444.

[Choi, 2007] Choi, A. (2007). Chart translate @ONLINE.

[Crook, 2002] Crook, H. (2002). *How to Improvise*. Advance Music.

[de Mantaras and Arcos, 2002] de Mantaras, R. L. and Arcos, J. (2002). Ai and music from composition to expressive performance. *AI Mag.*, 23(3):43–57.

[de Mantaras et al., 2007] de Mantaras, R. L., Grachten, M., and Arcos, J. L. (2007). Playing with cases: Tempo transformations of jazz performances using case-based reasoning. In *FLAIRS Conference'07*, pages 8–11.

[Dixon, 2000] Dixon, S. (2000). Extraction of musical performance parameters from audio data. In *in Proc. IEEE Pacific-Rim Conf. on Multimedia*, pages 42–45.

[Dixon, 2004] Dixon, S. (2004). Analysis of musical content in digital audio. In *Computer Graphics and Multimedia: Applications, Problems, and Solutions*, pages 214–235. Idea Group.

[Dixon and Widmer, 2005] Dixon, S. and Widmer, G. (2005). Match: A music alignment tool chest. In *6 th International Conference on Music Information Retrival (ISMIR*.

[Dovey, 1995] Dovey, M. J. (1995). Analysis of rachmaninoff's piano performances using inductive logic programming (extended abstract). In *Proceed-*

ings of the 8th European Conference on Machine Learning, ECML '95, pages 279–282, London, UK, UK. Springer-Verlag.

[Eerola and Toiviainen, 2004] Eerola, T. and Toiviainen, P. (2004). *MIDI Toolbox: MATLAB Tools for Music Research*. University of Jyvskyl, Kopijyv, Jyvskyl, Finland.

[Ellis, 2007] Ellis, D. (2007). Beat tracking by dynamic programming. *J. New Music Research, Special Issue on Beat and Tempo Extraction*, 36(1):51–60.

[Friberg et al., 2006] Friberg, A., Bresin, R., and Sundberg, J. (2006). Overview of the KTH rule system for musical performance. *Advances in Cognitive Psychology, Special Issue on Music Performance*, 2(2-3):145–161.

[Gabrielsson, 1999] Gabrielsson, A. (1999). Music performance. In D., D., editor, *Psychology of Music*, pages 501–602. Academic Press, Sandiego.

[Giraldo and Ramirez, 2012] Giraldo, S. and Ramirez, R. (June 30, 2012). Modeling embellishment, timing and energy expressive transformations in jazz guitar. *Proc. of the 5th International Workshop on Machine Learning and Music (MML 2012)*.

[Goebl et al., 2008] Goebl, W., Dixon, S., Poli, G. D., Friberg, A., Bresin, R., and Widmer, G. (2008). sense in expressive music performance: Data acquisition, computational studies, and models. In *Sound to Sense, Sense to Sound A State of the Art in Sound and Music Computing*, pages 196–242. Polotti and Rocchesso.

[Gómez et al., 2003] Gómez, E., Klapuri, A., and Meudic, B. (2003). Melody description and extraction in the context of music content processing. *Journal of New Music Research*, 32. Content Based Retrieval.

[Grachten et al., 2006] Grachten, M., Arcos, J.-L., and Mántaras, R. L. (2006). A case based approach to expressivity-aware tempo transformation. *Mach. Learn.*, 65(2-3):411–437.

[Hall et al., 2009] Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., and Witten, I. H. (2009). The weka data mining software: an update. *SIGKDD Explor. Newsl.*, 11(1):10–18.

[Jones et al., 01] Jones, E., Oliphant, T., Peterson, P., et al. (2001–). SciPy: Open source scientific tools for Python.

[Juslin, 2001] Juslin, P. (2001). Communicating emotion in music performance: A review and a theoretical framework. In Juslin and Sloboda, editors, *Music and emotion: Theory and research.Series in affective science.*, pages 309–337. Oxford University Press, New York.

[Klapuri, 1999] Klapuri, A. (1999). Sound onset detection by applying psychoacoustic knowledge. In *Proceedings of the Acoustics, Speech, and Signal Processing, 1999. on 1999 IEEE International Conference, Volume 06*, ICASSP '99, pages 3089–3092, Washington, DC, USA. IEEE Computer Society.

[Laurson et al., 1999] Laurson, M., Hiipakka, J., Erkut, C., Karjalainen, M., Vlimki, V., and Kuuskankare, M. (1999). From expressive notation to model-based sound synthesis: a case study of the acoustic guitar. In *Proceedings of the International Computer Music Conference*, page 14.

[Laurson et al., 2010] Laurson, M., Vlimki, V., and Penttinen, H. (September 6-10, 2010). Simulating idiomatic playing styles in a classical guitar synthesizer: Rasgueado as a case study. *Proc. of the 13th International Conference on Digital Audio Effects (DAFx-10)*.

[Lerdahl and Jackendoff, 1983] Lerdahl, F. and Jackendoff, R. (1983). *A generative theory of tonal music / Fred Lerdahl, Ray Jackendoff*. MIT Press, Cambridge, Mass. :.

[Maestre et al., 2009] Maestre, E., Ramirez, R., Kersten, S., and Serra, X. (2009). Expressive concatenative synthesis by reusing samples from real performance recordings. *Computer Music Journal*, 33:23–42.

[Marshall, 2000] Marshall, W. (2000). *Best of Jazz*. Hall Leonard, Milwaukee, W, USA.

[MATLAB, 2011] MATLAB (2011). *version 7.12.0 (R2010a)*. The MathWorks Inc., Natick, Massachusetts.

[Narmour, 1990] Narmour, E. (1990). *The Analysis and Cognition of Basic Melodic Structures: The Implication-Realization Model*. University of Chicago Press.

[O'Grady and Rickard., 2009] O'Grady, P. D. and Rickard., S. T. (2009). Automatic hexaphonic guitar transcription using non-negative constraintsr. *Proc. IET Irish Signals Syst. Conf. (ISSC2009)*, page 16.

[Plaza and Arcos, 2002] Plaza, E. and Arcos, J. (2002). Constructive adaptation. *Advances in Case-Based Reasoning, in Lecture Notes in Artificial Intelligence*, (2416):306320.

[Ramirez and Hazan, 2006] Ramirez, R. and Hazan, A. (2006). A tool for generating and explaining expressive music performances of monophonic jazz melodies. *International Journal on Artificial Intelligence Tools*, 15:673–691.

[Ramirez et al., 2010] Ramirez, R., Maestre, E., and Serra, X. (2010). Automatic performer identification in commercial monophonic jazz performances. *Pattern Recognition Letters*, 31:1514–1523.

[Serra et al., 1997] Serra, X., Roads, P. I. C., Pope, S., Picialli, A., Poli, G. D., and musical Signal (1997). Musical sound modeling. In *in Musical Signal Processing*.

[Shalev-shwartz et al., 2004] Shalev-shwartz, S., Keshet, J., and Singer, Y. (2004). Learning to align polyphonic music. In *In Proceedings of the 5th International Conference on Music Information Retrieval*, pages 381–386.

[Soulez and et al., 2003] Soulez, F. and et al. (2003). Improving polyphonic and poly-instrumental music to score alignment. In *IN ISMIR*, pages 143–148.

[Temperley, 2001] Temperley, D. (2001). The cognition of basic musical structures. *Mass*.

[Turetsky and Ellis, 2003] Turetsky, R. J. and Ellis, D. P. (2003). Ground-truth transcriptions of real music from force-aligned midi syntheses.

[Widmer, 2001] Widmer, G. (2001). Discovering strong principles of expressive music performance with the plcg rule learning strategy. In De Raedt, L. and Flach, P., editors, *Machine Learning: ECML 2001*, volume 2167 of *Lecture Notes in Computer Science*, pages 552–563. Springer Berlin / Heidelberg.

[Widmer and Tobudic, 2002] Widmer, G. and Tobudic, A. (2002). Playing mozart by analogy: Learning multi-level timing and dynamics strategies. *JOURNAL OF NEW MUSIC RESEARCH*, 32:259–268.

Appendices

Appendix A

ALGORITHMS

In this section we present an overview and summary of the main script created in matlab and a brief explanation of each of the functions developed. All the scripts can be found at:

`http://sergiomusicresearch.wordpress.com/algorithms/`

A.1 File: embellish find.mat

This is the main script of the model. It follows the following stages:

- Get input information: Reads MIDI data and transforms it into numerical matrices. Also reads wav audio into a numerical vector, and reads the score to performance data previously stored in a text file.

- Beat tracking and tempo estimation: Performs beat tracking over the backing track audio file using Dan Ellis beat tracker function. After, it quantize the beats using linear regression, and estimates tempo.

- Calculate descriptors from audio file: It creates a structure array with descriptors calculated from the audio signal, using the function **audio_processing.mat**.
- Prepare MIDI data: MIDI files are quantized, and tempo is set to the one estimated previously
- Calculate descriptors from MIDI data: A descriptor set for each MIDI score is calculated using the function **midi2ds.mat**.
- Create embellishment data set: Embellishment data set is created using the function **embellish.m**. From this data base, we obtain the information to label each note in the training data set as embellished or not (y/n). Also the embellish field for the test data base is set to interrogation marks "?".
- Create energy and duration data sets: It combines the MIDI descriptors and the audio descriptors calculated previously. The fields for RMS ratio and duration are added to each data set respectively.
- Create arff files: Attributes names are created using the function **attributes.m**. After, the arff files are created using the function **arff_write.m**. With this we create arff files for embellishment, energy and duration. Train and test files are created separately.
- WEKA Experiment: A first experiment is performed for embellishment prediction. When the train and test files are the same it performs the 10 cross fold validation experiment 10 times, using the **cross_val.mat** function. Correctly classified instances are calculated each time, and the data with highest score is stored. If train and test are different it performs the train and test experiment normally using the function **weka_run**.

- Embellishment note concatenation: First structures are transformed to matrices, using **struct2matrix.m** function. Then note concatenation is performed using the following processes:

- KNN note search and transformation: The index of the notes predicted to be embellished are stored in a variable. Relevant descriptors and their weight are defined. The note descriptor matrix is normalized using the function **normalizesig.m**. KNN search is performed for each note predicted to be embellished, to find the closest note in the data base of embellishments. If the closest note distance is zero, we use the next closest one. This is done to force the system to use a different ornamentation if the note found is exactly the same as the one being searched. We got track of the similar notes found in order to check that the resulting pairs of notes had musical sense.

  Embellish transformation is done creating a structure with the length of the embellishment found. Each MIDI attribute (e.g. duration, pitc, onset, etc) of each of the notes of the the embellishment is transformed to a new value relative to the note which is being transformed, as explained in section 3.4. The resulting structure notes (corresponding to a single embellishment) is converted to a MIDI matrix format. The current embellishment has to be inserted, taking into account how much the total length of the new score has increased up to the moment, due to the embellishments inserted before.

- Note length correction: Searches for overlapping notes and returns a single melodic line. This is done using the function **dur_correct.m**

- Pitch correction: Is performed manually to certain notes on each case. An automated way based on musical theory will be developed as future work.

- Duration test set preparation: First we calculate descriptors using the **midi2ds.m** function. To this data set of descriptors we add a test column for the Duration Ratio prediction, filling it with question marks "?".

- Duration modeling: structure fields are used as attributes names, and relevant descriptors are indexed. Train and test data are created removing the irrelevant fields. Again arff files and attributes list are created, and the Machine Learning experiment is done with the **weka_run.m** function. Finally, each note's duration is transformed to the predicted duration ratio, both in seconds and in beats.

- Synthesis: For a more realistic synthesis we changed the onsets information adjusting it to the original tempo grid of the beat tracker. The function **unquantize.m** is used for this purpose. The function **sampler.m** is used for the concatenative synthesis. At this point we use this function to separately synthesize the induced score, the performance transcription score, the test score, and the train score.

- Visualization: Is the MIDI piano roll of the predicted score.

A.1.1 File: audio processing.m

This function create descriptors fron an audio signal,and a structure containing the MIDI notes information of the score transcription of the audio. The process follows the the following steps:

- Note Segmentation. Onsets in second are found in the melody audio file using the **essentia_onsets.m** function.

