Feature Extraction for ASR: Delta

Wantee Wang

2015-03-14 16:57:00 +0800

In order to use the time dynamic information of speech, One can calculate the *Deltas* and *Delta-Deltas* from the original features.

Also known as *differential* and *acceleration* coefficients, they are computed as,

$$d_t = \frac{\sum_{n=1}^{N} n(c_{t+n} - c_{t-n})}{2\sum_{n=1}^{N} n^2}$$

where d_t is a delta coefficient, from frame t computed in terms of the static coefficients c_{t-N} to c_{t+N}. A typical value for N is 2. Delta-Delta (Acceleration) coefficients are calculated in the same way, but they are calculated from the deltas, not the static coefficients.