Learning Harmonic Relationships in Digital Audio with Dirichlet-Based Hidden Markov Models

J. Ashley Burgoyne and Lawrence K. Saul

Harmonic analysis is a standard musicological tool for understanding many pieces of Western classical music and making comparisons among them. Traditionally, this analysis is done on paper scores, and most past research in machine-assisted analysis has begun with digital representations of them. Human music students are also taught to hear their musical analyses, however, in both musical recordings and performances. Our approach attempts to teach machines to do the same, beginning with a corpus of recorded Mozart symphonies.

Pitch Class Profiles

The audio files are first transformed into an ordered series of normalized pitch class profile (PCP) vectors. PCP vectors represent the total amount of spectral energy within each pitch class during a short time frame. The PCP vectors in this model are normalized to sum to unity so that they represent the proportion of spectral energy within each pitch class during the time frame.

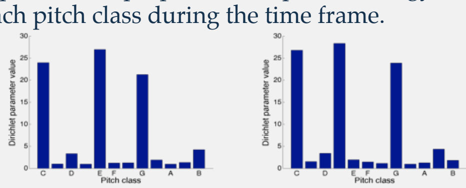


Figure 1a Initial Dirichlet parameters, rotated for C major and minor. The shapes of the distributions illustrate the mean distribution while the parameter values are inversely proportional to the variance.

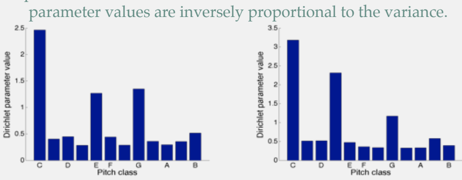


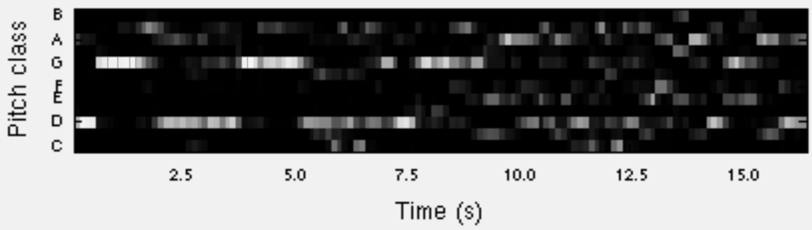
Figure 1b Trained Dirichlet parameters, rotated for C major and minor. The distribution shapes remain unmistakably major and minor triads while the variance has increased considerably.

Dirichlet Distributions

Dirichlet distributions are probability distributions over sets of discrete probability distributions. They are more attractive than Gaussian models for systems where the relations among outputs are more important than their magnitude, such as normalized PCP vectors. The formula for a Dirichlet distribution is:

$$Dir(\vec{p}, \vec{u}) \triangleq \frac{1}{Z(\vec{u})} \prod_{i=1}^{n} p_i^{u_i - 1}$$

a Reduced musical score.



b PCP distributions (first repeat only), scaled linearly in time. Note the lack of obvious triads.

		1		2	3		4	5	6		7		8		9			10				11			12			13		14
Ground truth:	D **g**		g **g**	D **g**	g **g**		g **g**	D **g**	g **g**	D **g**		E♭ **g/d**		A **d**	d **d**	g **d**	A **d**	B♭ **d**	C **d**		D **d**	g **d**	A **d**	d **d**	E **d**	A **d**	B♭ **d**	e **d**	A **d**	d **d**
First repeat:	D **g**	G **g**	g **g**	D **d**	G **d**	g **d**	g **d**	D **d**	g **d**	D **d**	G **d**	g **d**	c♯ **d**	A **d**	d **d**	g **d**	A **d**	d **d**	C **d**	g **d**	D **d**	g **d**	c♯ **d**	d **d**	E **d**	A **d**	d **d**	e **d**	A **d**	D **d**
Second repeat:	D **d**	G **d**	g **d**	D **d**	g **d**	G **d**	g **d**	D **d**	g **d**	D **d**	G **d**	g **g**	e **g**	A **g**	d **g**	g **g**	A **g**	d **g**	C **g**	g **g**		g **g**	b♭ **g**	d **g**	E **g**	A **g**	d **g**	e **g**	A **g**	D **g**

c Ground truth harmonic labels and machine analysis. The results are especially impressive considering the few well defined triads above: clearly, the harmonic model is working.

Figure 2 W. A. Mozart, Symphony No. 40 (K. 550), III, mm. 1–14

Harmonic Mixture

Relationships among tonal chords and keys are modeled as three classes:

	Major	Minor
I	none	primary
i	primary	none
♭II	primary	none
♭ii	double	secondary
II	secondary	double
ii	none	primary
♭III	primary	none
♭iii	double	secondary
III	secondary	double
iii	none	primary
IV	none	none
iv	primary	none
V	none	none
v	primary	none
♭VI	primary	none
♭vi	double	secondary
VI	secondary	double
vi	none	primary
♭VII	primary	none
♭vii	double	secondary
VII	secondary	double
vii	secondary	double

Hidden Markov Model

A hiden Markov model (HMM) forms the backbone of the system. It derives its transition matrix from the three classes of harmonic mixture described above, accepts a series of normalized PCP vectors as input, and parameterizes its observation distributions as Dirichlet distributions tied to the underlying states.

Real-Time Results

The training corpus comprised professional compact disc recordings of five Mozart symphonies in fifteen movements altogether. After processing, the expectation maximization (EM) algorithm was used to tune the Dirichlet parameters.

The results of a test on a recording of Mozart's K 550 are displayed above. Dirichlet distributions on PCP vectors proved to be an effective means for chord recognition in recorded performances of symphonic music, and their efficiency will allow for implementation in real time.

McGill

Penn