- Manual correction can be performed when ever there are more or less notes detected than the MIDI score transcription. If user choose to not perform the manunal correction the system reads the last saved version of the onset text file. Manual correction is performed on a separate function: **onset_manual_corr.m**.

- Duration is calculated taking the difference between onsets and offsets. Onsets are calculated using Essentia, but offsets are manually annotated. Some times, offsets can be taken as the next note onset. However, this assumption can't be generalized, because this can induce errors when short notes are followed by silences. An automated way to calculate onsets and offsets will be left as part of future work.

- Previous duration uses shifts the duration row of the MIDI matrix by one step and copy it in a new field.

- Next duration is calculated as before, the shift is performed on the opposite direction.

- Energy calculation the RMS value for each note is calculated by the following formula:

$$RMS = \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n}}$$

- RMS Ratio Calculation: The same formula is used to calculate the overall RMS value, but this time all the wave form is parsed. Finally, the RMS value of each note is divided by the overall RMS value to obtain the RMS ratio.

A.1.2 File: midi2ds.m

This function transform a MIDI matrix (nmat) to a structure array, and calculate descriptors, based on note information, key information and chord information. It uses the function **chordBBread.m** to read the chord and key information from a band in a box type file.

- Tempo
- onset_b: onsets in beats
- duration_b: Duration in beats
- pitch: MIDI note (1 to 127)
- vel: MIDI velocity, refers to the note loudness (1 to 127)
- onset_s: onset in seconds
- duration_s: duration in seconds
- bar: Bar at which the note occurs
- pre_dur_b: Duration of the previous note calculated by shifting the note duration by -1, in beats.
- nxt_dur_b: Duration of the next note calculated by shifting the note duration by 1, in beats.
- onser_b_mod: Onset beat mod: at which beat of the current bar the note occurs. (0 for 1st beat, 1 for 2nd beat, etc.)

- pitch_mod: pitch in chroma measured in 12 semi tones (0=C, 1=C#Db, 2=D, ... ,11=B).

- note2key: The function creates a string for note labeling, using the ”s” character to handle enharmonics (e.g. C# and Db) as: `notes=CsDsEFsGsAsB`. This way is possible to find the numerical index of a certain note from its name as a string (e.g G# = 7). Notice that notes are indexed from zero to correctly perform intervals operations. After calculating the key index from the Band in a Box type file, this calculates the absolute value of the subtraction of each note chroma and the key index.

- chord: After organizing the chord names and indexes into a row, it maps each chord name to the corresponding notes of the bar/beat at which the chord occurs.

- chord_id: Same as before with the chord index.

- chord_type: It maps the chord type (e.g maj7, m, m7, etc) to its respective note. For doing so, it reads the chord type information from the chord string, discharging the root of the chord.

- note2chord: This is the interval from each note to its corresponding chord root.

- isChordN: Is a chord note? It searches in a data base of chords (see table 3.2) and checks if the current note belongs to this chord description.

- mtr: Is the metrical strength as explained in section 3.3. It uses the beat onset mod information and uses a switch to label when the note occurs on a

certain beat, according to the mtr definition of the mentioned section.

- narmour: The function **narmour_sig.m** was implemented to calculate narmour structures of each note.

A.1.3 File: embellish.m

Given a midi score and a midi performance of that score, we use the midi2ds to calculate note descriptors for each file, and then characterize the embelishments (transformations done) to each note of the score. Each note of the original score can be transformed to a set of notes that correspond to a certain ornamentation. Bebore hand we have labeled on a data base wich notes of the plain score, correspond to the performance transcription. The transformation is stored by beat offset (boff), interval offset (ioff), and duration in beats (already calculated as dur), with respect of the original note. We also calculate ornament duration difference (odd), and number of notes used to embellish. The embellish data base is then created calculating the following fields for each note:

- ioff: Interval offset with respect to original note: here we subtract the pitch of each performance note and the pitch of its corresponding note in the original score.

- boff: Beat offset with respect to original note: here we rest the beat onset of each performed note and its corresponding note in the score. This way we measure anticipations or retardation.

- boff_r: Beat offset ratio. Same as before but we calculate the ratio instead.

- odd: Duration difference with respect to the original note. The problem here are the ornament notes that are the same duration regardless the original duration. In this case we will prefer to use the absolute value of the embellished note, instead of a relative one.

- odd_r: Same as above but we calculate the ratio instead of the difference.

A.1.4 File: attributes.m

This function uses the field names of the train and test structures and convert them in to a cell array with the attribute list following the arff format. After the cell array is used to create the arff file. If checks the class of the data stored on each field and creates a numeric attribute, or a text attribute. In the second case it searches for the unique values in both train and test data bases.

A.1.5 File: arff_write.m

This function writes a arff file form a structure array and a cell array of attributes. The trainOrtest indicator defines the format of the last column whether it is numerical or text. For the experiments purpose the tempo field is not taken into account as we are only working with one song as a training set. Field names are concatenated in a single string as headers. Format is defined as string or float depending on the cell class of each field of the structure. Cell class is searched by the **struc_class.m** function. Finally the data is written to text using Matlab fopen and fprintf functions.

A.1.6 File: cross_val.m

This function performs the cross validation for a machine learning experiment. It uses as input parameters a structure array, the number of folds, a seed indicator (true or false) to use the same seed each time the experiment is performed or a different one otherwise, and finally the classifier or algorithm to use for the experiment. The process is performed in the followin way:

- Randomization: It creates a vector of the size of the data to index the each instance. Another vector is created randomly of the size of the sample, the structure and both first index and randomized index are used to randomize the data in the structure with the function **perm_ds.m**.

- Create folds indexes: Each fold is created using in and out indexes, taking into account the the division of sample size over number of folds may have a remaining. The size of a number of fold samples equal to this remaining is increased by one to have an even distribution. In and out indexes for each fold is stored in a temporal matrix.

- Run machine learning experiment: After attributes are created using the attribute function explained before and we perform the Weka experiment for each fold. The data is reorganized each time using the **subset_ds.m** function.

- Un randomize: After running the experiment for all folds the predictions are reorganized to the original order.

- Evaluation: We compute correctly classified instances by comparing the prediction and the last column of the structure data and increasing a counter

each time a prediction is equal to the actual data. After the percentage is calculated. We also implemented a confusion matrix, it looks for the unique classes in the last column of the data. Then it creates a matrix of the size of those unique classes. On each cell it calculates by increasing a counter how many times the current pair of classes was predicted to be the same.

A.1.7 File: weka_run

This function creates the OS commands to run Weka form the command line, and also to run the Phython script **Clean_PredictionFile.py** [1] to clean the output file of Weka, and returns an array with the predictions. It uses as inputs the train structure data, the test structure data , the attributes list, and a indicator "read data" which is used to read whether the data of the experiment or the previous data stored. It first deletes any previous prediction txt files created and any arff files created. Then it creates the new train and test files with the arff_write.m function. The java command is run from a bat file which is run from Matlab using the system command. Then the output file is cleaned and the predictions are read into a array.

A.1.8 File: struct2matrix.m

This function transforms the structure array data into a regular matrix in order to make possible the use of the knnsearch function. It also changes everything into numerical data: y=1, n=0, mtr:0,1,2,3; Narmour structures: 0 to 8. Chord extensions are converted to a numerical index based on the extension data base described Table 3.2, using the function **chordExtensions.m**. After the structure

[1] S. Gulati, 2012 MTG, Pompeu Fabra University

is converted to cell and then from cell to matrix.

A.1.9 File: normalizesig.m

A simple finction to normalize a matrix. It divides each row cell value by the maximun value of its corresponding column.

A.1.10 File: dur_correct.m

This function searches for notes with a duration longer than the onset difference from the current note and the next note, and adjust the length of the note to this value.

A.1.11 File unquantize.m

This function is to adjust the tempo resolution to a beat detection grid given by a beat detection function. It uses as inputs the MIDI matrix (nmat) and the beat detection vector, with beats onsets in seconds. The process follows these steps:

- First we move beats to zero position.
- After we get the tempo from the midi file.
- We calculate the duration of one beat in seconds.
- Calculate the real duration of each beat by differentiating the beat vector
- Calculate the beat ratio shift of each real beat compared to the midi beat duration.

- Calculate the beat fraction of each note (e.g. if note is at beat 2.3, the beat fraction is 0.3)
- Calculate beat fraction in seconds by multiplying it by the beat duration.
- Calculate the current beat at which the note occurs. Do same calculation in seconds as above.
- Create current beat index by summing 1 to the current beat vector (to avoid zero indexing)
- Calculate the beat offset by subtracting each note's beat indexed by the current beat position in the beat detection grid, from the current beat calculated before.
- calculate the new beat fraction by multiplying the beat ratio indexed by the current beat position by the beat fraction calculated before.
- Calculate new onset by performing the following operation: $Newonset = Currentbeat + beatoffset + newbeatfraction$
- Assign the new onset to the corresponding column of the MIDI matrix.

A.1.12 File: sampler.m

This function synthetize a midi matrix using a sample audio signal. For doing so, we need a data base of the audio descriptors and the midi score of the audio. The process follows the approach of concatenative synthesis by Maestre et al.. It is performed as following:

- Set relevant descriptors and weights

- Normalize the matrix of relevant descriptors using the function explained before.

- Define initial parameters for the sinusoidal plus stochastic model [Bonada, 2011]

- For each note find the closest note on a data base of note samples using KNN search

- Get the pitch and duration of the note found from the MIDI transcription.

- Define a pitch scaling factor form the pitch of the note found in the data base and the note we want to synthesize

- Get the note audio in a vector and apply the transformation using the sps model.

- Adjust the duration: a temporal note is defined to have a duration equal to the onset difference of two consecutive notes (except for the last note). If the transformed note duration is shorter than the length of the temporal note, it is assigned to the temporal note and the rest is filled with silence. On the contrary, the transformed note is shortened, and the **fadeOut.m** function is used to create a decay of 10% if the length of the note. Finally each note is concatenated in a resulting audio vector following the onset information in seconds of the MIDI file.

A.1.13 File: essentia_onsets.m

This function run essentia onset detection function from a the Python scritp **python onset_std.py**. After the essentia function is run, it read the onset from a Yaml file, and returns the onsets in seconds in a vector.

A.1.14 File: onset_manual_corr.m

This function is designed to perform manual correction of the onsets calculated by the Essentia onset detection. It plots the audio using the function **onsets_plot.m** wave and a vertical line over the onsets detected. From the visual inspection the user can manually delete, create, and move a marker. Options for listening, and save the markers into a text file are also available.

A.1.15 File: chordBBread.m

This function uses charttranslate by Choi [Choi, 2007], to create a readable chord chart from band in a box type files. It returns an struct array with the chords and information like name, meter, style and key. Also the tempo, and the number of bars are extracted. Aspects like how many chords are present at a current bar are taken into account. The maximum number of chords per bar allowed are 4. Also inversions or hybrid chords are changed to its corresponding root position. Chord root is indexed using the same enharmonic approach explained in A.1.2.

A.1.16 File: narmour_sig.m

This function calculates the Narmour strucutres on a midi file in nmat format. It parses the nmat matrix and for each note it classify the corresponding narmour

structure for each of the three posible positions. The classification is based on the work by Ramirez et al. 2006 and we have add two more classifiers: two long intervals in the same direction, and, a short interval followed by a long interval in the opposite direction. Long and short interval limit was set to 6 semitones. The process is performed creating a 3 note window and for each window we diferentiate the pitch to get the intervals. This information is parsed using the functions **intsize2.m** for labeling each of the two interval small or large (e.g. both intervals small: ”ss”, large interval after a small interval: ”sl”, etc), and **samedir.m** which retrieves 1 if notes go to the same direction or 0 if notes move in opposite directions. This two conditions are evaluated simultaneusly and all the possible combinations are labeled as showed in Table **??**

A.1.17 File: struc_class.m

This function retrieves a vector of the same length of the number of fields of a structure, containing and indicator of the class of the cells of a particular field, being 1 for cell class, and zero other wise.

A.1.18 File: perm_ds.m

This function reorganizes the data in a structure given an index of the initial order of the data, and a random index order of the data. Tempo field is not used, so is removed. Each field of the structure is reorganized by indexing each element to the same field indexed by the randomized index.

