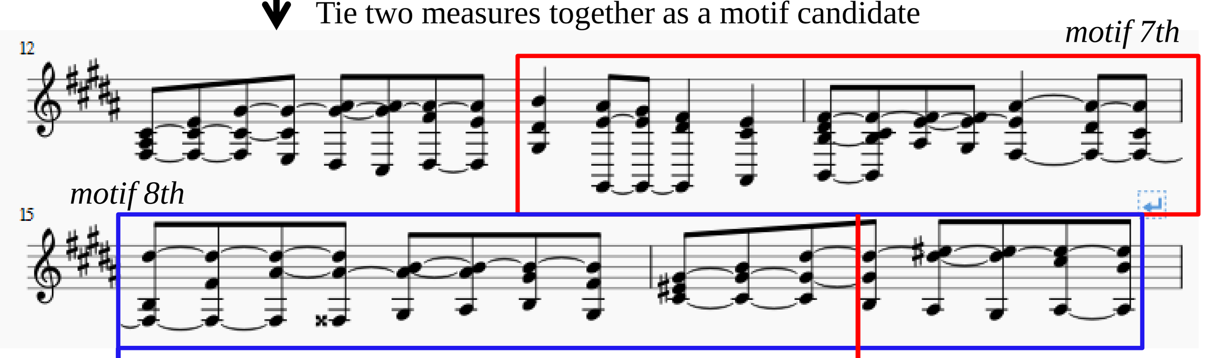
Digital Media Lab
School of Electrical Engineering, Korea University
지도교수 : 설 상 훈 교수님

Composer detection based on the harmonic motif

2014170966 Jongyeong Lee

Feature extraction

Combine treble clef and base clef
Tie two measures together as a motif candidate



candidate
[8, 11, 3, 4, 8, 10, 4, 8, 3, 6, 8, 10, 1, 4, 11, 3, 6, 6, 8, 1, 4, 6, 10, 4, 6, 8, 4, 6, 10, 3, 6, 10, 6, 10, 1]

candidate
[11, 3, 6, 3, 6, 3, 6, 10, 3, 7, 10, 8, 10, 11, 10, 11, 8, 11, 6, 8, 11, 1, 5, 8, 8, 11, 1, 1, 3, 8, 8, 11, 3, 3, 5, 10, 3, 5, 8, 10, 1, 5, 5, 10, 11]

$$\tau \sum m_i[n]m_i[n] > \sum m_i[n]m_j'[n]$$

m'[n] means distinct harmonic motif, and m[n] means a motif candidate. And τ is threshold value which is a variable.

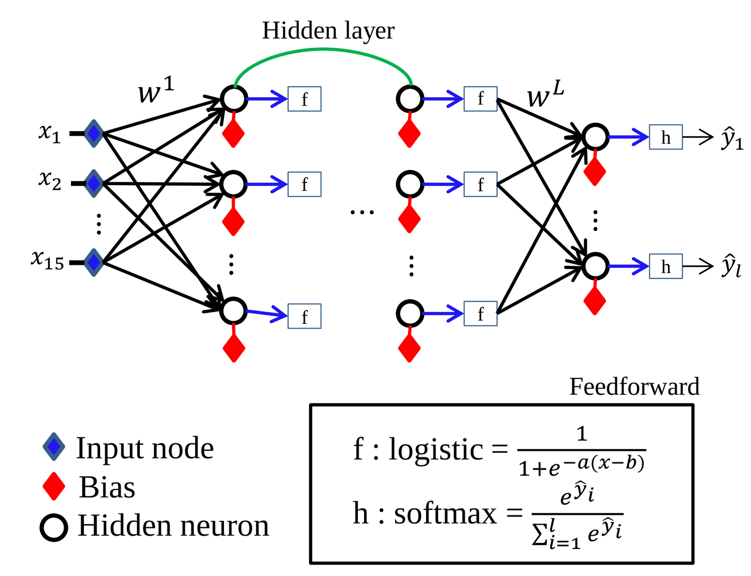
If this inequality is satisfied for all j, then it is considered to be one of the harmonic motif.

