Word Recognition

M. Harding, S. Lane, E. Metheney, K. Stoutz, and R. Whitaker

July 28, 2011
Overview

- Scaling Function and Haar Wavelet
- Discrete Fourier Transform
- Research - “Am”, “Blue”, and “Car”
- Self Comparison with Rhyming Words
- Two Person Comparison
- Other Methods
- Outlook
Scaling Function and Haar Wavelet

Scaling Function

\[ \psi(t) = \begin{cases} 
1 & 0 \leq t < 1/2, \\
-1 & 1/2 \leq t < 1, \\
0 & \text{otherwise.} 
\end{cases} \]

Haar Wavelet

\[ \phi(t) = \begin{cases} 
1 & 0 \leq t < 1, \\
0 & \text{otherwise.} 
\end{cases} \]
Discrete Fourier Transform

Given a sequence of \( N \) samples \( f(n) \), indexed by \( n = 0, 1, \ldots, N - 1 \), the DFT is defined as \( F(k) \), where \( k = 0, 1, \ldots, N - 1 \):

\[
F(k) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} f(n) e^{i2\pi kn/N}
\]
Algorithm

1. Create a library of recordings
2. Perform a wavelet decomposition
3. Compute error between the sample and library recordings
4. Display and evaluate results
Wavelet Decomposition

- Note the change in the range from the original to the detail.
Computing Error

Example of a Mat on a Signal
“Am”, “Blue”, and “Car”

Test for the Word Car

Error

Grid Size
"Fat"
Rhyming Words
"Hat"
“Rat”
The Issue

- Although this algorithm is effective it is not practical.
- It is necessary to be able to compare samples to a library created by another individual.
- When comparing two different people the problem becomes more complicated.
Example: Am

- These graphs are the original signals before DFT, filtering, or compression have been applied.
Need for Discrete Fourier Transform
Distinguishing Features
Algorithm

1. Analyze the first third of the signal.
2. Compare to the library of recordings of the word Car.
3. Calculate the error between the sample and the recordings in the library as previously described in the first algorithm.
4. Repeat using the middle third of the signal and compare to recordings of the word Blue.
5. Repeat using the final third of the signal and compare to recordings of the word Am.
6. Graph the errors together on the same plot.
Two Person Comparison “Car”
Linear Predictive Coding

- Models the process of speech production
- Predicts the next speech sample using a linear combination of previous samples
- Used in telephone systems, voice-mail systems, and telephone answering machines

Equation of Predicted Signal

$$\hat{x}(n) = \sum_{i=1}^{p} a_i x(n - i)$$
Outlook

- Time Efficiency
- Accents
Acknowledgement

We would like to thank Dr. G. Berkolaiko, Dr. C. Liaw, and Joe Bartley for their patience and contribution to our research.