A.1.19 File: subset_ds.m

This function reorganizes the data of a structure given a range of instances that will be used as training set. This range is set to test structure array and the rest of the data is concatenated in the training structure data.

A.1.20 File: chordExtensions.m

This function reads a chord type database was build with intervals describing note composition.

A.1.21 File: fadeOut.m

This function performs a fade out process on the audio sample by creating a linear function from 1 to zero, and multipliying it to the portion of the wave that will be transformed.

A.1.22 File: onset_std.py

This function is implemented in Python.

A.1.23 File: onsets_plot.m

This function plots the audio wave and on the top of it draw a vertical line numbered at the place where onsets were calculated.

A.1.24 File: intsize2.m

Given three notes' pitch, this function classify each of the two intervals between the notes as large or small, defining the limit in 6 semitones.

A.1.25 File: samedir.m

Given a set of three notes' pitch, this function returns 1 if the intervals betwen the two notes go in the same direction and zero otherwise.

Appendix B

ONLINE SURVEY

B.0.26 Table of responses

B.0.27 Responses summary

Timestamp	How old are you? / Cuantos años tienes?	Gender / Sexo	Years of Musical Training / Años de Formación Musical	Do you play any musical instument? / Tocas algún instrumento musical?		Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?	Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?	Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?	Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?	Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?	Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?	Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?	Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?	Comments / Comentarios	Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?	Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?
8/17/2012 20:24:38	37	Male / Hombre	From 5 to 8 years / De 5 a 8 años	Yes, I have formal education / Si, tengo formación profesional		Both sounds equally "human". / Ambas suenan igualmente "humanas"	SAMPLE_4 sounds more human than SAMPLE_3 / El SAMPLE_4 suena mas humano que el SAMPLE_3.	SAMPLE_5 sounds more human than SAMPLE_6 / El SAMPLE_5 suena mas humano que el SAMPLE_6.	Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_9 sounds more human than SAMPLE_10 / El SAMPLE_9 suena mas humano que el SAMPLE_10.	SAMPLE_11 sounds more human than SAMPLE_12 / El SAMPLE_11 suena mas humano que el SAMPLE_12.	Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_20 sounds more human than SAMPLE_19 / El SAMPLE_20 suena mas humano que el SAMPLE_19.	By Sergio, test data.	SAMPLE_16 sounds more human than SAMPLE_15 / El SAMPLE_16 suena mas humano que el SAMPLE_15.	Both sounds equally human / Ambos suenan igualemente humanos
8/17/2012 20:36:30	17	Male / Hombre	0 years / 0 años	Yes, I play a bit / Si, toco un poco		SAMPLE_2 sounds more human than SAMPLE_1 / El SAMPLE_2 suena mas humano que el SAMPLE_1.	SAMPLE_3 sounds more human than SAMPLE_4 / El SAMPLE_3 suena mas humano que el SAMPLE_4.	SAMPLE_6 sounds more human than SAMPLE_5 / El SAMPLE_6 suena mas humano que el SAMPLE_5.	Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_10 sounds more human than SAMPLE_9 / El SAMPLE_10 suena mas humano que el SAMPLE_9.	SAMPLE_12 sounds more human than SAMPLE_11 / El SAMPLE_12 suena mas humano que el SAMPLE_11.	Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_19 sounds more human than SAMPLE_20 / El SAMPLE_19 suena mas humano que el SAMPLE_20.	Muy bien !!	SAMPLE_15 sounds more human than SAMPLE_16 / El SAMPLE_15 suena mas humano que el SAMPLE_16.	SAMPLE_18 sounds more human than SAMPLE_17 / El SAMPLE_18 suena mas humano que el SAMPLE_17.
8/20/2012 9:35:06	37	Male / Hombre	From 5 to 8 years / De 5 a 8 años	Yes, I have formal education / Si, tengo formación profesional		SAMPLE_1 sounds more human than SAMPLE_2 / El SAMPLE_1 suena mas humano que el SAMPLE_2.	SAMPLE_3 sounds more human than SAMPLE_4 / El SAMPLE_3 suena mas humano que el SAMPLE_4.	SAMPLE_5 sounds more human than SAMPLE_6 / El SAMPLE_5 suena mas humano que el SAMPLE_6.	SAMPLE_7 sounds more human than SAMPLE_8 / El SAMPLE_7 suena mas humano que el SAMPLE_8.	SAMPLE_9 sounds more human than SAMPLE_10 / El SAMPLE_9 suena mas humano que el SAMPLE_10.	SAMPLE_11 sounds more human than SAMPLE_12 / El SAMPLE_11 suena mas humano que el SAMPLE_12.	SAMPLE_13 sounds more human than SAMPLE_14 / El SAMPLE_13 suena mas humano que el SAMPLE_14.	SAMPLE_19 sounds more human than SAMPLE_20 / El SAMPLE_19 suena mas humano que el SAMPLE_20.	last test checking (erase data)	SAMPLE_15 sounds more human than SAMPLE_16 / El SAMPLE_15 suena mas humano que el SAMPLE_16.	SAMPLE_17 sounds more human than SAMPLE_18 / El SAMPLE_17 suena mas humano que el SAMPLE_18.
8/20/2012 10:12:17	25	Male / Hombre	0 years / 0 años	Yes, I play a bit / Si, toco un poco		SAMPLE_2 sounds more human than SAMPLE_1 / El SAMPLE_2 suena mas humano que el SAMPLE_1.	SAMPLE_3 sounds more human than SAMPLE_4 / El SAMPLE_3 suena mas humano que el SAMPLE_4.	SAMPLE_5 sounds more human than SAMPLE_6 / El SAMPLE_5 suena mas humano que el SAMPLE_6.	Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_11 sounds more human than SAMPLE_12 / El SAMPLE_11 suena mas humano que el SAMPLE_12.	Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_20 sounds more human than SAMPLE_19 / El SAMPLE_20 suena mas humano que el SAMPLE_19.	Me ha costado mucho diferenciar cual es la humana i cual la artificial en practicamente todos los casos.	SAMPLE_16 sounds more human than SAMPLE_15 / El SAMPLE_16 suena mas humano que el SAMPLE_15.	Both sounds equally human / Ambos suenan igualemente humanos
8/20/2012 10:51:46	30	Male / Hombre	From 0 to 3 years / De 0 a 3 años	Yes, I´ve taken lessons / Si, he tomado lecciones		None of them sound "human". / Ninguna suena "humana".	SAMPLE_3 sounds more human than SAMPLE_4 / El SAMPLE_3 suena mas humano que el SAMPLE_4.	SAMPLE_6 sounds more human than SAMPLE_5 / El SAMPLE_6 suena mas humano que el SAMPLE_5.	SAMPLE_8 sounds more human than SAMPLE_7 / El SAMPLE_8 suena mas humano que el SAMPLE_7.	SAMPLE_10 sounds more human than SAMPLE_9 / El SAMPLE_10 suena mas humano que el SAMPLE_9.	SAMPLE_12 sounds more human than SAMPLE_11 / El SAMPLE_12 suena mas humano que el SAMPLE_11.	Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_19 sounds more human than SAMPLE_20 / El SAMPLE_19 suena mas humano que el SAMPLE_20.		SAMPLE_15 sounds more human than SAMPLE_16 / El SAMPLE_15 suena mas humano que el SAMPLE_16.	None of them sound human / Ninguno suena human
8/20/2012 11:03:31	49	Male / Hombre	More than 8 years / Más de 8 años	Yes, I have formal education / Si, tengo formación profesional		SAMPLE_2 sounds more human than SAMPLE_1 / El SAMPLE_2 suena mas humano que el SAMPLE_1.	SAMPLE_3 sounds more human than SAMPLE_4 / El SAMPLE_3 suena mas humano que el SAMPLE_4.	SAMPLE_5 sounds more human than SAMPLE_6 / El SAMPLE_5 suena mas humano que el SAMPLE_6.	SAMPLE_8 sounds more human than SAMPLE_7 / El SAMPLE_8 suena mas humano que el SAMPLE_7.	SAMPLE_10 sounds more human than SAMPLE_9 / El SAMPLE_10 suena mas humano que el SAMPLE_9.	SAMPLE_11 sounds more human than SAMPLE_12 / El SAMPLE_11 suena mas humano que el SAMPLE_12.	SAMPLE_14 sounds more human than SAMPLE_13 / El SAMPLE_14 suena mas humano que el SAMPLE_13.	SAMPLE_19 sounds more human than SAMPLE_20 / El SAMPLE_19 suena mas humano que el SAMPLE_20.	note length or sustain in some cases was a factor. Also speed of grace notes.	SAMPLE_16 sounds more human than SAMPLE_15 / El SAMPLE_16 suena mas humano que el SAMPLE_15.	Both sounds equally human / Ambos suenan igualemente humanos
8/20/2012 13:08:08	34	Male / Hombre	From 0 to 3 years / De 0 a 3 años	Yes, I play a bit / Si, toco un poco		None of them sound "human". / Ninguna suena "humana".	SAMPLE_3 sounds more human than SAMPLE_4 / El SAMPLE_3 suena mas humano que el SAMPLE_4.	SAMPLE_6 sounds more human than SAMPLE_5 / El SAMPLE_6 suena mas humano que el SAMPLE_5.	Both sounds equally human / Ambos suenan igualemente humanos	None of them sound human / Ninguno suena human	SAMPLE_11 sounds more human than SAMPLE_12 / El SAMPLE_11 suena mas humano que el SAMPLE_12.	SAMPLE_13 sounds more human than SAMPLE_14 / El SAMPLE_13 suena mas humano que el SAMPLE_14.	SAMPLE_19 sounds more human than SAMPLE_20 / El SAMPLE_19 suena mas humano que el SAMPLE_20.		Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos
8/20/2012 15:16:53	24	Male / Hombre	0 years / 0 años	Yes, I play a bit / Si, toco un poco		SAMPLE_2 sounds more human than SAMPLE_1 / El SAMPLE_2 suena mas humano que el SAMPLE_1.	SAMPLE_3 sounds more human than SAMPLE_4 / El SAMPLE_3 suena mas humano que el SAMPLE_4.	None of them sound human / Ninguno suena human	SAMPLE_8 sounds more human than SAMPLE_7 / El SAMPLE_8 suena mas humano que el SAMPLE_7.	Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_12 sounds more human than SAMPLE_11 / El SAMPLE_12 suena mas humano que el SAMPLE_11.	Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_19 sounds more human than SAMPLE_20 / El SAMPLE_19 suena mas humano que el SAMPLE_20.		Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos
8/20/2012 16:24:21	35	Male / Hombre	More than 8 years / Más de 8 años	Yes, I have formal education / Si, tengo formación profesional		SAMPLE_2 sounds more human than SAMPLE_1 / El SAMPLE_2 suena mas humano que el SAMPLE_1.	SAMPLE_3 sounds more human than SAMPLE_4 / El SAMPLE_3 suena mas humano que el SAMPLE_4.	None of them sound human / Ninguno suena human	SAMPLE_8 sounds more human than SAMPLE_7 / El SAMPLE_8 suena mas humano que el SAMPLE_7.	Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_12 sounds more human than SAMPLE_11 / El SAMPLE_12 suena mas humano que el SAMPLE_11.	Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_19 sounds more human than SAMPLE_20 / El SAMPLE_19 suena mas humano que el SAMPLE_20.	Algunos samples dan la impresion de que son mitad tocados, mitad sampleados. Hay uno, por ejemplo, que es el mismo motivo tocado dos veces. La segunda suena distinta de la primera - aparece por primera vez el motivo, despues hay un ruido, cambia el sonido y luego repiten el motivo.	SAMPLE_15 sounds more human than SAMPLE_16 / El SAMPLE_15 suena mas humano que el SAMPLE_16.	None of them sound human / Ninguno suena human
8/20/2012 16:57:05	28	Male / Hombre	More than 8 years / Más de 8 años	Yes, I have formal education / Si, tengo formación profesional		Both sounds equally "human". / Ambas suenan igualmente "humanas"	Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_6 sounds more human than SAMPLE_5 / El SAMPLE_6 suena mas humano que el SAMPLE_5.	Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_10 sounds more human than SAMPLE_9 / El SAMPLE_10 suena mas humano que el SAMPLE_9.	Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos	Overall I experienced that all of the sounds could have been played by humans although in different contexts. I tended to say that one sound was more human when this was perceived as more interesting and thus more likely to be played by a human in the context of "solo improvisation". But I must say that the adjective "human" may be a bit vague. Since the played excerpts are sometime totally different maybe you should specify if one have to judge the performance or the composition.	SAMPLE_16 sounds more human than SAMPLE_15 / El SAMPLE_16 suena mas humano que el SAMPLE_15.	Both sounds equally human / Ambos suenan igualemente humanos
8/20/2012 18:53:35	36	Female / Mujer	From 0 to 3 years / De 0 a 3 años	Yes, I play a bit / Si, toco un poco		Both sounds equally "human". / Ambas suenan igualmente "humanas"	Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos		Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos
8/20/2012 19:43:14	34	Female / Mujer	0 years / 0 años	No		SAMPLE_1 sounds more human than SAMPLE_2 / El SAMPLE_1 suena mas humano que el SAMPLE_2.	SAMPLE_3 sounds more human than SAMPLE_4 / El SAMPLE_3 suena mas humano que el SAMPLE_4.	SAMPLE_5 sounds more human than SAMPLE_6 / El SAMPLE_5 suena mas humano que el SAMPLE_6.	SAMPLE_7 sounds more human than SAMPLE_8 / El SAMPLE_7 suena mas humano que el SAMPLE_8.	SAMPLE_9 sounds more human than SAMPLE_10 / El SAMPLE_9 suena mas humano que el SAMPLE_10.	SAMPLE_11 sounds more human than SAMPLE_12 / El SAMPLE_11 suena mas humano que el SAMPLE_12.	SAMPLE_13 sounds more human than SAMPLE_14 / El SAMPLE_13 suena mas humano que el SAMPLE_14.	SAMPLE_19 sounds more human than SAMPLE_20 / El SAMPLE_19 suena mas humano que el SAMPLE_20.		SAMPLE_15 sounds more human than SAMPLE_16 / El SAMPLE_15 suena mas humano que el SAMPLE_16.	SAMPLE_17 sounds more human than SAMPLE_18 / El SAMPLE_17 suena mas humano que el SAMPLE_18.
8/21/2012 0:04:01	44	Male / Hombre	From 5 to 8 years / De 5 a 8 años	Yes, I´ve taken lessons / Si, he tomado lecciones		SAMPLE_2 sounds more human than SAMPLE_1 / El SAMPLE_2 suena mas humano que el SAMPLE_1.	SAMPLE_4 sounds more human than SAMPLE_3 / El SAMPLE_4 suena mas humano que el SAMPLE_3.	SAMPLE_5 sounds more human than SAMPLE_6 / El SAMPLE_5 suena mas humano que el SAMPLE_6.	SAMPLE_8 sounds more human than SAMPLE_7 / El SAMPLE_8 suena mas humano que el SAMPLE_7.	Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_11 sounds more human than SAMPLE_12 / El SAMPLE_11 suena mas humano que el SAMPLE_12.	SAMPLE_14 sounds more human than SAMPLE_13 / El SAMPLE_14 suena mas humano que el SAMPLE_13.	SAMPLE_20 sounds more human than SAMPLE_19 / El SAMPLE_20 suena mas humano que el SAMPLE_19.		SAMPLE_16 sounds more human than SAMPLE_15 / El SAMPLE_16 suena mas humano que el SAMPLE_15.	Both sounds equally human / Ambos suenan igualemente humanos

