

Bài Báo Cáo

PHÂN ĐOẠN TÍN HIỆU THÀNH TIẾNG NÓI VÀ KHOẢNG LẶNG

NHÓM 7

- *Thành viên :*
- *-Bùi Duy Phương*
- *-Thái Văn Thọ*
- *-Nguyễn Thị Nga*
- *-Nguyễn Văn Hải*

I. ĐẶT VẤN ĐỀ

-Phân đoạn tín hiệu tiếng nói (speech signal segmentation) là quá trình xác định ranh giới giữa các từ, âm tiết hoặc âm vị trong các ngôn ngữ tự nhiên được nói



II. LÝ THUYẾT XỬ LÝ TÍN HIỆU TIẾNG NÓI VÀ CÁC THUẬT TOÁN

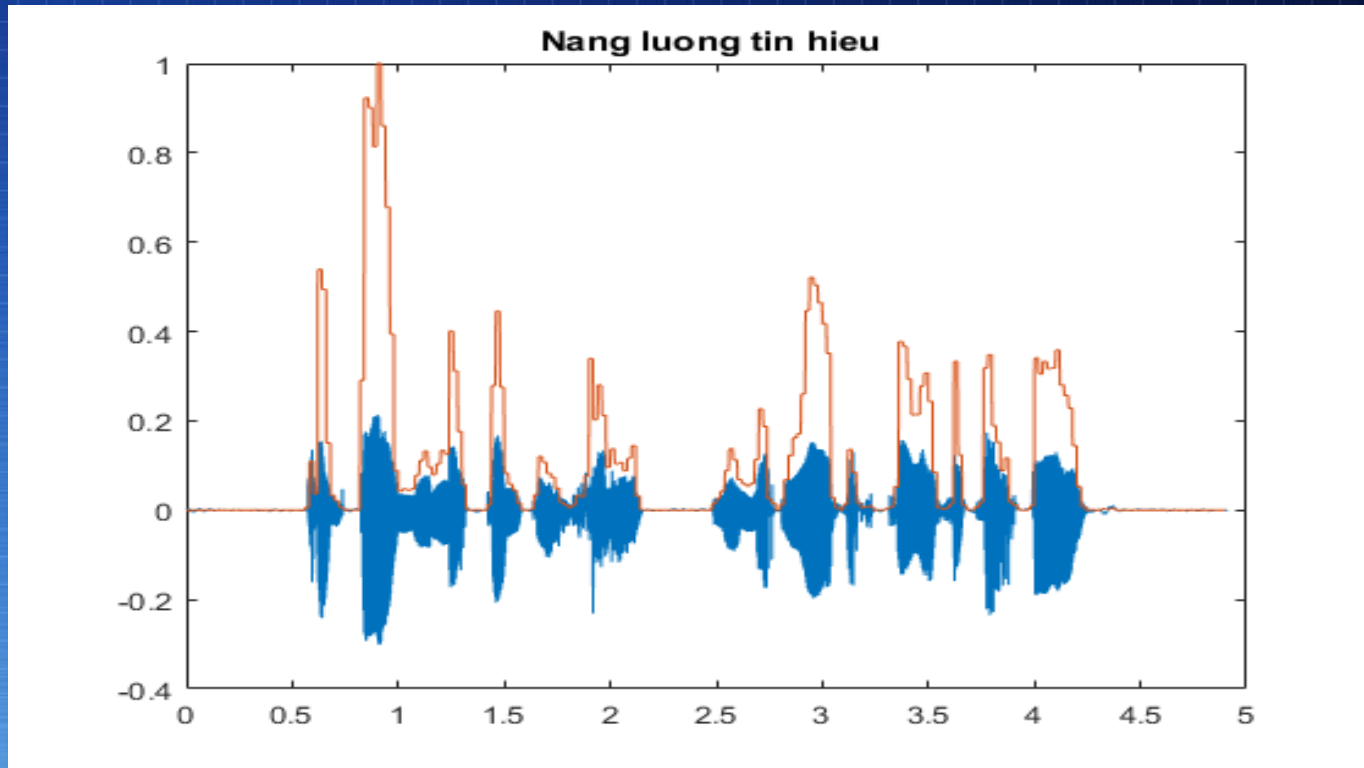
1. Thuật toán Short-Time Energy

- Năng lượng là tổng bình phương biên độ của các mẫu tín hiệu:

$$E_x = \sum_{-\infty}^{+\infty} |x[n]|^2$$

II. LÝ THUYẾT XỬ LÝ TÍN HIỆU TIẾNG NÓI VÀ CÁC THUẬT TOÁN

1. Thuật toán Short-Time Energy

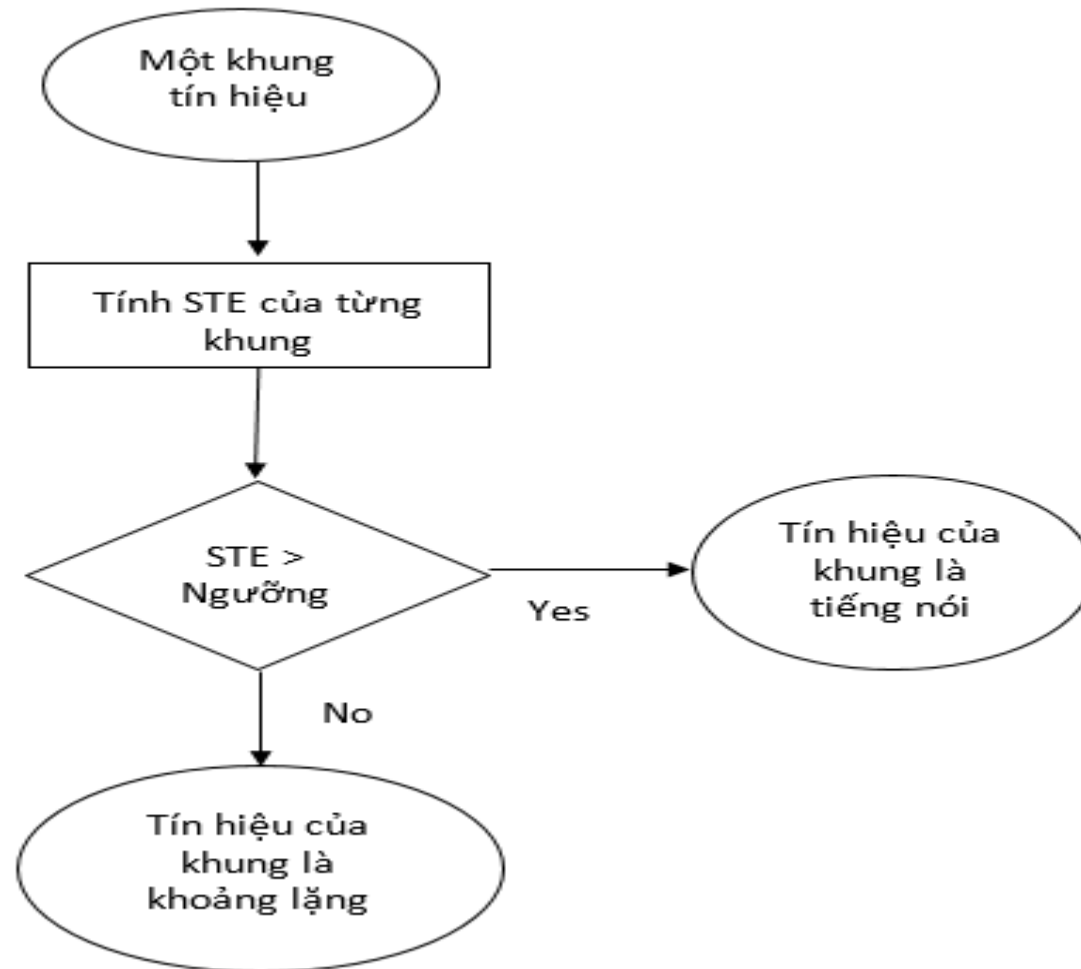


Hình 1. Ví dụ thuật toán Short-Time Energy trên file studio_female.wav

II. LÝ THUYẾT XỬ LÝ TÍN HIỆU TIẾNG NÓI VÀ CÁC THUẬT TOÁN

1. Thuật toán Short-Time Energy

2. Sơ đồ thuật toán



II. LÝ THUYẾT XỬ LÝ TÍN HIỆU TIẾNG NÓI VÀ CÁC THUẬT TOÁN

2. Thuật toán kết hợp Zero Crossing Rate và Short-Time Energy

Công thức hàm Zero-Crossing Rate có thể được biểu diễn như sau:

$$ZCR_n = \sum_{m=-\infty}^{\infty} |\operatorname{sgn}[x(m)] - \operatorname{sgn}[x(m-1)]| w(n-m)$$

Với:

$$\operatorname{sgn}[x(n)] = \begin{cases} 1 & x(n) \geq 0 \\ -1 & x(n) < 0 \end{cases}$$

$$w(n) = \begin{cases} \frac{1}{2N} & 0 \leq n \leq N-1 \\ 0 & \text{ngược lại} \end{cases}$$

