

MUMT621 – Final Project – Hidden Markov Model for Symbolic Key Detection

Néstor Nápoles López

Abstract

In this paper, we introduce a model for the task of key detection based on a Hidden Markov Model (HMM). Each of the 24 major and minor keys available in the tonal system represent a hidden state in this model. Additionally, the 12 pitch-classes represent the possible observation symbols generated by a hidden state. The model consists of 24 hidden states and 12 observation symbols, the latter are shared for all hidden states.

At first, the model is initialized with the same probability of starting in any of the 24 keys. In the following transitions, the probability distribution is computed from a geometric model of key-distance where each *layer* of neighbor keys is exponentially less-likely. In the case of the observation symbols, the probability distributions are set from well-known key-profiles used for key detection.

The algorithm has been tested in a total of 124 MIDI files that comprehend music from J. S. Bach, F. Chopin, A. Scriabin and D. Shostakovich.

Using the best combination of the key-profiles that were used for evaluation, the algorithm is able to guess the main key of 103 of the MIDI files in the dataset, which account roughly for 83% of the music pieces. Other attempts have been made of detecting the key of a musical fragment using HMMs, however, this is the first approach in the domain of symbolic music information retrieval, as far as we are concerned.

1 Introduction

Finding the key of a piece of Western art music has been in the interest of the Music Information Retrieval (MIR) community for several years already. Since they were introduced by Krumhansl and Kessler (1982), the design or acquisition of key-profiles have been the preferred methodology to solve this problem. In this project, we combine the use of key-profiles with the capabilities of Hidden Markov Models (HMMs) to model the time-varying aspects of music to find the key of a piece of music by considering the pitch-class of every note in the piece as the observation symbol of an HMM.

2 Design of the Hidden Markov Model

As mentioned, the proposed HMM consists of 24 hidden states and 12 observation symbols. Figure 1 shows a basic overview of the HMM.

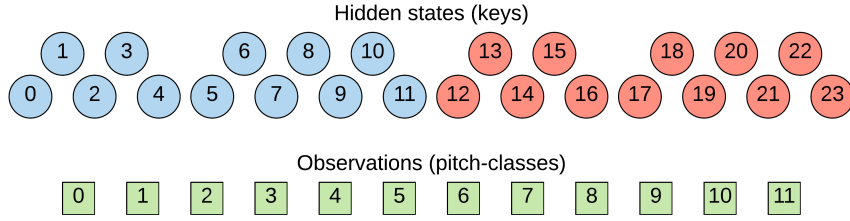


Figure 1: States and observation symbols of the proposed HMM. The blue states represent major keys, the red states represent minor keys, the green boxes represent the twelve pitch-classes.

2.1 Keys as hidden states

The hidden states correspond to 24 different major and minor keys (i.e., no distinction between enharmonics), each of these keys is represented as a hidden state in the model. All of the keys may transition to any of the other 23 keys, however, the probability of transitioning to neighbor keys in the circle of fifths is preferred over distant keys.

2.2 Pitch-classes as observation symbols

All major and minor keys are able to *emit* any of the twelve pitch-classes before they transition to a new state, however, by acquiring the probability distributions from common key-profiles used for the task of key detection, the emission of *diatonic* tones of the key are preferred over accidental tones.

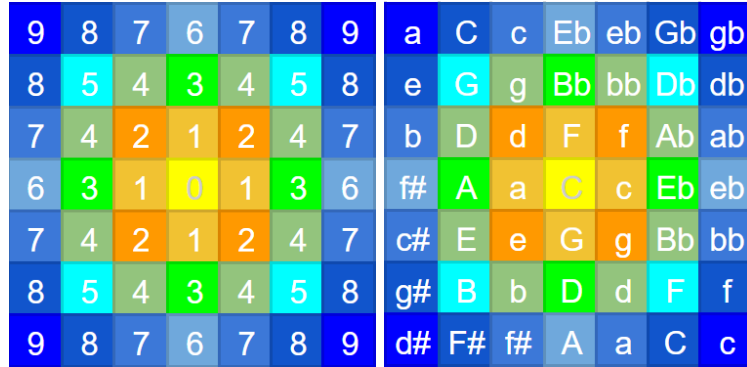


Figure 2: A generic overview of the geometric model of key distance, starting from any tonic (left). The key distances of C Major (right)

3 Parameters of the model

3.1 Initial probabilities

The initial probabilities of the model are the same for each key, $p(state) = \frac{1}{24}$.