Timestamp	How old are you? / Cuantos años tienes?	Gender / Sexo	Years of Musical Training / Años de Formación Musical	Do you play any musical instument? / Tocas algún instrumento musical?		Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?	Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?	Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?	Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?	Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?	Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?	Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?	Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?	Comments / Comentarios	Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?	Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?
8/21/2012 0:14:34	40	Female / Mujer	0 years / 0 años	No		SAMPLE_2 sounds more human than SAMPLE_1 / El SAMPLE_2 suena mas humano que el SAMPLE_1.	SAMPLE_3 sounds more human than SAMPLE_4 / El SAMPLE_3 suena mas humano que el SAMPLE_4.	SAMPLE_6 sounds more human than SAMPLE_5 / El SAMPLE_6 suena mas humano que el SAMPLE_5.	SAMPLE_8 sounds more human than SAMPLE_7 / El SAMPLE_8 suena mas humano que el SAMPLE_7.	SAMPLE_10 sounds more human than SAMPLE_9 / El SAMPLE_10 suena mas humano que el SAMPLE_9.	SAMPLE_11 sounds more human than SAMPLE_12 / El SAMPLE_11 suena mas humano que el SAMPLE_12.	SAMPLE_14 sounds more human than SAMPLE_13 / El SAMPLE_14 suena mas humano que el SAMPLE_13.	SAMPLE_20 sounds more human than SAMPLE_19 / El SAMPLE_20 suena mas humano que el SAMPLE_19.		SAMPLE_15 sounds more human than SAMPLE_16 / El SAMPLE_15 suena mas humano que el SAMPLE_16.	SAMPLE_17 sounds more human than SAMPLE_18 / El SAMPLE_17 suena mas humano que el SAMPLE_18.
8/21/2012 0:28:47	30	Female / Mujer	More than 8 years / Más de 8 años	Yes, I have formal education / Si, tengo formación profesional		SAMPLE_2 sounds more human than SAMPLE_1 / El SAMPLE_2 suena mas humano que el SAMPLE_1.	Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_6 sounds more human than SAMPLE_5 / El SAMPLE_6 suena mas humano que el SAMPLE_5.	SAMPLE_8 sounds more human than SAMPLE_7 / El SAMPLE_8 suena mas humano que el SAMPLE_7.	Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_12 sounds more human than SAMPLE_11 / El SAMPLE_12 suena mas humano que el SAMPLE_11.	SAMPLE_13 sounds more human than SAMPLE_14 / El SAMPLE_13 suena mas humano que el SAMPLE_14.	SAMPLE_19 sounds more human than SAMPLE_20 / El SAMPLE_19 suena mas humano que el SAMPLE_20.	En los samples que me han parecido más humanos no había notas cuyo sonido desaparecía de repente y tampoco había demasiadas notas de alturas muy diferentes en poco tiempo.	SAMPLE_15 sounds more human than SAMPLE_16 / El SAMPLE_15 suena mas humano que el SAMPLE_16.	SAMPLE_17 sounds more human than SAMPLE_18 / El SAMPLE_17 suena mas humano que el SAMPLE_18.
8/21/2012 1:37:23	37	Female / Mujer	0 years / 0 años	No		SAMPLE_2 sounds more human than SAMPLE_1 / El SAMPLE_2 suena mas humano que el SAMPLE_1.	SAMPLE_4 sounds more human than SAMPLE_3 / El SAMPLE_4 suena mas humano que el SAMPLE_3.	SAMPLE_6 sounds more human than SAMPLE_5 / El SAMPLE_6 suena mas humano que el SAMPLE_5.	SAMPLE_8 sounds more human than SAMPLE_7 / El SAMPLE_8 suena mas humano que el SAMPLE_7.	Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos		Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos
8/21/2012 11:34:41	31	Male / Hombre	From 3 to 5 years / De 3 a 5 años	Yes, I´ve taken lessons / Si, he tomado lecciones		SAMPLE_2 sounds more human than SAMPLE_1 / El SAMPLE_2 suena mas humano que el SAMPLE_1.	SAMPLE_3 sounds more human than SAMPLE_4 / El SAMPLE_3 suena mas humano que el SAMPLE_4.	SAMPLE_6 sounds more human than SAMPLE_5 / El SAMPLE_6 suena mas humano que el SAMPLE_5.	SAMPLE_8 sounds more human than SAMPLE_7 / El SAMPLE_8 suena mas humano que el SAMPLE_7.	SAMPLE_10 sounds more human than SAMPLE_9 / El SAMPLE_10 suena mas humano que el SAMPLE_9.	Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_20 sounds more human than SAMPLE_19 / El SAMPLE_20 suena mas humano que el SAMPLE_19.		SAMPLE_16 sounds more human than SAMPLE_15 / El SAMPLE_16 suena mas humano que el SAMPLE_15.	SAMPLE_18 sounds more human than SAMPLE_17 / El SAMPLE_18 suena mas humano que el SAMPLE_17.
8/21/2012 13:01:00	25	Male / Hombre	From 3 to 5 years / De 3 a 5 años	Yes, I´ve taken lessons / Si, he tomado lecciones		SAMPLE_2 sounds more human than SAMPLE_1 / El SAMPLE_2 suena mas humano que el SAMPLE_1.	SAMPLE_3 sounds more human than SAMPLE_4 / El SAMPLE_3 suena mas humano que el SAMPLE_4.	SAMPLE_6 sounds more human than SAMPLE_5 / El SAMPLE_6 suena mas humano que el SAMPLE_5.	SAMPLE_8 sounds more human than SAMPLE_7 / El SAMPLE_8 suena mas humano que el SAMPLE_7.	SAMPLE_10 sounds more human than SAMPLE_9 / El SAMPLE_10 suena mas humano que el SAMPLE_9.	SAMPLE_12 sounds more human than SAMPLE_11 / El SAMPLE_12 suena mas humano que el SAMPLE_11.	SAMPLE_13 sounds more human than SAMPLE_14 / El SAMPLE_13 suena mas humano que el SAMPLE_14.	SAMPLE_19 sounds more human than SAMPLE_20 / El SAMPLE_19 suena mas humano que el SAMPLE_20.		SAMPLE_15 sounds more human than SAMPLE_16 / El SAMPLE_15 suena mas humano que el SAMPLE_16.	SAMPLE_17 sounds more human than SAMPLE_18 / El SAMPLE_17 suena mas humano que el SAMPLE_18.
8/21/2012 13:31:03	34	Female / Mujer	From 0 to 3 years / De 0 a 3 años	Yes, I play a bit / Si, toco un poco		SAMPLE_2 sounds more human than SAMPLE_1 / El SAMPLE_2 suena mas humano que el SAMPLE_1.	SAMPLE_4 sounds more human than SAMPLE_3 / El SAMPLE_4 suena mas humano que el SAMPLE_3.	SAMPLE_5 sounds more human than SAMPLE_6 / El SAMPLE_5 suena mas humano que el SAMPLE_6.	SAMPLE_8 sounds more human than SAMPLE_7 / El SAMPLE_8 suena mas humano que el SAMPLE_7.	SAMPLE_9 sounds more human than SAMPLE_10 / El SAMPLE_9 suena mas humano que el SAMPLE_10.	SAMPLE_11 sounds more human than SAMPLE_12 / El SAMPLE_11 suena mas humano que el SAMPLE_12.	SAMPLE_13 sounds more human than SAMPLE_14 / El SAMPLE_13 suena mas humano que el SAMPLE_14.	SAMPLE_20 sounds more human than SAMPLE_19 / El SAMPLE_20 suena mas humano que el SAMPLE_19.		SAMPLE_16 sounds more human than SAMPLE_15 / El SAMPLE_16 suena mas humano que el SAMPLE_15.	Both sounds equally human / Ambos suenan igualemente humanos
8/21/2012 18:20:25	28	Female / Mujer	0 years / 0 años	No		Both sounds equally "human". / Ambas suenan igualmente "humanas"	SAMPLE_3 sounds more human than SAMPLE_4 / El SAMPLE_3 suena mas humano que el SAMPLE_4.	SAMPLE_5 sounds more human than SAMPLE_6 / El SAMPLE_5 suena mas humano que el SAMPLE_6.	SAMPLE_8 sounds more human than SAMPLE_7 / El SAMPLE_8 suena mas humano que el SAMPLE_7.	None of them sound human / Ninguno suena human	SAMPLE_11 sounds more human than SAMPLE_12 / El SAMPLE_11 suena mas humano que el SAMPLE_12.	SAMPLE_13 sounds more human than SAMPLE_14 / El SAMPLE_13 suena mas humano que el SAMPLE_14.	None of them sound human / Ninguno suena human		SAMPLE_15 sounds more human than SAMPLE_16 / El SAMPLE_15 suena mas humano que el SAMPLE_16.	SAMPLE_17 sounds more human than SAMPLE_18 / El SAMPLE_17 suena mas humano que el SAMPLE_18.
8/21/2012 20:14:09	24	Female / Mujer	From 0 to 3 years / De 0 a 3 años	Yes, I play a bit / Si, toco un poco		Both sounds equally "human". / Ambas suenan igualmente "humanas"	SAMPLE_4 sounds more human than SAMPLE_3 / El SAMPLE_4 suena mas humano que el SAMPLE_3.	SAMPLE_6 sounds more human than SAMPLE_5 / El SAMPLE_6 suena mas humano que el SAMPLE_5.	SAMPLE_7 sounds more human than SAMPLE_8 / El SAMPLE_7 suena mas humano que el SAMPLE_8.	SAMPLE_10 sounds more human than SAMPLE_9 / El SAMPLE_10 suena mas humano que el SAMPLE_9.	SAMPLE_11 sounds more human than SAMPLE_12 / El SAMPLE_11 suena mas humano que el SAMPLE_12.	SAMPLE_13 sounds more human than SAMPLE_14 / El SAMPLE_13 suena mas humano que el SAMPLE_14.	SAMPLE_20 sounds more human than SAMPLE_19 / El SAMPLE_20 suena mas humano que el SAMPLE_19.		Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos
8/21/2012 20:50:31	35	Male / Hombre	From 3 to 5 years / De 3 a 5 años	Yes, I play a bit / Si, toco un poco		SAMPLE_2 sounds more human than SAMPLE_1 / El SAMPLE_2 suena mas humano que el SAMPLE_1.	SAMPLE_3 sounds more human than SAMPLE_4 / El SAMPLE_3 suena mas humano que el SAMPLE_4.	SAMPLE_6 sounds more human than SAMPLE_5 / El SAMPLE_6 suena mas humano que el SAMPLE_5.	None of them sound human / Ninguno suena human	SAMPLE_10 sounds more human than SAMPLE_9 / El SAMPLE_10 suena mas humano que el SAMPLE_9.	SAMPLE_12 sounds more human than SAMPLE_11 / El SAMPLE_12 suena mas humano que el SAMPLE_11.	SAMPLE_14 sounds more human than SAMPLE_13 / El SAMPLE_14 suena mas humano que el SAMPLE_13.	SAMPLE_19 sounds more human than SAMPLE_20 / El SAMPLE_19 suena mas humano que el SAMPLE_20.	Si en algo ayuda, lo que me hace sentir poco humanas las interpretaciones son algunos cambios bruscos de notas, que aunque podrían hacerse a esa velocidad usando ciertas técnicas (hammer-on, pull-off o simplemente pulsando muy rápido) suenan muy cortados y poco naturales.	SAMPLE_15 sounds more human than SAMPLE_16 / El SAMPLE_15 suena mas humano que el SAMPLE_16.	None of them sound human / Ninguno suena human
8/22/2012 0:06:50	34	Male / Hombre	More than 8 years / Más de 8 años	Yes, I have formal education / Si, tengo formación profesional		None of them sound "human". / Ninguna suena "humana".	Both sounds equally human / Ambos suenan igualemente humanos	None of them sound human / Ninguno suena human	SAMPLE_8 sounds more human than SAMPLE_7 / El SAMPLE_8 suena mas humano que el SAMPLE_7.	Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_12 sounds more human than SAMPLE_11 / El SAMPLE_12 suena mas humano que el SAMPLE_11.	SAMPLE_13 sounds more human than SAMPLE_14 / El SAMPLE_13 suena mas humano que el SAMPLE_14.	SAMPLE_19 sounds more human than SAMPLE_20 / El SAMPLE_19 suena mas humano que el SAMPLE_20.		Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_18 sounds more human than SAMPLE_17 / El SAMPLE_18 suena mas humano que el SAMPLE_17.
8/22/2012 0:20:07	40	Male / Hombre	More than 8 years / Más de 8 años	Yes, I have formal education / Si, tengo formación profesional		SAMPLE_2 sounds more human than SAMPLE_1 / El SAMPLE_2 suena mas humano que el SAMPLE_1.	SAMPLE_3 sounds more human than SAMPLE_4 / El SAMPLE_3 suena mas humano que el SAMPLE_4.	SAMPLE_6 sounds more human than SAMPLE_5 / El SAMPLE_6 suena mas humano que el SAMPLE_5.	SAMPLE_8 sounds more human than SAMPLE_7 / El SAMPLE_8 suena mas humano que el SAMPLE_7.	Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_20 sounds more human than SAMPLE_19 / El SAMPLE_20 suena mas humano que el SAMPLE_19.		SAMPLE_15 sounds more human than SAMPLE_16 / El SAMPLE_15 suena mas humano que el SAMPLE_16.	SAMPLE_17 sounds more human than SAMPLE_18 / El SAMPLE_17 suena mas humano que el SAMPLE_18.
8/22/2012 21:10:47	27	Male / Hombre	More than 8 years / Más de 8 años	Yes, I have formal education / Si, tengo formación profesional		None of them sound "human". / Ninguna suena "humana".	SAMPLE_3 sounds more human than SAMPLE_4 / El SAMPLE_3 suena mas humano que el SAMPLE_4.	SAMPLE_6 sounds more human than SAMPLE_5 / El SAMPLE_6 suena mas humano que el SAMPLE_5.	SAMPLE_8 sounds more human than SAMPLE_7 / El SAMPLE_8 suena mas humano que el SAMPLE_7.	SAMPLE_10 sounds more human than SAMPLE_9 / El SAMPLE_10 suena mas humano que el SAMPLE_9.	SAMPLE_12 sounds more human than SAMPLE_11 / El SAMPLE_12 suena mas humano que el SAMPLE_11.	SAMPLE_14 sounds more human than SAMPLE_13 / El SAMPLE_14 suena mas humano que el SAMPLE_13.	SAMPLE_19 sounds more human than SAMPLE_20 / El SAMPLE_19 suena mas humano que el SAMPLE_20.		SAMPLE_16 sounds more human than SAMPLE_15 / El SAMPLE_16 suena mas humano que el SAMPLE_15.	SAMPLE_18 sounds more human than SAMPLE_17 / El SAMPLE_18 suena mas humano que el SAMPLE_17.
8/23/2012 13:15:21	38	Male / Hombre	More than 8 years / Más de 8 años	Yes, I have formal education / Si, tengo formación profesional		SAMPLE_2 sounds more human than SAMPLE_1 / El SAMPLE_2 suena mas humano que el SAMPLE_1.	SAMPLE_3 sounds more human than SAMPLE_4 / El SAMPLE_3 suena mas humano que el SAMPLE_4.	SAMPLE_6 sounds more human than SAMPLE_5 / El SAMPLE_6 suena mas humano que el SAMPLE_5.	SAMPLE_8 sounds more human than SAMPLE_7 / El SAMPLE_8 suena mas humano que el SAMPLE_7.	SAMPLE_10 sounds more human than SAMPLE_9 / El SAMPLE_10 suena mas humano que el SAMPLE_9.	SAMPLE_12 sounds more human than SAMPLE_11 / El SAMPLE_12 suena mas humano que el SAMPLE_11.	SAMPLE_13 sounds more human than SAMPLE_14 / El SAMPLE_13 suena mas humano que el SAMPLE_14.	Both sounds equally human / Ambos suenan igualemente humanos	Actually some answers were biased by the "musical feeling" of the riffs. I've tried to decide between pairs of samplers, avoiding "both equal" options whenever possible. Good work, all the best!! Enric:-)	SAMPLE_16 sounds more human than SAMPLE_15 / El SAMPLE_16 suena mas humano que el SAMPLE_15.	SAMPLE_17 sounds more human than SAMPLE_18 / El SAMPLE_17 suena mas humano que el SAMPLE_18.

