Bird syllable detection

Mohamed Abrash Paulina Ibek

April 2024

Table of Contents

The Ship of Theseus

The Ship of Theseus (Identity)

The Ship of Theseus paradox questions the identity of an object as its parts are gradually replaced over time. It asks whether the object remains the same if all its original parts are replaced, challenging notions of identity and continuity.

The Ship of Theseus

The ship of Theseus (Similarity)

How much should we change one ship before we can say that the pair is different? what aspects/ features can be changed while still considering the two ships to be similar?

- color?
- shape?
- function?

Bird Song Data and Objectives

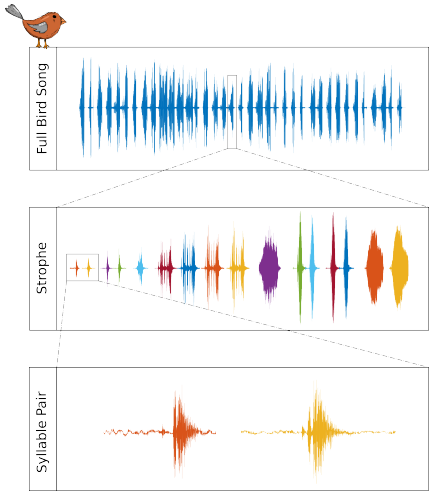
Data:

- 6 bird songs broken into individual syllables
- Similarity labels for consecutive syllables

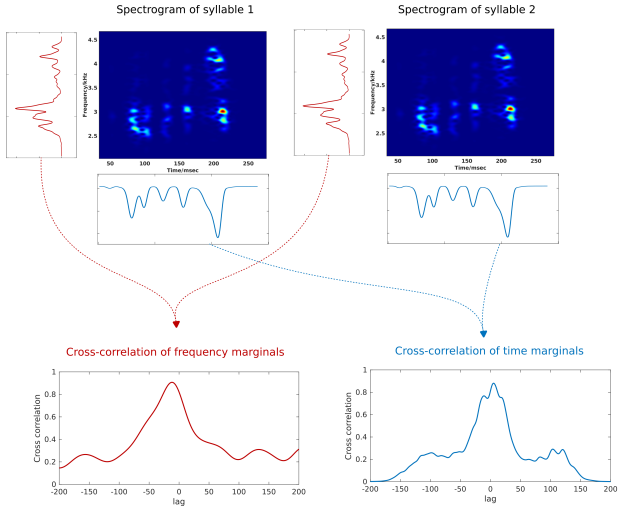
Objective:

- Construct a measure of similarity that can automate the classification of bird syllables

$$\mu(.,.) : X_1(t), X_2(t) \rightarrow \mu(X_1, X_2)$$



The Spectrogram Cross-Correlation Method



Spectrogram Cross-Correlation Results

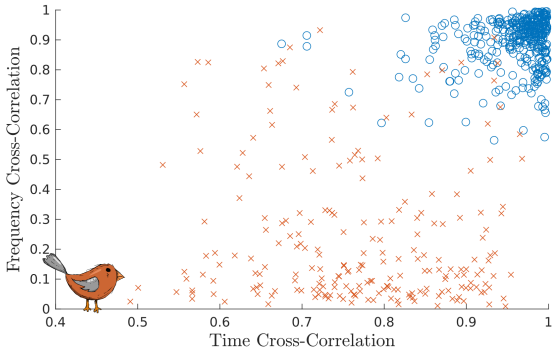


Figure: "o" = similar pairs, "x" = dissimilar pair

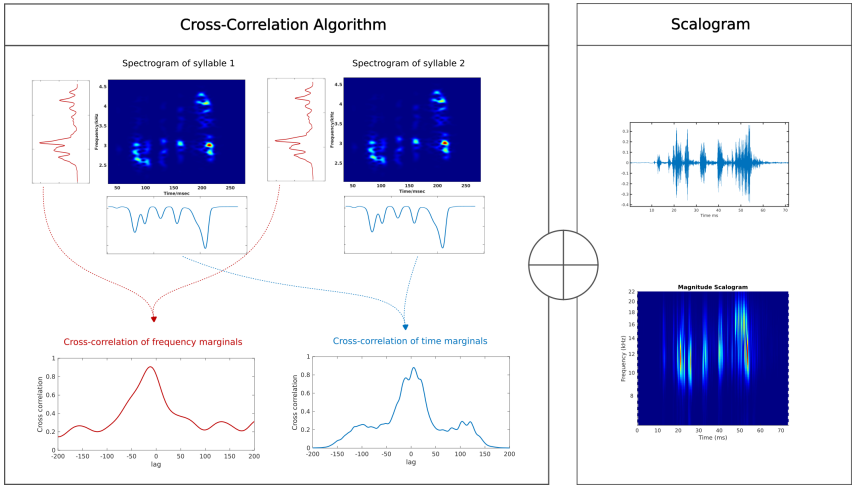
Spectrogram Cross-Correlation Results

Table: SPCC performance on the entire data set comprised of six bird songs.