3.2 State transitions

The probability distributions for state transitions that happen after the initialization have been taken from a geometric model of key distance. The probability of a transition to another key in the next group of neighbors decreases exponentially. The geometric model of key distance can be observed in Figure 2

Using this model, we can observe, for example, a list of the groups of neighbor keys for C Major, in descending order:

Group	Keys
0	C
1	G F a c
2	d e f g
3	D Eb A Bb
4	E Ab bb b
5	Db B
6	eb f#
7	c# ab
8	F#

This structure of nine groups of neighbours repeats for all major and minor keys. We can use these groups to compute the transition probability for any key according to the following formula:

$$p(s) = \alpha^{(G-1)-s_g}$$

where G is the number of groups of neighbor keys to the current tonic, according to the geometric model (i.e., $G = 9$), s_g is the group to which the key s belongs, α is the probability ratio between a key of one group and a key from a contiguous group.

Figure 3 shows a plot of the probability distribution for the transition to the next key, if the current key is C Major.

3.3 Emission probabilities

The emission probability distributions have been taken from the common key-profiles used by other key detection algorithms. Particularly, we considered the same five pairs of key-profiles used in the comparison by Albrecht and Shanahan (2013).

4 Dataset

The model has been evaluated using sets of short musical compositions in all major and minor keys. Each set follows the format of the *Well-Tempered Clavier* by Johann Sebastian Bach. The sets are: The Well-Tempered Clavier, Volume I, by Johann Sebastian Bach (24 MIDI files); The Well-Tempered Clavier (Part II), by Johann Sebastian Bach (24 MIDI

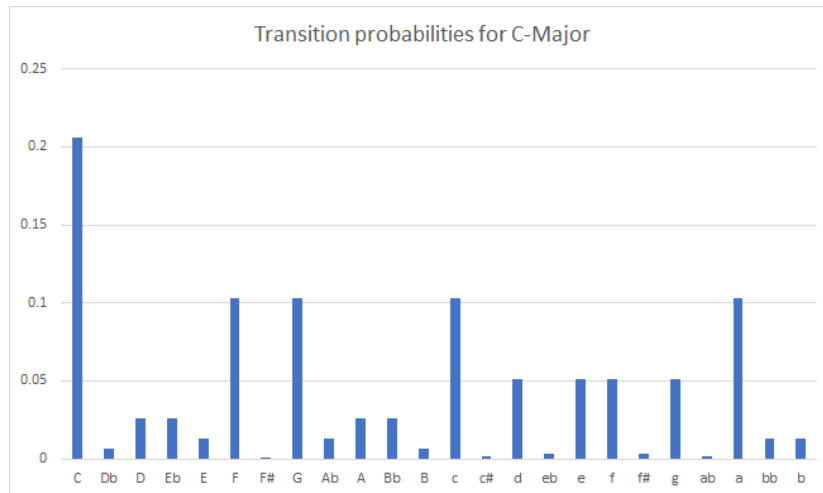


Figure 3: Probability distribution for the next state if the current state is C Major

files); Preludes Op.28, by Frédéric Chopin (24 midi files); 4 of the 24 preludes from Op.11, by Alexander Scriabin (4 MIDI files); Preludes and Fugues Op.87, by Dmitri Shostakovich (48 MIDI files). In total, 124 MIDI files were used for testing the algorithm.

Even if these MIDI files have not been explicitly annotated with their keys and modulations, the main key of each of the selected pieces has been clearly established by the composer in the original scores, meaning an evaluation process of the algorithm is possible at least against the main key of the musical piece.

5 Results

The best performance of the algorithm used a ratio of $\alpha = 10$, the key-profiles from Temperley for major keys, and the key-profiles from Sapp for minor keys. With this configuration, the algorithm guessed correctly 103 out of 124 keys in the MIDI files from the dataset.

The worst performance of the algorithm used a ratio of $\alpha = 2$, the key-profiles from Krumhansl and Kessler for major keys, and the key-profiles from Aarden and Essen for the minor keys. With this configuration, the algorithm guessed 49 out of 124 keys in the MIDI files from the dataset.

A full report of the evaluation is given as an appendix to this paper.

References

- Aarden, Bret J. 2003. “Dynamic Melodic Expectancy.” PhD diss., The Ohio State University.
- Albrecht, Joshua, and Daniel Shanahan. 2013. “The use of large corpora to train a new type of key-finding algorithm: An improved treatment of the minor mode.” *Music Perception: An Interdisciplinary Journal* 31 (1): 59–67.