For all harmonic motif
Data input size = 15

- x_1 : average number of chords with Z-Relation
- x_2 : average number of dominant chords
- x_3 : average number of tonic chords
- x_4 : average number of repeated diatonic notes
- x_5 : average number of consonant chords
- x_6 : average number of the forte set class / 10 (normalized)
- x_7 : standard deviation of the number of the forte set class / $\sqrt{10}$ (normalized)
- x_8 : average number of the cardinality of the multiset(= # of pitch values) / 10
- x_9 : standard deviation of the cardinality of the multiset / $\sqrt{10}$
- x_{10} : average number of minor 2nd / major 7th
- x_{11} : average number of major 2nd / minor 7th
- x_{12} : average number of minor 3rd / major 6th
- x_{13} : average number of major 3rd / minor 6th
- x_{14} : average number of perfect 4th / perfect 5th
- x_{15} : average number of tritones

* x_{10}~x_{15} is interval vector

Multi-layer Perceptron



Feedforward

f : logistic = $\frac{1}{1+e^{-a(x-b)}}$

h : softmax = $\frac{e^{\hat{y}_i}}{\sum_{i=1}^{l} e^{\hat{y}_i}}$

Backpropagation

$$\boldsymbol{w}^r \mathrel{+}= \Delta \boldsymbol{w}^r$$

Where, $\Delta \boldsymbol{w}^r = -\mu \boldsymbol{\delta}^{r^T} \boldsymbol{y}^{r-1}$, and
μ is learning factor

$\boldsymbol{\delta}^L = \hat{\boldsymbol{y}} - \boldsymbol{y}$

$\boldsymbol{\delta}^r = \mathrm{a}(\boldsymbol{\delta}^{r+1^T} \cdot \boldsymbol{w}^{r+1})^T * \hat{\boldsymbol{y}}(\boldsymbol{1} - \hat{\boldsymbol{y}})$, r-1 < L

Cross-entropy cost function

$J = -\sum_{i=1}^{l} y_i \log \hat{y}_i$

Define 'loss' as ratio of J(new) to J(old)

Adaptive learning

μ = 0.001:0.301:0.001 = [0.001, 0.002, ⋯, 2.999, 3.000]

Initialize μ to a reasonable value. Ex) μ[two thirds of the length]
step = (loss-1)/lf_step , where lf_step means learning factor step

Case 1:
Step > 0 , which means an increase in cost
So, we have to decrease learning factor -> move to left by log(step)

Case 2:
Step < 0 , which means an decrease in cost
So, we have to increase learning factor -> move to right by log(-step)

Results

τ = 0.8, initial weight/bias std = 0.2/0.4, lf_step=0.04
Hidden node = 80, layer = 5, train/ test data =2210/474
a = 2, b = 0.5, Beethoven and Bach

Error rate : 20.25%

60% performance improvement over random selection(50%)

τ = 0.9, initial weight/bias std = 0.2/0.4, lf_step=0.04
Hidden node = 80, layer = 5, train/ test data =4763/1313
a = 2, b = 0.5, Beethoven, Bach and Chopin

Error rate : 38.46%

84.6% performance improvement over random selection(67%)

τ = 0.9, initial weight/bias std = 0.2/0.4, lf_step=0.04
Hidden node = 80, layer = 5, train/ test data =6770/2099
a = 2, b = 0.5, Beethoven, Bach, Chopin and Alkan

Error rate : 51.74%

93% performance improvement over random selection(75%)

τ = 0.8, train/test data = 3695/1128

Error rate : 51.51%

On average, each score has 10 to 20 harmonic motifs, depending on τ. Therefore, if the accuracy of a motif is over 50%, the composer of the score can be detected.

This means that each composer has a different tendency to use chords, and that the computer can recognize and distinguish it.

Feature selection

To check if a feature is valid or not, t-test was used.

$\overline{z_i} = \overline{x_{i,A}} - \overline{x_{i,B}}$, i-th feature in class A/B

Hypothesis : $E[x_{i,A}] = E[x_{i,B}]$

# of class A sample data = N_1 and # of class B sample data = N_2

$\widehat{\sigma^2} = \frac{1}{N_1+N_2-2}(N_1\sigma_1^2 + N_2\sigma_2^2)$: Chi-square distribution $N_1 + N_2 - 2$ degree of freedom

$q = \frac{\overline{z_i}}{\hat{\sigma}\sqrt{\frac{N_1+N_2}{N_1 N_2}}}$, if q lies outside an interval D, which is obtained from t-test table for $N_1 + N_2 - 2$ degrees of freedom and significance level ρ (ex, 0.05), then select

Features that distinguish Beethoven and Bach (ρ = 0.05)
: forte set class and cardinality of the multiset
frequency of dominant use, consonant use, minor 3rd / major 6th , major 3rd / minor 6th , perfect 4th / perfect 5th and tritones.
(feature # : 2, 5, 6, 7, 8, 9, 12, 13, 14, 15)

Because only 15 features are used, There is no significant difference in performance even with only distinctive features

Space for another student