Why is adaboost good for this task ?

Their model uses simple decision stumps (if the parameter is in a certain range, it returns one, else it returns 0), each of which operates on a single feature dimension. By doing this for every parameter, and asking for each genre « is this extract from this genre or not ? », they could then select the genre which received the most votes.

Also, the parameters can be ranked by « which had the most accuracy » so they could perform feature selection in parallel with classification, which is one of the main challenges (each iteration the most confusing parameters were left out). Once only the best parameters were kept, they are attributed weight in decision, and we get a quite robust classifier (between 75 and 85% accuracy depending on dataset, results obtained in a competition)

Second, ADABOOST scales linearly with the number of training points provided. If we limit the number of learning iterations, our model has the potential to deal with very large datasets. Unlike SVM for exemple (quadratic)

Features that were pre selected

frame level parameters
zero-crossing rate
spectral centroid
spectral roll-off
spectral spread

short(~50ms)framesofdata

hyper-parameters (data engineered, details at the end)
256 RCEPS
64 MFCC
32 LPC
the lowest 32 of 512 Fourier coefficients
16 spectral rolloff
1 LPC error,
zero-crossing rate

Though it is possible to extract data from small frame and then decide by voting, West and Cox (2004), note a significant improvement in performance when they use a “memory” feature, which is the mean and variance of frame-level features over the previous second

the optimal segment length was around 3.5 seconds for them (though they had stricter restriction for computation complexity)

Trying to reach an equilibrium between number of datapoint and size of datapoint, they found 3.5 seconds to be the best size.

They said the challenge was mainly to extend this to larger databases, as they were working for a project which had to be trained under 24 hours (and in 2005, so their computation power was quite limited), we can probably find better solutions now with the advance of technology.

Fast Fourier transform coefficients (FFTCs).

Fourier analysis is used to analyze the spectral characteristics of each frame of data. In general we computed a 512-point Fourier transform F(s) of each 1024-point frame s. The 32 lowest frequency points were retained for our experiments.

Real cepstral coefficients (RCEPS).

Cepstral analysis is commonly used in speech recognition to separate vocal excitation from the effects of the vocal tract. See Gold and Morgan (2000) for an overview. RCEPS is defined as real(F#(log(|F(s)|))) where F is the Fourier transform and F# is the inverse Fourier transform.

Mel-frequency cepstral coefficients (MFCC).

These features are similar to RCEPS, except that the input x is first projected according to the Mel-scale Junqua and Haton (1996), a psycho-acoustic frequency scale on which a change of 1 unit carries the same perceptual significance, regardless of the position on the scale.

Zero-crossing rate (ZCR).

The zero-crossing rate (ZCR) is the rate of sign-changes along the signal. In a signal with a single pitched instrument, the ZCR is correlated with the dominant frequency (Kedem, 1986).

Spectralspread, spectral centroid, spectral roll off

Spectral spread and spectral centroid are measures of how power is distributed over frequency. Spectral spread is the variance of this distribution. The spectral centroid is the center of mass of this distribution and is thus positively correlated with brightness.Spectral roll off feature is the a-quantile of the total energy in |Fs|

Auto-regression coefficients (LPC).

Thek linearpredictivecoefficients(LPC)ofasignal sare the product of an autoregressive compression of the spectral envelope of a signal. These coefficients can be computed efficiently from the signal's autocorrelation by the Levinson-Durbin recursion (Makhoul, 1975).

1 feature extractor is available as a C program from the first author's website:
http://www-etud. iro.umontreal.ca/~bergstrj. Springer