- Bellmann, Héctor. 2006. "About the Determination of Key of a Musical Excerpt." In *Computer Music Modeling and Retrieval*, edited by Richard Kronland-Martinet, Thierry Voinier, and Sølvi Ystad, 76–91. Berlin, Heidelberg: Springer Berlin Heidelberg. ISBN: 978-3-540-34028-7.
- Cannam, Chris, Emmanouil Benetos, Matthias Mauch, Matthew E.P. Davies, Simon Dixon, Christian Landone, Katy Noland, and Dan Stowell. 2017. "MIREX 2017: VAMP Plugins from the Centre for Digital Music." In *Submission at the Music Information Retrieval Evaluation eXchange (MIREX 2017)*, CN1. Champaign, IL: University of Illinois at Urbana-Champaign.
- Chai, Wei, and Barry Vercoe. 2005. "Detection of Key Change in Classical Piano Music." In *Proceedings of the 6th International Conference on Music Information Retrieval*, edited by Joshua D. Reiss and Geraint A. Wiggins, 468–473. London, UK: University of London. ISBN: 0-9551179-0-9.
- Chew, Elaine. 2000. "Towards a Mathematical Model of Tonality." PhD diss., Massachusetts Institute of Technology.
- Chuan, Ching-Hua, and Elaine Chew. 2005. "Polyphonic Audio Key Finding Using the Spiral Array CEG Algorithm." In *2005 IEEE International Conference on Multimedia and Expo*, 21–24. Amsterdam, Netherlands: IEEE. ISBN: 0-7803-9331-7. doi:10.1109/ICME.2005.1521350.
- Gómez, Emilia. 2006. "Tonal Description of Music Audio Signals." PhD diss., Universitat Pompeu Fabra.
- Huron, David, and Richard Parncutt. 1993. "An Improved Model of Tonality Perception Incorporating Pitch Salience and Echoic Memory." *Psychomusicology: A Journal of Research in Music Cognition* 12 (2): 154.
- Hyer, Brian. *Tonality*. <http://www.oxfordmusiconline.com/view/10.1093/gmo/9781561592630.001.0001/omo-9781561592630-e-0000028102>.
- İzmirli, Özgür. 2005. "An Algorithm for Audio Key Finding." In *Submission at the Music Information Retrieval Evaluation eXchange (MIREX 2017)*, Rank 1. Champaign, IL: University of Illinois at Urbana-Champaign.
- Krumhansl, Carol L., and Edward J. Kessler. 1982. "Tracing the Dynamic Changes in Perceived Tonal Organization in a Spatial Representation of Musical Keys." *Psychological Review* 89 (4): 334–368.
- Longuet-Higgins, Hugh Christopher. 1976. "Perception of Melodies." *Nature* 263 (5579): 646–653.

- Noland, Katy, and Mark Sandler. 2007. "Signal Processing Parameters for Tonality Estimation." In *Audio Engineering Society Convention 122*. Vienna, Austria: Curran Associates, Inc. ISBN: 9-7816042314-0-3. <http://www.aes.org/e-lib/browse.cfm?elib=14140>.
- Papadopoulos, Hélène, and Geoffroy Peeters. 2009. "Local Key Estimation Based on Harmonic and Metric Structures." In *12th Int. Conference on Digital Audio Effects (DAFx-09)*, edited by Fabio Antonacci, 408–415. Como, Italy: Politecnico de Milano. <https://hal.archives-ouvertes.fr/hal-00511452>.
- Schedl, Markus, Emilia Gómez, and Julián Urbano. 2014. "Music Information Retrieval: Recent Developments and Applications." *Foundations and Trends in Information Retrieval* 8 (2-3): 127–261. ISSN: 1554-0669. doi:10.1561/15000000042. <http://dx.doi.org/10.1561/15000000042>.
- Temperley, David. 1999. "What's Key for Key? The Krumhansl-Schmuckler Key-Finding Algorithm Reconsidered." *Music Perception: An Interdisciplinary Journal* 17 (1): 65–100. ISSN: 07307829, 15338312. <http://www.jstor.org/stable/40285812>.
- . 2002. "A Bayesian Approach to Key-Finding." In *Music and Artificial Intelligence*, edited by Christina Anagnostopoulou, Miguel Ferrand, and Alan Smaill, 195–206. Berlin, Heidelberg: Springer. ISBN: 978-3-540-45722-0.
- Thompson, William F, and Lola L Cuddy. 1992. "Perceived Key Movement in Four-Voice Harmony and Single Voices." *Music Perception: An Interdisciplinary Journal* 9 (4): 427–438.