Timestamp	How old are you? / Cuantos años tienes?	Gender / Sexo	Years of Musical Training / Años de Formación Musical	Do you play any musical instument? / Tocas algún instrumento musical?		Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?	Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?	Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?	Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?	Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?	Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?	Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?	Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?	Comments / Comentarios	Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?	Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?
8/24/2012 17:12:48	38	Male / Hombre	From 0 to 3 years / De 0 a 3 años	Yes, I play a bit / Si, toco un poco		SAMPLE_2 sounds more human than SAMPLE_1 / El SAMPLE_2 suena mas humano que el SAMPLE_1.	Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_8 sounds more human than SAMPLE_7 / El SAMPLE_8 suena mas humano que el SAMPLE_7.	SAMPLE_10 sounds more human than SAMPLE_9 / El SAMPLE_10 suena mas humano que el SAMPLE_9.	SAMPLE_12 sounds more human than SAMPLE_11 / El SAMPLE_12 suena mas humano que el SAMPLE_11.	SAMPLE_13 sounds more human than SAMPLE_14 / El SAMPLE_13 suena mas humano que el SAMPLE_14.	Both sounds equally human / Ambos suenan igualemente humanos		Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_17 sounds more human than SAMPLE_18 / El SAMPLE_17 suena mas humano que el SAMPLE_18.
8/25/2012 4:11:51	33	Female / Mujer	From 0 to 3 years / De 0 a 3 años	Yes, I´ve taken lessons / Si, he tomado lecciones		None of them sound "human". / Ninguna suena "humana".	None of them sound human / Ninguno suena humano	None of them sound human / Ninguno suena human	SAMPLE_8 sounds more human than SAMPLE_7 / El SAMPLE_8 suena mas humano que el SAMPLE_7.	SAMPLE_10 sounds more human than SAMPLE_9 / El SAMPLE_10 suena mas humano que el SAMPLE_9.	SAMPLE_12 sounds more human than SAMPLE_11 / El SAMPLE_12 suena mas humano que el SAMPLE_11.	SAMPLE_13 sounds more human than SAMPLE_14 / El SAMPLE_13 suena mas humano que el SAMPLE_14.	SAMPLE_19 sounds more human than SAMPLE_20 / El SAMPLE_19 suena mas humano que el SAMPLE_20.		SAMPLE_16 sounds more human than SAMPLE_15 / El SAMPLE_16 suena mas humano que el SAMPLE_15.	SAMPLE_17 sounds more human than SAMPLE_18 / El SAMPLE_17 suena mas humano que el SAMPLE_18.
8/26/2012 14:48:55	54	Male / Hombre	More than 8 years / Más de 8 años	Yes, I have formal education / Si, tengo formación profesional		None of them sound "human". / Ninguna suena "humana".	SAMPLE_3 sounds more human than SAMPLE_4 / El SAMPLE_3 suena mas humano que el SAMPLE_4.	SAMPLE_6 sounds more human than SAMPLE_5 / El SAMPLE_6 suena mas humano que el SAMPLE_5.	SAMPLE_8 sounds more human than SAMPLE_7 / El SAMPLE_8 suena mas humano que el SAMPLE_7.	SAMPLE_10 sounds more human than SAMPLE_9 / El SAMPLE_10 suena mas humano que el SAMPLE_9.	SAMPLE_12 sounds more human than SAMPLE_11 / El SAMPLE_12 suena mas humano que el SAMPLE_11.	Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_19 sounds more human than SAMPLE_20 / El SAMPLE_19 suena mas humano que el SAMPLE_20.	Los glissandos son lo que ha marcado "mi" diferencia. El criterio ha ido variando del sample 1 al 20 porque mi oido se ha ido acostumbrando.	SAMPLE_15 sounds more human than SAMPLE_16 / El SAMPLE_15 suena mas humano que el SAMPLE_16.	SAMPLE_18 sounds more human than SAMPLE_17 / El SAMPLE_18 suena mas humano que el SAMPLE_17.
8/28/2012 20:31:31	31	Male / Hombre	From 3 to 5 years / De 3 a 5 años	Yes, I play a bit / Si, toco un poco		SAMPLE_2 sounds more human than SAMPLE_1 / El SAMPLE_2 suena mas humano que el SAMPLE_1.	SAMPLE_3 sounds more human than SAMPLE_4 / El SAMPLE_3 suena mas humano que el SAMPLE_4.	SAMPLE_5 sounds more human than SAMPLE_6 / El SAMPLE_5 suena mas humano que el SAMPLE_6.	SAMPLE_8 sounds more human than SAMPLE_7 / El SAMPLE_8 suena mas humano que el SAMPLE_7.	Both sounds equally human / Ambos suenan igualemente humanos	None of them sound human / Ninguno suena human	SAMPLE_13 sounds more human than SAMPLE_14 / El SAMPLE_13 suena mas humano que el SAMPLE_14.	Both sounds equally human / Ambos suenan igualemente humanos		SAMPLE_16 sounds more human than SAMPLE_15 / El SAMPLE_16 suena mas humano que el SAMPLE_15.	Both sounds equally human / Ambos suenan igualemente humanos
8/28/2012 20:45:15	30	Male / Hombre	0 years / 0 años	Yes, I´ve taken lessons / Si, he tomado lecciones		SAMPLE_2 sounds more human than SAMPLE_1 / El SAMPLE_2 suena mas humano que el SAMPLE_1.	SAMPLE_4 sounds more human than SAMPLE_3 / El SAMPLE_4 suena mas humano que el SAMPLE_3.	SAMPLE_6 sounds more human than SAMPLE_5 / El SAMPLE_6 suena mas humano que el SAMPLE_5.	Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_10 sounds more human than SAMPLE_9 / El SAMPLE_10 suena mas humano que el SAMPLE_9.	SAMPLE_12 sounds more human than SAMPLE_11 / El SAMPLE_12 suena mas humano que el SAMPLE_11.	SAMPLE_13 sounds more human than SAMPLE_14 / El SAMPLE_13 suena mas humano que el SAMPLE_14.	Both sounds equally human / Ambos suenan igualemente humanos	Divertido!	SAMPLE_15 sounds more human than SAMPLE_16 / El SAMPLE_15 suena mas humano que el SAMPLE_16.	SAMPLE_18 sounds more human than SAMPLE_17 / El SAMPLE_18 suena mas humano que el SAMPLE_17.
8/28/2012 20:56:30	30	Female / Mujer	From 5 to 8 years / De 5 a 8 años	Yes, I have formal education / Si, tengo formación profesional		SAMPLE_1 sounds more human than SAMPLE_2 / El SAMPLE_1 suena mas humano que el SAMPLE_2.	None of them sound human / Ninguno suena humano	Both sounds equally human / Ambos suenan igualemente humanos	None of them sound human / Ninguno suena human	Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_19 sounds more human than SAMPLE_20 / El SAMPLE_19 suena mas humano que el SAMPLE_20.		Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_18 sounds more human than SAMPLE_17 / El SAMPLE_18 suena mas humano que el SAMPLE_17.
8/28/2012 21:15:07	38	Male / Hombre	From 0 to 3 years / De 0 a 3 años	Yes, I play a bit / Si, toco un poco		Both sounds equally "human". / Ambas suenan igualmente "humanas"	Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos	gracias por recordarme lo sordo que soy :)	Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos
8/28/2012 21:16:10	35	Female / Mujer	0 years / 0 años	No		SAMPLE_2 sounds more human than SAMPLE_1 / El SAMPLE_2 suena mas humano que el SAMPLE_1.	SAMPLE_4 sounds more human than SAMPLE_3 / El SAMPLE_4 suena mas humano que el SAMPLE_3.	SAMPLE_6 sounds more human than SAMPLE_5 / El SAMPLE_6 suena mas humano que el SAMPLE_5.	SAMPLE_8 sounds more human than SAMPLE_7 / El SAMPLE_8 suena mas humano que el SAMPLE_7.	Both sounds equally human / Ambos suenan igualemente humanos	None of them sound human / Ninguno suena human	Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_20 sounds more human than SAMPLE_19 / El SAMPLE_20 suena mas humano que el SAMPLE_19.		SAMPLE_16 sounds more human than SAMPLE_15 / El SAMPLE_16 suena mas humano que el SAMPLE_15.	SAMPLE_18 sounds more human than SAMPLE_17 / El SAMPLE_18 suena mas humano que el SAMPLE_17.
8/28/2012 21:24:45	34	Male / Hombre	More than 8 years / Más de 8 años	Yes, I have formal education / Si, tengo formación profesional		Both sounds equally "human". / Ambas suenan igualmente "humanas"	SAMPLE_4 sounds more human than SAMPLE_3 / El SAMPLE_4 suena mas humano que el SAMPLE_3.	Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos		None of them sound human / Ninguno suena human	SAMPLE_17 sounds more human than SAMPLE_18 / El SAMPLE_17 suena mas humano que el SAMPLE_18.
8/28/2012 21:46:22	37	Male / Hombre	From 3 to 5 years / De 3 a 5 años	Yes, I´ve taken lessons / Si, he tomado lecciones		SAMPLE_2 sounds more human than SAMPLE_1 / El SAMPLE_2 suena mas humano que el SAMPLE_1.	SAMPLE_3 sounds more human than SAMPLE_4 / El SAMPLE_3 suena mas humano que el SAMPLE_4.	SAMPLE_6 sounds more human than SAMPLE_5 / El SAMPLE_6 suena mas humano que el SAMPLE_5.	SAMPLE_8 sounds more human than SAMPLE_7 / El SAMPLE_8 suena mas humano que el SAMPLE_7.	Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_12 sounds more human than SAMPLE_11 / El SAMPLE_12 suena mas humano que el SAMPLE_11.	Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos	Luego me explica de que se trata Sergio? Que le vaya bien, un abrazo. Santiago Arteaga	SAMPLE_15 sounds more human than SAMPLE_16 / El SAMPLE_15 suena mas humano que el SAMPLE_16.	SAMPLE_17 sounds more human than SAMPLE_18 / El SAMPLE_17 suena mas humano que el SAMPLE_18.
8/28/2012 22:09:08	48	Female / Mujer	0 years / 0 años	No		SAMPLE_2 sounds more human than SAMPLE_1 / El SAMPLE_2 suena mas humano que el SAMPLE_1.	Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_6 sounds more human than SAMPLE_5 / El SAMPLE_6 suena mas humano que el SAMPLE_5.	Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_10 sounds more human than SAMPLE_9 / El SAMPLE_10 suena mas humano que el SAMPLE_9.	SAMPLE_12 sounds more human than SAMPLE_11 / El SAMPLE_12 suena mas humano que el SAMPLE_11.	SAMPLE_13 sounds more human than SAMPLE_14 / El SAMPLE_13 suena mas humano que el SAMPLE_14.	Both sounds equally human / Ambos suenan igualemente humanos	Suenan muy bajito, o es mi ordenador?	SAMPLE_16 sounds more human than SAMPLE_15 / El SAMPLE_16 suena mas humano que el SAMPLE_15.	Both sounds equally human / Ambos suenan igualemente humanos
8/28/2012 22:09:09	35	Male / Hombre	0 years / 0 años	No		SAMPLE_2 sounds more human than SAMPLE_1 / El SAMPLE_2 suena mas humano que el SAMPLE_1.	None of them sound human / Ninguno suena humano	SAMPLE_5 sounds more human than SAMPLE_6 / El SAMPLE_5 suena mas humano que el SAMPLE_6.	Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_10 sounds more human than SAMPLE_9 / El SAMPLE_10 suena mas humano que el SAMPLE_9.	SAMPLE_12 sounds more human than SAMPLE_11 / El SAMPLE_12 suena mas humano que el SAMPLE_11.	Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos		Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_17 sounds more human than SAMPLE_18 / El SAMPLE_17 suena mas humano que el SAMPLE_18.
8/28/2012 22:18:04	25	Female / Mujer	0 years / 0 años	No		SAMPLE_2 sounds more human than SAMPLE_1 / El SAMPLE_2 suena mas humano que el SAMPLE_1.	SAMPLE_4 sounds more human than SAMPLE_3 / El SAMPLE_4 suena mas humano que el SAMPLE_3.	Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_8 sounds more human than SAMPLE_7 / El SAMPLE_8 suena mas humano que el SAMPLE_7.	Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_11 sounds more human than SAMPLE_12 / El SAMPLE_11 suena mas humano que el SAMPLE_12.	Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_20 sounds more human than SAMPLE_19 / El SAMPLE_20 suena mas humano que el SAMPLE_19.		SAMPLE_16 sounds more human than SAMPLE_15 / El SAMPLE_16 suena mas humano que el SAMPLE_15.	SAMPLE_17 sounds more human than SAMPLE_18 / El SAMPLE_17 suena mas humano que el SAMPLE_18.
8/29/2012 0:19:57	36	Male / Hombre	0 years / 0 años	Yes, I play a bit / Si, toco un poco		Both sounds equally "human". / Ambas suenan igualmente "humanas"	Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos		Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos