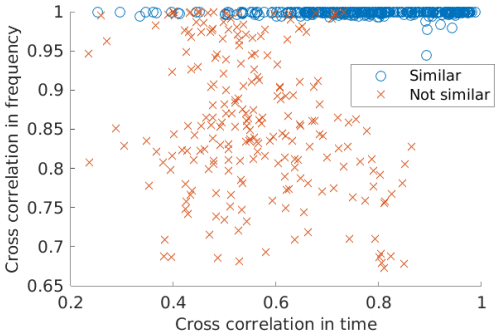
Song	13B04	19A04	23B05	14A03	19A01	23A01
Success rate	97.15%	97.07%	96.90%	96.81%	94.70%	97.26%
Number of evaluations	596	409	419	502	377	401

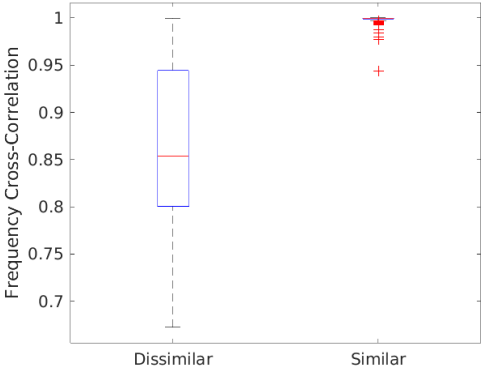
In total: 96.7 % success rate over 2704 evaluations.

The Scalogram Cross-Correlation

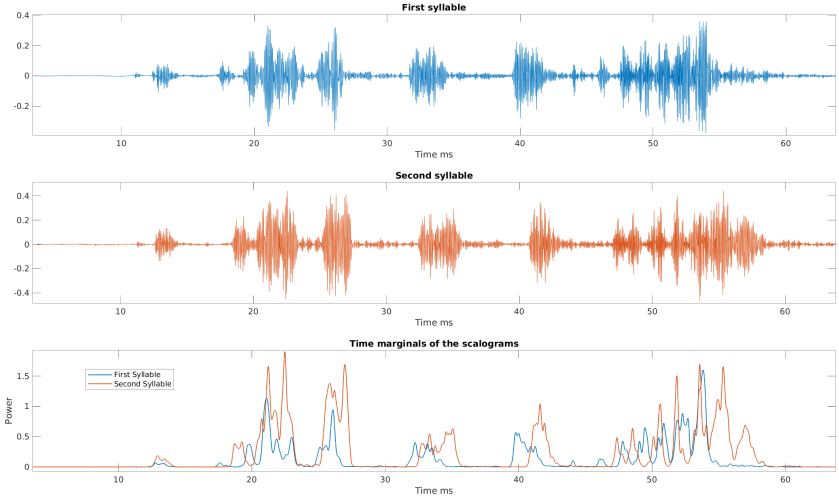


Scalogram Cross-Correlation





Scalogram Cross-Correlation, Miss-Classifications



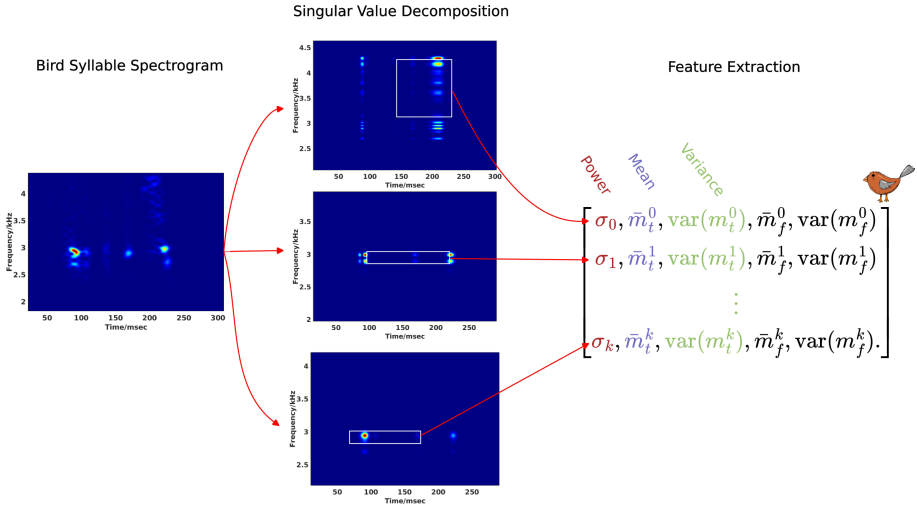
Scalogram Cross-Correlation, Results

Table: WTCC performance on the entire data set comprised of six bird songs.

