

zero_crossing_rate, rms mel_spectrogram



What is ZCR?

- ZCR cao thường liên quan đến âm thanh có tần số cao hoặc nhiều thành phần tần số cao, như các âm thanh sắc bén, tiếng ồn trắng hoặc giọng nói trong trạng thái căng thẳng.
- ZCR thấp thường xuất hiện ở các âm thanh có tần số thấp hoặc ít thay đổi tần số, như các âm thanh mềm mại, giọng nói trầm hoặc trong các trạng thái cảm xúc như buồn bã.

đo lường tần suất
mà tín hiệu
chuyển từ giá trị
dương sang giá
trị âm và ngược
lại trong một
đoạn tín hiệu



Formula

$$\text{ZCR} = \frac{1}{N-1} \sum_{n=1}^{N-1} 1_{\text{sign}(x[n]) \neq \text{sign}(x[n-1])}$$

Trong đó:

- N là số lượng mẫu trong một khung thời gian.
- $x[n]$ là giá trị của mẫu tại thời điểm n .
- $\text{sign}(x)$ là hàm dấu, xác định dấu của tín hiệu.



Hiệu quả

Phân biệt giọng nói cảm xúc: ZCR có thể phản ánh mức độ "sắc bén" của tín hiệu giọng nói. Ví dụ, trong trạng thái tức giận hoặc phấn khích, giọng nói thường có xu hướng nhanh và sắc bén hơn, có thể dẫn đến ZCR cao hơn. Ngược lại, các trạng thái như buồn bã hay mệt mỏi có thể dẫn đến ZCR thấp hơn do giọng nói chậm và ít thay đổi dấu.

Attention

việc trích xuất đặc trưng ZCR có thể mang lại giá trị trong việc nhận diện cảm xúc qua giọng nói, nhưng để đạt được độ chính xác cao nhất, nên sử dụng nó cùng với các đặc trưng âm thanh khác.

RMS

là một tham số quan trọng trong xử lý âm thanh, được sử dụng để đo mức độ năng lượng hoặc cường độ trung bình của tín hiệu âm thanh. RMS cung cấp một giá trị số học đại diện cho biên độ tín hiệu, giúp đánh giá mức độ "mạnh" hoặc "yếu" của âm thanh.

$$RMS_t = \sqrt{\frac{1}{K} \cdot \sum_{k=t \cdot K}^{(t+1) \cdot K - 1} s(k)^2}$$

Example

Ví Dụ Minh Họa

Giả sử chúng ta có một đoạn tín hiệu âm thanh với các giá trị biên độ mẫu như sau:

$$x = [0.5, -0.3, 0.4, -0.6, 0.2, -0.1]$$

Để tính RMS của đoạn tín hiệu này, ta thực hiện các bước sau:

1. Bình phương từng giá trị biên độ: $[0.25, 0.09, 0.16, 0.36, 0.04, 0.01]$
2. Tính giá trị trung bình của các bình phương: $\frac{0.25+0.09+0.16+0.36+0.04+0.01}{6} = 0.1517$
3. Lấy căn bậc hai của giá trị trung bình này: $\text{RMS} = \sqrt{0.1517} \approx 0.3895$

Giá trị RMS 0.3895 cho biết mức năng lượng trung bình của tín hiệu âm thanh trong đoạn này.

