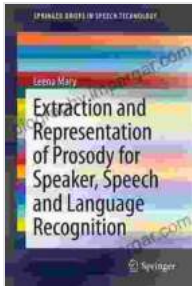


Extraction and Representation of Prosody for Speaker Speech and Language



Extraction and Representation of Prosody for Speaker, Speech and Language Recognition (SpringerBriefs in Speech Technology) by Leena Mary

★★★★☆ 4.7 out of 5

Language : English
File size : 1873 KB
Text-to-Speech : Enabled
Screen Reader : Supported
Enhanced typesetting : Enabled
Print length : 70 pages



Prosody is a key component of human speech and language. It refers to the variations in pitch, loudness, and timing that occur over the course of an utterance. These variations can convey a wide range of information, including the speaker's emotional state, their intentions, and the structure of the utterance.

The extraction and representation of prosody is a challenging task, but it is essential for the development of natural-sounding speech synthesis and recognition systems. In this article, we will provide a comprehensive overview of the different types of prosodic features, the methods used to extract these features, and the various ways in which they can be represented.

Types of Prosodic Features

There are a wide range of prosodic features that can be extracted from speech. These features can be divided into three main categories: pitch, loudness, and timing.

- **Pitch** refers to the frequency of the vocal cords' vibration. It is measured in hertz (Hz). The average pitch of human speech is around 120 Hz for women and 150 Hz for men.
- **Loudness** refers to the amplitude of the sound waves produced by the vocal cords. It is measured in decibels (dB). The average loudness of human speech is around 60 dB.
- **Timing** refers to the duration of speech sounds. It is measured in milliseconds (ms). The average duration of a syllable in English is around 100 ms.

Methods for Extracting Prosodic Features

There are a variety of methods that can be used to extract prosodic features from speech. These methods can be divided into two main categories: acoustic analysis and articulatory analysis.

- **Acoustic analysis** involves the analysis of the sound waves produced by the vocal cords. This can be done using a variety of techniques, including:
 - **Time-domain analysis** measures the amplitude and frequency of the sound waves over time.
 - **Frequency-domain analysis** measures the distribution of energy across different frequencies.

- **Cepstral analysis** measures the relationship between the time domain and frequency domain representations of the sound waves.
- **Articulatory analysis** involves the analysis of the movements of the articulators (i.e., the lips, tongue, and jaw) during speech production. This can be done using a variety of techniques, including:
 - **Electromyography (EMG)** measures the electrical activity of the muscles that control the articulators.
 - **Electromagnetic articulography (EMA)** measures the movements of the articulators using small magnets attached to the lips, tongue, and jaw.
 - **Optical tracking** measures the movements of the articulators using a camera.

Representation of Prosodic Features

Once prosodic features have been extracted from speech, they need to be represented in a way that can be used by speech synthesis and recognition systems. There are a variety of different ways to represent prosodic features, including:

- **Symbolic representations** use symbols to represent different prosodic features. For example, a high pitch might be represented by the symbol "H", a low pitch by the symbol "L", and a rising pitch by the symbol "↑".
- **Numeric representations** use numbers to represent different prosodic features. For example, a high pitch might be represented by the

number 1, a low pitch by the number 0, and a rising pitch by the number 0.5.

- **Graphical representations** use graphs to represent different prosodic features. For example, a pitch contour might be represented by a line graph, a loudness contour might be represented by a bar graph, and a timing contour might be represented by a scatter plot.

Applications of Prosody

Prosody has a wide range of applications in speech synthesis and recognition. In speech synthesis, prosody can be used to make synthetic speech sound more natural and expressive. In speech recognition, prosody can be used to improve the accuracy of recognition systems.

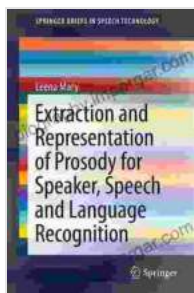
In addition to speech synthesis and recognition, prosody has also been used in a variety of other applications, including:

- **Forensic analysis:** Prosody can be used to identify speakers and to detect deception.
- **Medical diagnosis:** Prosody can be used to diagnose certain medical conditions, such as Parkinson's disease and autism.
- **Music analysis:** Prosody can be used to analyze the rhythm and melody of music.

Prosody is a key component of human speech and language. The extraction and representation of prosody is a challenging task, but it is essential for the development of natural-sounding speech synthesis and recognition systems. In this article, we have provided a comprehensive overview of the different types of prosodic features, the methods used to

extract these features, and the various ways in which they can be represented.

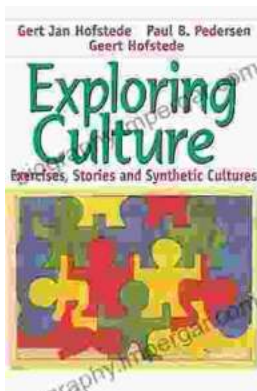
We hope that this article has been helpful. If you have any questions, please feel free to contact us.



Extraction and Representation of Prosody for Speaker, Speech and Language Recognition (SpringerBriefs in Speech Technology) by Leena Mary

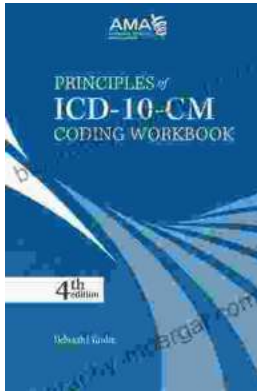
★★★★☆ 4.7 out of 5

Language : English
File size : 1873 KB
Text-to-Speech : Enabled
Screen Reader : Supported
Enhanced typesetting : Enabled
Print length : 70 pages



Exploring Culture: Exercises, Stories, and Synthetic Cultures

Culture is a complex and multifaceted concept that shapes our lives in countless ways. It influences our beliefs, values, behaviors, and even our physical appearance. In...



Principles of ICD-10 Coding Workbook: Your Comprehensive Guide to Accurate and Efficient Medical Documentation

Empower Yourself with the Knowledge and Skills for Expert ICD-10 Coding In today's healthcare landscape, accurate and efficient medical coding is...