

Sinh viên thực hiện: Lê Thị Bình 19TCLC\_DT2

Giáo viên hướng dẫn: TS.Ninh Khánh Duy

# I. Tổng quan

## a. Khái niệm về voiced và unvoiced

- Voiced (âm hữu thanh): là những âm sẽ làm rung thanh quản khi phát âm.
- Unvoiced(âm vô thanh):là những âm phát ra nhưng không tạo độ rung từ thanh quản.

#### b. Nhiệm vụ

- Phân tích đoạn tín hiệu huấn luyện thành các khoảng voiced và unvoiced bằng việc sử dụng các đặc trưng tín hiệu đã học.

## II. Các đặc trưng tín hiệu

### a. Short-Time Energy (STE)

- Là tổng bình phương các giá trị dạng sóng trên một số lượng mẫu hữu hạn thuộc một khung (20-25ms).

$$STE[n] = \sum_{m=0}^{N-1} x^2[n-m]$$

n: frame index

m: sample index

N: frame length (samples)

## b. Zero-Crossing Rate (ZCR)

- Là tốc độ mà dạng sóng đi qua trục 0.

$$ZCR[n] = \sum_{m=0}^{N-1} |sgn(x[n-m]) - sgn(x[n-m-1])|$$

n: frame index

m: sample index

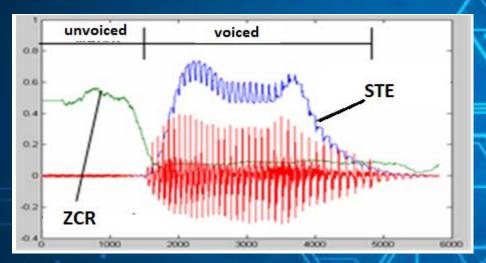
N: frame length

where sgn(.) is the signum function,

$$\operatorname{sgn}(x) = \begin{cases} 1 & x \ge 0, \\ -1 & x < 0. \end{cases}$$

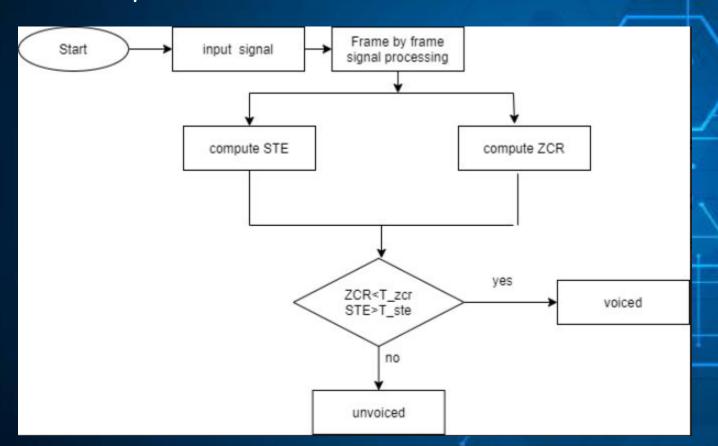
## c. Sự tương quan dữ liệu giữa 2 đặc trưng STE Và ZCR

Xét 1 đoạn tín hiệu hữu thanh (voiced)



 Khi đi qua đoạn tín hiệu voiced thì ZCR có biên độ nhỏ, còn STE có biên độ lớn.

# III. Thuật toán



## IV. Thuật toán tìm ngưỡng

#### a. Ngưỡng STE (T\_ste)

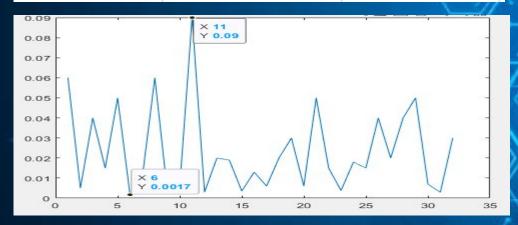
- Dựa trên đặc trưng STE khi đi qua đoạn tín hiệu voiced có biên độ lớn.
- Xét ngưỡng cho STE là giá trị biên độ có được khi khảo sát 4 tín hiệu trong các đoạn voiced với những biên độ nhỏ nhất.
- Hàm STE được chuẩn hóa về hàm có giới hạn [0;1]:

STE= STE/max(STE)

 Biểu đồ thống kê các biên độ có giá trị nhỏ nhất trong các khoảng voiced của 4 tín hiệu

N No. 10	meanV_STE	stdV_STE
Studio_M1	0.0286	0.0247
Studio_F1	0.0253	0.0313
Phone_M1	0.0174	0.014
Phone_M2	0.0237	0.0151

=> T\_ste= 0.0236



Dựa vào thống kê cùng với đó là việc thử các tín hiệu huấn luyện thì ngưỡng tìm được phù hợp: 0.02

Biểu đồ các phần tử được xét của 4 tín hiệu.