Timestamp	How old are you? / Cuantos años tienes?	Gender / Sexo	Years of Musical Training / Años de Formación Musical	Do you play any musical instument? / Tocas algún instrumento musical?		Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?	Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?	Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?	Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?	Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?	Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?	Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?	Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?	Comments / Comentarios	Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?	Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?
8/29/2012 1:14:38	34	Female / Mujer	From 0 to 3 years / De 0 a 3 años	Yes, I play a bit / Si, toco un poco		None of them sound "human". / Ninguna suena "humana".	Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_6 sounds more human than SAMPLE_5 / El SAMPLE_6 suena mas humano que el SAMPLE_5.	SAMPLE_8 sounds more human than SAMPLE_7 / El SAMPLE_8 suena mas humano que el SAMPLE_7.	SAMPLE_9 sounds more human than SAMPLE_10 / El SAMPLE_9 suena mas humano que el SAMPLE_10.	SAMPLE_12 sounds more human than SAMPLE_11 / El SAMPLE_12 suena mas humano que el SAMPLE_11.	SAMPLE_14 sounds more human than SAMPLE_13 / El SAMPLE_14 suena mas humano que el SAMPLE_13.	Both sounds equally human / Ambos suenan igualemente humanos	Que interesante! Good luck with your research.	None of them sound human / Ninguno suena human	None of them sound human / Ninguno suena human
8/29/2012 1:38:39	28	Male / Hombre	From 5 to 8 years / De 5 a 8 años	Yes, I´ve taken lessons / Si, he tomado lecciones		SAMPLE_1 sounds more human than SAMPLE_2 / El SAMPLE_1 suena mas humano que el SAMPLE_2.	SAMPLE_4 sounds more human than SAMPLE_3 / El SAMPLE_4 suena mas humano que el SAMPLE_3.	SAMPLE_5 sounds more human than SAMPLE_6 / El SAMPLE_5 suena mas humano que el SAMPLE_6.	SAMPLE_7 sounds more human than SAMPLE_8 / El SAMPLE_7 suena mas humano que el SAMPLE_8.	None of them sound human / Ninguno suena human	SAMPLE_12 sounds more human than SAMPLE_11 / El SAMPLE_12 suena mas humano que el SAMPLE_11.	Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_19 sounds more human than SAMPLE_20 / El SAMPLE_19 suena mas humano que el SAMPLE_20.		SAMPLE_15 sounds more human than SAMPLE_16 / El SAMPLE_15 suena mas humano que el SAMPLE_16.	SAMPLE_18 sounds more human than SAMPLE_17 / El SAMPLE_18 suena mas humano que el SAMPLE_17.
8/29/2012 9:18:44	27	Female / Mujer	More than 8 years / Más de 8 años	Yes, I have formal education / Si, tengo formación profesional		None of them sound "human". / Ninguna suena "humana".	None of them sound human / Ninguno suena humano	SAMPLE_5 sounds more human than SAMPLE_6 / El SAMPLE_5 suena mas humano que el SAMPLE_6.	SAMPLE_8 sounds more human than SAMPLE_7 / El SAMPLE_8 suena mas humano que el SAMPLE_7.	SAMPLE_9 sounds more human than SAMPLE_10 / El SAMPLE_9 suena mas humano que el SAMPLE_10.	None of them sound human / Ninguno suena human	SAMPLE_13 sounds more human than SAMPLE_14 / El SAMPLE_13 suena mas humano que el SAMPLE_14.	SAMPLE_20 sounds more human than SAMPLE_19 / El SAMPLE_20 suena mas humano que el SAMPLE_19.		Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_17 sounds more human than SAMPLE_18 / El SAMPLE_17 suena mas humano que el SAMPLE_18.
8/29/2012 9:34:09	26	Male / Hombre	From 5 to 8 years / De 5 a 8 años	Yes, I have formal education / Si, tengo formación profesional		SAMPLE_2 sounds more human than SAMPLE_1 / El SAMPLE_2 suena mas humano que el SAMPLE_1.	SAMPLE_4 sounds more human than SAMPLE_3 / El SAMPLE_4 suena mas humano que el SAMPLE_3.	SAMPLE_5 sounds more human than SAMPLE_6 / El SAMPLE_5 suena mas humano que el SAMPLE_6.	SAMPLE_8 sounds more human than SAMPLE_7 / El SAMPLE_8 suena mas humano que el SAMPLE_7.	SAMPLE_10 sounds more human than SAMPLE_9 / El SAMPLE_10 suena mas humano que el SAMPLE_9.	SAMPLE_12 sounds more human than SAMPLE_11 / El SAMPLE_12 suena mas humano que el SAMPLE_11.	SAMPLE_13 sounds more human than SAMPLE_14 / El SAMPLE_13 suena mas humano que el SAMPLE_14.	SAMPLE_20 sounds more human than SAMPLE_19 / El SAMPLE_20 suena mas humano que el SAMPLE_19.		SAMPLE_15 sounds more human than SAMPLE_16 / El SAMPLE_15 suena mas humano que el SAMPLE_16.	SAMPLE_17 sounds more human than SAMPLE_18 / El SAMPLE_17 suena mas humano que el SAMPLE_18.
8/29/2012 10:37:39	36	Male / Hombre	More than 8 years / Más de 8 años	Yes, I have formal education / Si, tengo formación profesional		SAMPLE_2 sounds more human than SAMPLE_1 / El SAMPLE_2 suena mas humano que el SAMPLE_1.	SAMPLE_3 sounds more human than SAMPLE_4 / El SAMPLE_3 suena mas humano que el SAMPLE_4.	SAMPLE_6 sounds more human than SAMPLE_5 / El SAMPLE_6 suena mas humano que el SAMPLE_5.	SAMPLE_8 sounds more human than SAMPLE_7 / El SAMPLE_8 suena mas humano que el SAMPLE_7.	SAMPLE_10 sounds more human than SAMPLE_9 / El SAMPLE_10 suena mas humano que el SAMPLE_9.	SAMPLE_12 sounds more human than SAMPLE_11 / El SAMPLE_12 suena mas humano que el SAMPLE_11.	SAMPLE_13 sounds more human than SAMPLE_14 / El SAMPLE_13 suena mas humano que el SAMPLE_14.	SAMPLE_19 sounds more human than SAMPLE_20 / El SAMPLE_19 suena mas humano que el SAMPLE_20.	It's always the grace notes that makes the computer obvious! ;-) Good luck with your research...	SAMPLE_15 sounds more human than SAMPLE_16 / El SAMPLE_15 suena mas humano que el SAMPLE_16.	SAMPLE_17 sounds more human than SAMPLE_18 / El SAMPLE_17 suena mas humano que el SAMPLE_18.
8/29/2012 17:45:13	39	Female / Mujer	0 years / 0 años	No		Both sounds equally "human". / Ambas suenan igualmente "humanas"	SAMPLE_3 sounds more human than SAMPLE_4 / El SAMPLE_3 suena mas humano que el SAMPLE_4.	SAMPLE_6 sounds more human than SAMPLE_5 / El SAMPLE_6 suena mas humano que el SAMPLE_5.	SAMPLE_8 sounds more human than SAMPLE_7 / El SAMPLE_8 suena mas humano que el SAMPLE_7.	Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_12 sounds more human than SAMPLE_11 / El SAMPLE_12 suena mas humano que el SAMPLE_11.	SAMPLE_13 sounds more human than SAMPLE_14 / El SAMPLE_13 suena mas humano que el SAMPLE_14.	SAMPLE_19 sounds more human than SAMPLE_20 / El SAMPLE_19 suena mas humano que el SAMPLE_20.		SAMPLE_16 sounds more human than SAMPLE_15 / El SAMPLE_16 suena mas humano que el SAMPLE_15.	SAMPLE_17 sounds more human than SAMPLE_18 / El SAMPLE_17 suena mas humano que el SAMPLE_18.
8/29/2012 18:23:22	22	Male / Hombre	From 5 to 8 years / De 5 a 8 años	Yes, I have formal education / Si, tengo formación profesional		SAMPLE_1 sounds more human than SAMPLE_2 / El SAMPLE_1 suena mas humano que el SAMPLE_2.	Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_6 sounds more human than SAMPLE_5 / El SAMPLE_6 suena mas humano que el SAMPLE_5.	Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_10 sounds more human than SAMPLE_9 / El SAMPLE_10 suena mas humano que el SAMPLE_9.	Both sounds equally human / Ambos suenan igualemente humanos	None of them sound human / Ninguno suena human	SAMPLE_19 sounds more human than SAMPLE_20 / El SAMPLE_19 suena mas humano que el SAMPLE_20.		None of them sound human / Ninguno suena human	SAMPLE_17 sounds more human than SAMPLE_18 / El SAMPLE_17 suena mas humano que el SAMPLE_18.
8/29/2012 20:55:17	28	Male / Hombre	More than 8 years / Más de 8 años	Yes, I have formal education / Si, tengo formación profesional		SAMPLE_2 sounds more human than SAMPLE_1 / El SAMPLE_2 suena mas humano que el SAMPLE_1.	SAMPLE_3 sounds more human than SAMPLE_4 / El SAMPLE_3 suena mas humano que el SAMPLE_4.	SAMPLE_6 sounds more human than SAMPLE_5 / El SAMPLE_6 suena mas humano que el SAMPLE_5.	SAMPLE_8 sounds more human than SAMPLE_7 / El SAMPLE_8 suena mas humano que el SAMPLE_7.	Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_19 sounds more human than SAMPLE_20 / El SAMPLE_19 suena mas humano que el SAMPLE_20.		SAMPLE_15 sounds more human than SAMPLE_16 / El SAMPLE_15 suena mas humano que el SAMPLE_16.	None of them sound human / Ninguno suena human
8/30/2012 0:15:19	34	Male / Hombre	More than 8 years / Más de 8 años	Yes, I have formal education / Si, tengo formación profesional		None of them sound "human". / Ninguna suena "humana".	SAMPLE_3 sounds more human than SAMPLE_4 / El SAMPLE_3 suena mas humano que el SAMPLE_4.	SAMPLE_6 sounds more human than SAMPLE_5 / El SAMPLE_6 suena mas humano que el SAMPLE_5.	None of them sound human / Ninguno suena human	SAMPLE_10 sounds more human than SAMPLE_9 / El SAMPLE_10 suena mas humano que el SAMPLE_9.	SAMPLE_12 sounds more human than SAMPLE_11 / El SAMPLE_12 suena mas humano que el SAMPLE_11.	SAMPLE_13 sounds more human than SAMPLE_14 / El SAMPLE_13 suena mas humano que el SAMPLE_14.	SAMPLE_19 sounds more human than SAMPLE_20 / El SAMPLE_19 suena mas humano que el SAMPLE_20.	buena!!!!!! abrazo!!	SAMPLE_15 sounds more human than SAMPLE_16 / El SAMPLE_15 suena mas humano que el SAMPLE_16.	None of them sound human / Ninguno suena human
8/30/2012 0:20:40	37	Male / Hombre	0 years / 0 años	No		Both sounds equally "human". / Ambas suenan igualmente "humanas"	SAMPLE_3 sounds more human than SAMPLE_4 / El SAMPLE_3 suena mas humano que el SAMPLE_4.	SAMPLE_5 sounds more human than SAMPLE_6 / El SAMPLE_5 suena mas humano que el SAMPLE_6.	SAMPLE_8 sounds more human than SAMPLE_7 / El SAMPLE_8 suena mas humano que el SAMPLE_7.	SAMPLE_10 sounds more human than SAMPLE_9 / El SAMPLE_10 suena mas humano que el SAMPLE_9.	SAMPLE_12 sounds more human than SAMPLE_11 / El SAMPLE_12 suena mas humano que el SAMPLE_11.	Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos	Checho buena suerte con el maestrado, mis oidos incultos casi no notaban diferencia entre las opciones	SAMPLE_15 sounds more human than SAMPLE_16 / El SAMPLE_15 suena mas humano que el SAMPLE_16.	Both sounds equally human / Ambos suenan igualemente humanos
8/30/2012 18:49:34	35	Female / Mujer	0 years / 0 años	No		SAMPLE_1 sounds more human than SAMPLE_2 / El SAMPLE_1 suena mas humano que el SAMPLE_2.	SAMPLE_3 sounds more human than SAMPLE_4 / El SAMPLE_3 suena mas humano que el SAMPLE_4.	Both sounds equally human / Ambos suenan igualemente humanos	Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_9 sounds more human than SAMPLE_10 / El SAMPLE_9 suena mas humano que el SAMPLE_10.	SAMPLE_11 sounds more human than SAMPLE_12 / El SAMPLE_11 suena mas humano que el SAMPLE_12.	None of them sound human / Ninguno suena human	SAMPLE_20 sounds more human than SAMPLE_19 / El SAMPLE_20 suena mas humano que el SAMPLE_19.	Creo que los samples deberian ser unos segundos mas largos, si no hay un poquito mas de melodia, para la gente que no sabemos nada de musica no son faciles de escuchar. La encuesta es un poquito tendenciosa porque no son visibles las 4 opciones de respuesta, en las primeras respuestas uno puede despistarse y no darse cuenta de que hay la opcion de decir que suenan igual de inhumanos o humanos. La solucion seria indicar al principio de la encuesta las opciones que hay. Gracias.	SAMPLE_16 sounds more human than SAMPLE_15 / El SAMPLE_16 suena mas humano que el SAMPLE_15.	SAMPLE_17 sounds more human than SAMPLE_18 / El SAMPLE_17 suena mas humano que el SAMPLE_18.
8/30/2012 19:13:56	31	Female / Mujer	0 years / 0 años	No		SAMPLE_2 sounds more human than SAMPLE_1 / El SAMPLE_2 suena mas humano que el SAMPLE_1.	SAMPLE_3 sounds more human than SAMPLE_4 / El SAMPLE_3 suena mas humano que el SAMPLE_4.	SAMPLE_6 sounds more human than SAMPLE_5 / El SAMPLE_6 suena mas humano que el SAMPLE_5.	SAMPLE_8 sounds more human than SAMPLE_7 / El SAMPLE_8 suena mas humano que el SAMPLE_7.	Both sounds equally human / Ambos suenan igualemente humanos	SAMPLE_12 sounds more human than SAMPLE_11 / El SAMPLE_12 suena mas humano que el SAMPLE_11.	SAMPLE_13 sounds more human than SAMPLE_14 / El SAMPLE_13 suena mas humano que el SAMPLE_14.	Both sounds equally human / Ambos suenan igualemente humanos		SAMPLE_16 sounds more human than SAMPLE_15 / El SAMPLE_16 suena mas humano que el SAMPLE_15.	Both sounds equally human / Ambos suenan igualemente humanos

