Linear Programming Boosting for Classification of Musical Genre

Tom Diethe (t.diethe@cs.ucl.ac.uk)

Department of Computer Science
University College London

Music, Brain & Cognition Workshop
NIPS 7/12/2007

Motivation

- Genre classifiaction from raw audio files is a fairly well researched area of music research

+ve's

- Valuable for audio information retrieval systems
- Chance to improve signal processing/feature extraction methods
- Tough Machine Learning problem

-ve's

- *"I have pretty serious doubts about genre classification in the first place, because of the seemingly arbitrary nature of the classes and how they are assigned."* (Dan Ellis)
- Genre classification task is to reproduce an arbitrary set of culturally assigned classes
- Genres are not neccessarily natural groupings
- Humans don't perform that well!

MIREX

- Music Information Retrieval Evaluation eXchange (MIREX) is part of International Conference on Music Information Retrieval (ISMIR)
- Classification of polyphonic musical audio into a single high-level genre per example
- Audio format: MP3, CD-quality (PCM, 16-bit, 44100 Hz), mono

Results for MIREX 2005:

Participant	Algorithm	Features	Score
Bergstra et al.	AdaBoost	Aggregated features	82.23%
Mandel & Ellis	SVM	KL-Divergence	78.81%
West	Trees,LDA	Spectral & Rhythmic	75.29%
Lidy & Rauber	SVM	Spectral & Rhythmic	75.27%
Pampalk et al.	1-NN	MFCC	75.14%
Scaringella	SVM	Texture & Rhythmic	73.11%
Ahrendt & Meng	SVM	Auto-Regression	71.55%
Burred	GMM/ML	Aggregated features	62.63%
Soares	GMM	Aggregated features	60.98%
Tzanetakis	LSVM	FFT/MFCC	60.72%

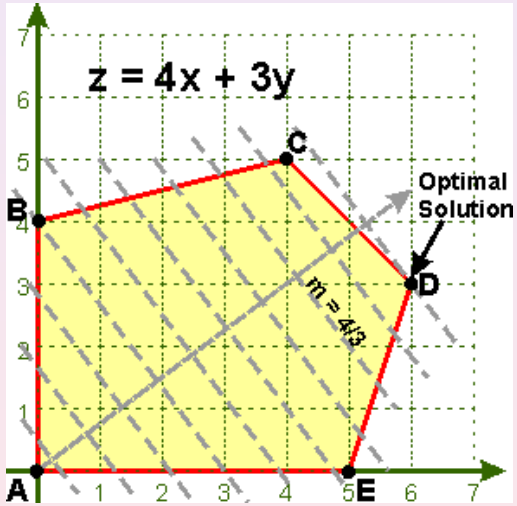
Feature Selection

- Techniques inspired by speech perception, signal processing theory, music theory etc.
- Frame length: 1024 samples (46.44ms @ 22050Hz)
- Discrete Fourier Transform, Real Cepstral Coefficients (RCEPS), Mel Frequency Cepstral Coefficients (MFCC), Zero Crossing Rate (ZCR), Spectral Rolloff, Autoregression (LPC, LPCE)
- Aggregation by by fitting individual Gaussians to each feature (diagonal covariance)
- Feature vector by concatenating 256 RCEPS, 64 MFCC, 32 LPC, 1 LPCE, 32 FFTC, 16 rolloff, and 1 ZCR.
- $\Rightarrow 402x2 = 804$ parameters for each segment.

Linear Programming Boosting

Intuition

- Taking a 1-norm of the slack variables and optimising the 1-norm of coefficients leads to a linear programme
- LPBoost converges in a finite number of iterations to a globally optimal solution
- Column generation generates only variables which have the potential to improve the objective function (i.e. -ve reduced cost)
- In the dual form the constraints are the weak learners
- Add a weak learner, check if the L.P. is solved
- If not find the weak learner that violates the constraints the most
- Repeat until the L.P. constraints are not violated
- $\rightarrow$ global optimum solution
- LPBoost iterations are typically slower than AdaBoost
- Converges much more quickly



Data

Magnatune 2004

- RWC Magnatune database used for the 2004 Audio desciption contest is still available

Anders Meng dataset d004

- 11 genres, 1100 training examples and 220 test examples
- The integrity of the data-set have been evaluated by humans (experts and non-experts) at a decision time horizon of 30seconds
- The genres are :
  Alternative, Country, Easy Listening, Electronica, Jazz, Latin, Pop&Dance, Rap/HipHop, R&B/Soul, Reggae, Rock

Results

Summary of experimental results

- The numbers in parentheses show the number of weak learners in the final solution
- Best reported performance on Meng 4 dataset: 44% (machine), 52% (human)
- AdaBoost stopping parameter selected by 5-fold c.v.
- Average test accuracy:

Algorithm	**Magnatune 6**	**Meng 2**	**Meng 4**
AdaBoost	61.3% (10000)	87.5% (10000)	43.3% (5000)
LPBoost	63.5% (585)	87.5% (401)	41.7% (452)

Conclusions & Further Work

Conclusions

- Many different approaches to genre classificaition in terms of feature selection and algorithm choice
- Boosting with an aggregrated feature set works well, but shows that we don't know what we're doing!
- LPBOOST competitive with ADABOOST, whilst giving much sparser solutions

Further Work

- More musical feature set?
- Multiclass LPBOOST