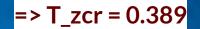
### b. Ngưỡng ZCR (T\_zcr)

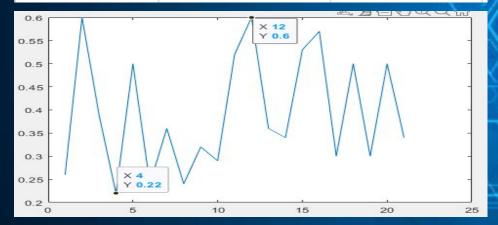
- Dựa trên đặc trưng ZCR khi đi qua đoạn tín hiệu voiced có biên độ nhỏ
- Xét ngưỡng cho ZCR là giá trị biên độ có được khi khảo sát 4 tín hiệu trong các đoạn voiced với những biên độ lớn nhất
- Hàm ZCR được chuẩn hóa về hàm có giới hạn [0;1]:

ZCR= ZCR/max(ZCR)

 Biểu đồ thống kê các biên độ có giá trị lớn nhất trong các khoảng voiced của 4 tín hiệu

	meanV_ZCR	stdV_ZCR
Studio_M1	0.4027	0.2327
Studio_F1	0.3025	0.0506
Phone_M1	0.43	0.1225
Phone_M2	0.4233	0.1268





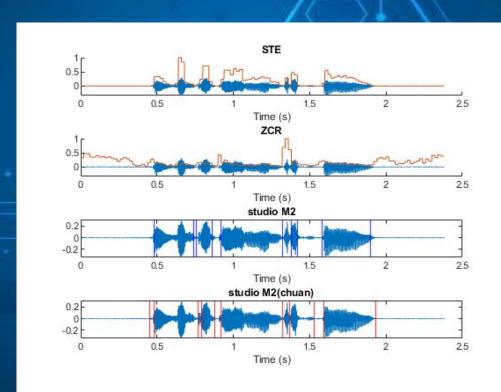
Dựa vào thống kê cùng với đó là việc thử các tín hiệu huấn luyện thì ngưỡng tìm được phù hợp: 0.4

Biểu đồ các phần tử được xét của 4 tín hiệu.

- Nhận xét về cách tìm ngưỡng và thuật toán của chương trình
- + Thuật toán
- Đơn giản, độ chính xác phụ thuộc vào ngưỡng được xác định.
- + Tìm ngưỡng
- Dựa vào thống kê của các tín hiệu huấn luyện.
- => Độ chính xác của thuật toán không cao
- => Gây ra sai lệch trong kết quả với các tín hiệu thực nghiệm, tùy vào từng tín hiệu mà độ sai lệch cũng khác nhau.

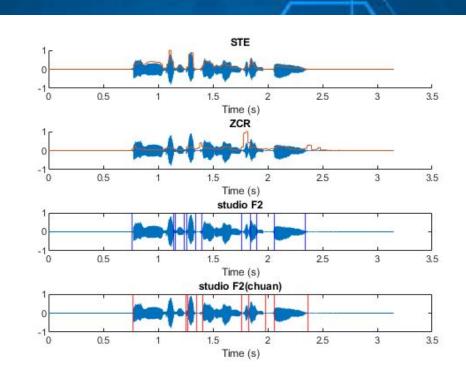
# V. Kết quả chạy các file tín hiệu thực nghiệm

- a. Tín hiệu studio\_M2
- Nhận xét
- Với tín hiệu này kết quả sai lệch các biên khá nhiều so với kết quả chuẩn
- Nguyên nhân
- Do các âm hữu thanh(voiced) trong đoạn lệch có STE thấp dẫn đến việc xác định bị sai.



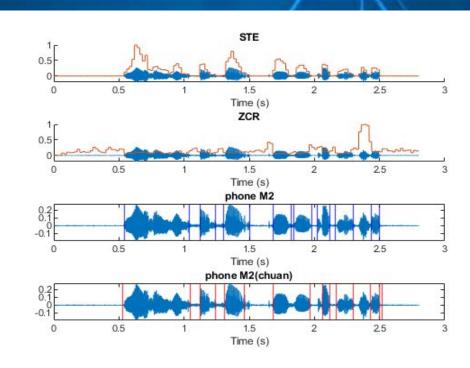
### b. Tín hiệu studio\_F2

- Nhận xét
- Với tín hiệu này kết quả sai lệch các biên ít so với kết quả chuẩn
- Một khoảng tín hiệu voiced bị nhầm thành unvoiced
- Nguyên nhân
- Do việc đặt ngưỡng gây ra



### c. Tín hiệu phone\_M2

- Nhận xét
- Với tín hiệu này kết quả sai lệch các biên ít so với kết quả chuẩn
- Một khoảng tín hiệu voiced bị nhầm thành unvoiced
- Nguyên nhân
- Do việc đặt ngưỡng gây ra



### c. Tín hiệu phone\_F2

- Nhận xét
- Với tín hiệu này kết quả sai lệch các biên ít so với kết quả chuẩn
- Một khoảng tín hiệu voiced bị nhầm thành unvoiced
- Nguyên nhân
- Do việc đặt ngưỡng gây ra