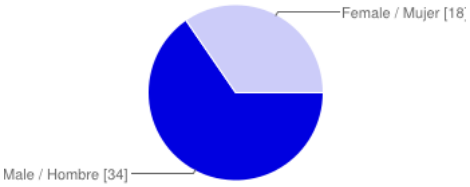
54 responses

Summary See complete responses

How old are you? / Cuantos años tienes?

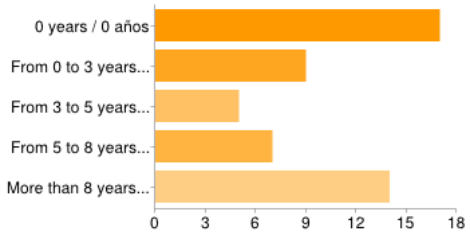
37 17 37 25 30 49 34 24 35 28 36 34 44 40 30 37 31 25 34 28 24 35 34 40 27 38 38 33 54 31 30 30 38 35 34 37

Gender / Sexo



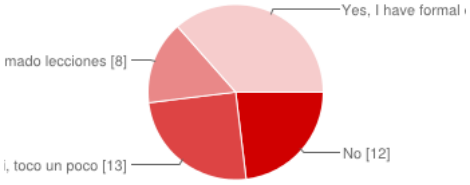
Male / Hombre	**34**	63%
Female / Mujer	**18**	33%

Years of Musical Training / Años de Formación Musical



0 years / 0 años	**17**	31%
From 0 to 3 years / De 0 a 3 años	**9**	17%
From 3 to 5 years / De 3 a 5 años	**5**	9%
From 5 to 8 years / De 5 a 8 años	**7**	13%
More than 8 years / Más de 8 años	**14**	26%

Do you play any musical instument? / Tocas algún instrumento musical?



