Practical classification method for birdsong with variable note sequences and its application to the whole day recordings

T. KOUMURA & K. OKANOYA (The Univ. of Tokyo)

Goals
Classifying notes in the songs with as little human effort as possible. Characterizing hour-scale modulation in the note sequence pattern.

Song Recording
Sound proof box
Light 13 h
Dark 11 h
Light
Dark
2 days
Sequence analysis
Evaluation of the automatic note classification
2/3 → training
1/3 → validation
3 birds
20523±5992 notes, 44.6±6.9 minutes in a day

Recognition
Output label sequence
\[ A A B B C \]
Correct label sequence
\[ A A B B C \]
Error backpropagation
ANN
Output layer (number of labels)
Full connection layer (120)
Subsampling layer (2×2)
Convolutional layer (5×5)
Subsampling layer (2×2)
Convolutional layer (5×5)
Transition scores
\[ 0.8 \]
\[ 0.8 \]
\[ -0.5 \]
Transition scores
\[ A A B B C C C C C C C C \]

Graph transformer network
Output label sequence
\[ A A B B C \]
Correct label sequence
\[ A A B B C \]
Error backpropagation

ANN
Output layer (number of labels)
Full connection layer (120)
Subsampling layer (2×2)
Convolutional layer (5×5)
Subsampling layer (2×2)
Convolutional layer (5×5)
Transition scores
\[ 0.8 \]
\[ 0.8 \]
\[ -0.5 \]
Transition scores
\[ A A B B C C C C C C C C \]

Graph transformer network

Modeling of transition probabilities at branch points
Motif: frequently appearing sequence pattern
Sound spectrogram
Note sequence
Motif sequence
Branch points in motif sequences
\[ P(x=a|Data, w) = \frac{1}{n_x} \]
\[ P(x=b|Data, w) = \frac{1}{n_x} \]
Estimation of the transition probability
by the moving average of the motif frequencies
\[ f_y = x(t; Data, w) \]
Determination of the most suitable window width
\[ \hat{w} = \arg \max_w \left( \sum P(x=Data, w) \right) \]

Validation score
Error ratio = \[ \frac{\text{Levenshtein distance between correct & output sequences}}{\text{Length of correct sequence}} \]
Transition scores
\[ +0.33±0.15\% \]
\[ -0.21±0.14\% \]
Note positions
\[ +0.33±0.15\% \]
\[ -0.21±0.14\% \]
(Ave. ±sd over 3 birds)

Typical result
Detected note labels & intervals
\[ A B C D E F G H I J \]
Better with transition scores note positions

Result summary

Bird A | 5 | 5 | 12 | 5
---|---|---|---|---
Bird B | 2 | 4 | 2 | 0
Bird C | 2 | 1 | 4 | 0