Song	13B04	19A04	23B05	14A03	19A01	23A01
Success rate	0.9547	0.9609	0.9260	0.9422	0.9390	0.9352
Number of evaluations	596	409	419	502	377	401

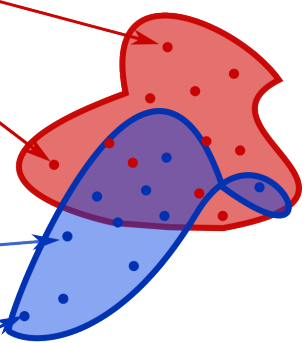
Total 94 % over 2704 evaluations.

Feature Extraction with SVD- algorithm



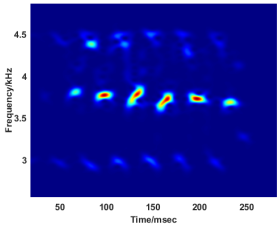
Feature Extraction with SVD- algorithm

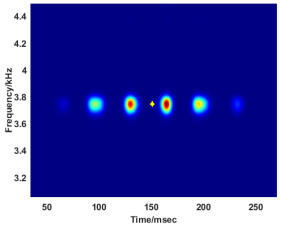
$$\begin{bmatrix} \sigma_0, \bar{m}_t^0, \mathrm{var}(m_t^0), \bar{m}_f^0, \mathrm{var}(m_f^0) \\ \sigma_1, \bar{m}_t^1, \mathrm{var}(m_t^1), \bar{m}_f^1, \mathrm{var}(m_f^1) \\ \vdots \\ \sigma_k, \bar{m}_t^k, \mathrm{var}(m_t^k), \bar{m}_f^k, \mathrm{var}(m_f^k). \end{bmatrix}$$

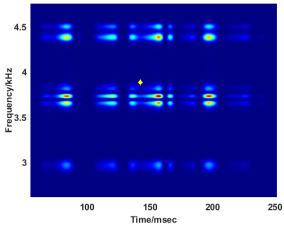


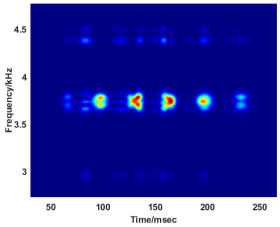
$$\begin{bmatrix} \sigma_0, \bar{m}_t^0, \mathrm{var}(m_t^0), \bar{m}_f^0, \mathrm{var}(m_f^0) \\ \sigma_1, \bar{m}_t^1, \mathrm{var}(m_t^1), \bar{m}_f^1, \mathrm{var}(m_f^1) \\ \vdots \\ \sigma_k, \bar{m}_t^k, \mathrm{var}(m_t^k), \bar{m}_f^k, \mathrm{var}(m_f^k) \end{bmatrix}$$

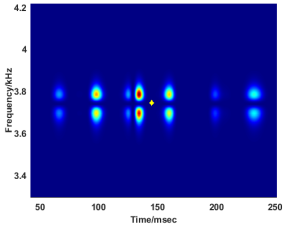
Feature Extraction with SVD

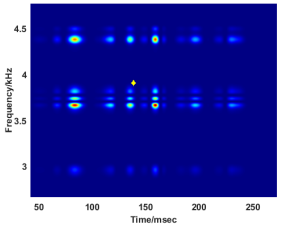




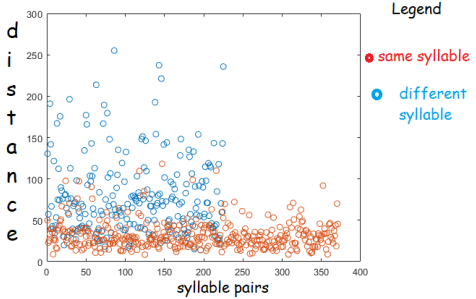


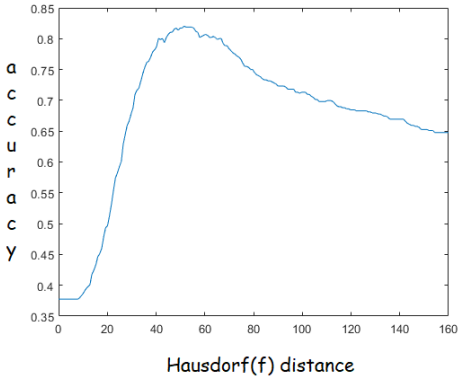






Training


distance
syllable pairs
Legend
same syllable
different syllable


accuracy
Hausdorf(f) distance

SVD results

Table: Accuaracies for the SVD method

Songs	13B05	19A04	23B05	14A03	19A01	23A01
Accuracies	0.8037	0.7995	0.8138	0.7470	0.7507	0.8329

An average success rate of 79 % over 2704 evaluations.

Conclusion

- All methods were successful, but to varying degrees.
- When comparing things, one must determine which features to include
- Having too little detail produces bad results
- Having too much detail made one measure of similarity useless
- Familiarize oneself with the problem and determine what is important in the context