No	**12**	22%
Yes, I play a bit / Si, toco un poco	**13**	24%
Yes, I´ve taken lessons / Si, he tomado lecciones	**8**	15%
Yes, I have formal education / Si, tengo formación profesional	**19**	35%

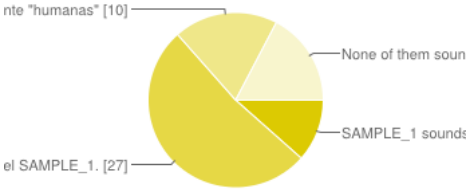
Music Samples / Muestras musicales

In the following URL link you will find some links to some sound samples, labeled as SAMPLE_1, SAMPLE_2 etc.(Please don´t listen no any sample yet!). Each sample is betwen 15 to 30 seconds long. Please read carefully the questions and listen to the samples only when indicated. Here is the link: http://sergiomusicresearch.wordpress.com Please copy and paste the url link in a new browser tab or window. When you have the page open please click "continue" _____ En el siguiente link URL encontrará algunos links a ciertos archivos de audio titulados SAMPLE_1, SAMPLE_2, etc. (Por favor no escuche ninguno aun!). Cada muestra musical tiene entre 15 y 30 segundos de duración. Lea atentamente las siguientes preguntas y escuche las muestras solamente cuando se le indique. Aquí esta el link: http://sergiomusicresearch.wordpress.com Por favor cópie y pegue la dirección url en una nueva pestaña o ventana de su explorador. Cuando tenga esta página abierta oprima "continue"

Question 1 / Pregunta 1

(Pag. 1/10) Please listen to SAMPLE_1 and to SAMPLE_2. _____ Por favor escuche el SAMPLE_1 y el SAMPLE_2.

Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?

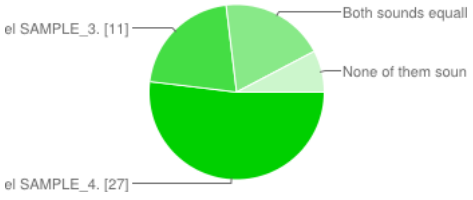


SAMPLE_1 sounds more human than SAMPLE_2 / El SAMPLE_1 suena mas humano que el SAMPLE_2.	**6**
SAMPLE_2 sounds more human than SAMPLE_1 / El SAMPLE_2 suena mas humano que el SAMPLE_1.	**27**
Both sounds equally "human". / Ambas suenan igualmente "humanas"	**10**
None of them sound "human". / Ninguna suena "humana".	**9**

Question 2 / Pregunta 2

(Pag. 2/10) Please listen to SAMPLE_3 and to SAMPLE_4. _____ Por favor escuche el SAMPLE_3 y el SAMPLE_4.

Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?

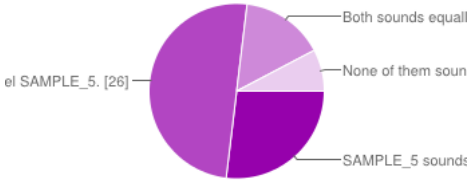


SAMPLE_3 sounds more human than SAMPLE_4 / El SAMPLE_3 suena mas humano que el SAMPLE_4.	27
SAMPLE_4 sounds more human than SAMPLE_3 / El SAMPLE_4 suena mas humano que el SAMPLE_3.	11
Both sounds equally human / Ambos suenan igualemente humanos	10
None of them sound human / Ninguno suena humano	4

Question 3 / Pregunta 3

(Pag 3/10) Please listen to SAMPLE_5 and to SAMPLE_6. _____ Por favor escuche el SAMPLE_5 y el SAMPLE_5.

Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?

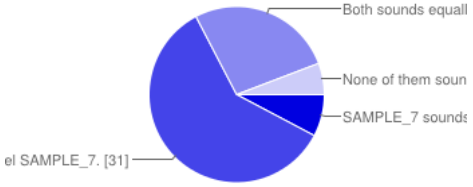


SAMPLE_5 sounds more human than SAMPLE_6 / El SAMPLE_5 suena mas humano que el SAMPLE_6.	14
SAMPLE_6 sounds more human than SAMPLE_5 / El SAMPLE_6 suena mas humano que el SAMPLE_5.	26
Both sounds equally human / Ambos suenan igualemente humanos	8
None of them sound human / Ninguno suena human	4

Question 4 / Pregunta 4

(Pag 4/10) Please listen to SAMPLE_7 and to SAMPLE_8. _____ Por favor escuche el SAMPLE_7 y el SAMPLE_8.

Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?

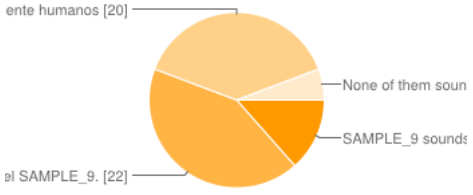


SAMPLE_7 sounds more human than SAMPLE_8 / El SAMPLE_7 suena mas humano que el SAMPLE_8.	4
SAMPLE_8 sounds more human than SAMPLE_7 / El SAMPLE_8 suena mas humano que el SAMPLE_7.	31
Both sounds equally human / Ambos suenan igualemente humanos	14
None of them sound human / Ninguno suena human	3

Question 5 / Pregunta 5

(Pag 5/10) Please listen to SAMPLE_9 and to SAMPLE_10. _____ Por favor escuche el SAMPLE_9 y el SAMPLE_10.

Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?

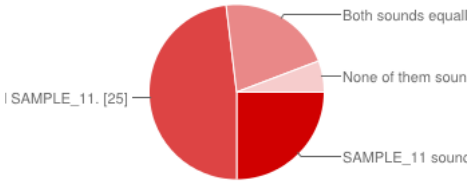


SAMPLE_9 sounds more human than SAMPLE_10 / El SAMPLE_9 suena mas humano que el SAMPLE_10.	
SAMPLE_10 sounds more human than SAMPLE_9 / El SAMPLE_10 suena mas humano que el SAMPLE_9.	2
Both sounds equally human / Ambos suenan igualemente humanos	2
None of them sound human / Ninguno suena human	

Question 6 / Pregunta 6

(Pag 6/10) Please listen to SAMPLE_11 and to SAMPLE_12. _____ Por favor escuche el SAMPLE_11 y el SAMPLE_12.

Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?

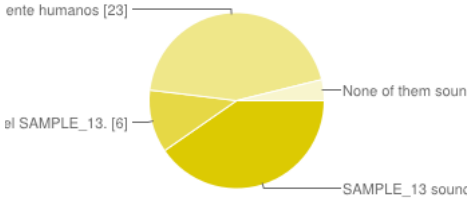


SAMPLE_11 sounds more human than SAMPLE_12 / El SAMPLE_11 suena mas humano que el SAMPLE_12.
SAMPLE_12 sounds more human than SAMPLE_11 / El SAMPLE_12 suena mas humano que el SAMPLE_11.
Both sounds equally human / Ambos suenan igualemente humanos
None of them sound human / Ninguno suena human

Question 7 / Pregunta 7

(Pag 7/10) Please listen to SAMPLE_13 and to SAMPLE_14. _____ Por favor escuche el SAMPLE_13 y el SAMPLE_14.

Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?

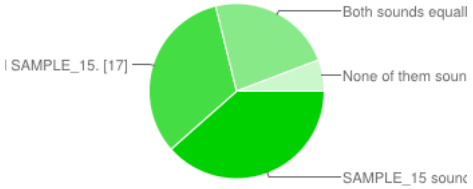


SAMPLE_13 sounds more human than SAMPLE_14 / El SAMPLE_13 suena mas humano que el SAMPLE_14.
SAMPLE_14 sounds more human than SAMPLE_13 / El SAMPLE_14 suena mas humano que el SAMPLE_13.
Both sounds equally human / Ambos suenan igualemente humanos
None of them sound human / Ninguno suena human

Question 8 / Pregunta 8

(Pag 8/10) Please listen to SAMPLE_15 and to SAMPLE_16. _____ Por favor escuche el SAMPLE_15 y el SAMPLE_16.

Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?

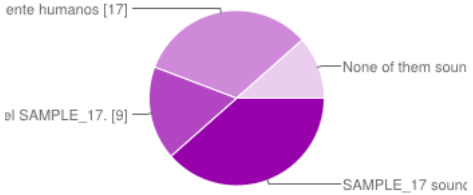


SAMPLE_15 sounds more human than SAMPLE_16 / El SAMPLE_15 suena mas humano que el SAMPLE_16.
SAMPLE_16 sounds more human than SAMPLE_15 / El SAMPLE_16 suena mas humano que el SAMPLE_15.
Both sounds equally human / Ambos suenan igualemente humanos
None of them sound human / Ninguno suena human

Question 9 / Pregunta 9

(Pag 9/10) Please listen to SAMPLE_17 and to SAMPLE_18. _____ Por favor escuche el SAMPLE_17 y el SAMPLE_18.

Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?

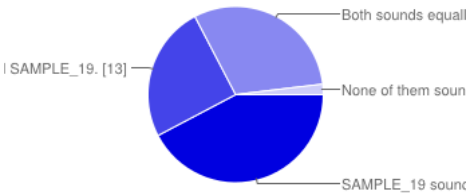


SAMPLE_17 sounds more human than SAMPLE_18 / El SAMPLE_17 suena mas humano que el SAMPLE_18.
SAMPLE_18 sounds more human than SAMPLE_17 / El SAMPLE_18 suena mas humano que el SAMPLE_17.
Both sounds equally human / Ambos suenan igualemente humanos
None of them sound human / Ninguno suena human

Question 10 / Pregunta 10

(Pag 10/10) Please listen to SAMPLE_19 and to SAMPLE_20. _____ Por favor escuche el SAMPLE_19 y el SAMPLE_20.

Which one sounds closer to a "human" performance? / Cual de estas suena mas cercana a una interpretación "humana"?



SAMPLE_19 sounds more human than SAMPLE_20 / El SAMPLE_19 suena mas humano que el SAMPLE_20.

SAMPLE_20 sounds more human than SAMPLE_19 / El SAMPLE_20 suena mas humano que el SAMPLE_19.

Both sounds equally human / Ambos suenan igualemente humanos

None of them sound human / Ninguno suena human

Comments / Comentarios

We are interested in your comments and impressions. However this question is not mandatory, if you don't have any, just leave it in blank. _____ Estamos interesados en sus comentarios e impresiones. No obstante esta pregunta no es obligatoria, si no tiene ninguno, simplemente deje la respuesta en blanco.

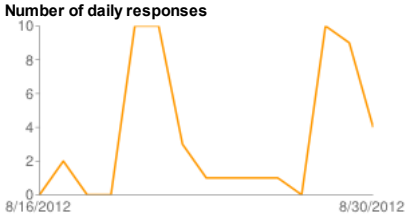
Comments / Comentarios

By Sergio, test data. Muy bien !! last test checking (erase data) Me ha costado mucho diferenciar cual es la humana i cual la artificial en practicamente todos los casos. note length or sustain in some cases was a factor. Also speed of grace notes. Algunos samples dan la impresion de que son mitad tocados, mitad sampleados. Hay uno, por ejemplo, que es el mismo motivo tocado dos veces. La segunda suena distinta de la primera - aparece por primera vez el motivo, despues hay un ruido, cambia el sonido y luego repiten el motivo. Overall I experienced that all of the sounds could have been played by h ...

That's all / Eso es todo

Thanks for your collaboration, Don't forget to submit your answers by clicking on "submit" button. _____ Gracias por su colaboración. No olvide enviar sus respuestas pulsando el botón "submit / enviar".

Number of daily responses



Number of responses without dates: 2