II. LÝ THUYẾT XỬ LÝ TÍN HIỆU TIẾNG NÓI VÀ CÁC THUẬT TOÁN

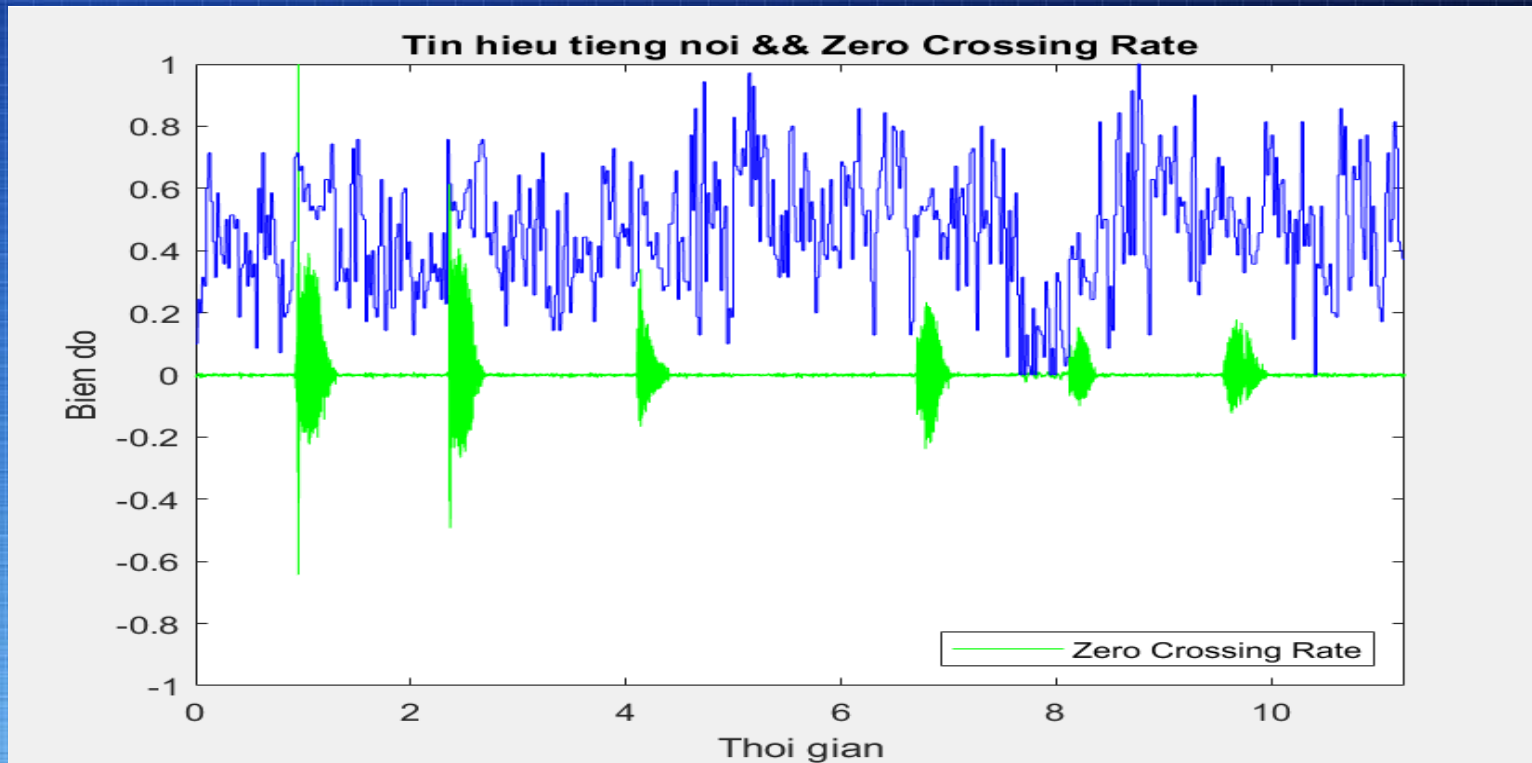
2. Thuật toán kết hợp Zero Crossing Rate và Short-Time Energy

Sau khi chia tín hiệu thành các khung nhỏ, với mỗi khung m , ZCR_m là số lần biên độ tín hiệu bằng qua 0 đối với các mẫu thuộc khung :

$$ZCR_m = \lim_{n \rightarrow N} \left(\frac{\text{số lần } (x[n] * x[n+1]) < 0}{2L+1} \right) \quad \text{với } N \in [-L, L]$$

II. LÝ THUYẾT XỬ LÝ TÍN HIỆU TIẾNG NÓI VÀ CÁC THUẬT TOÁN

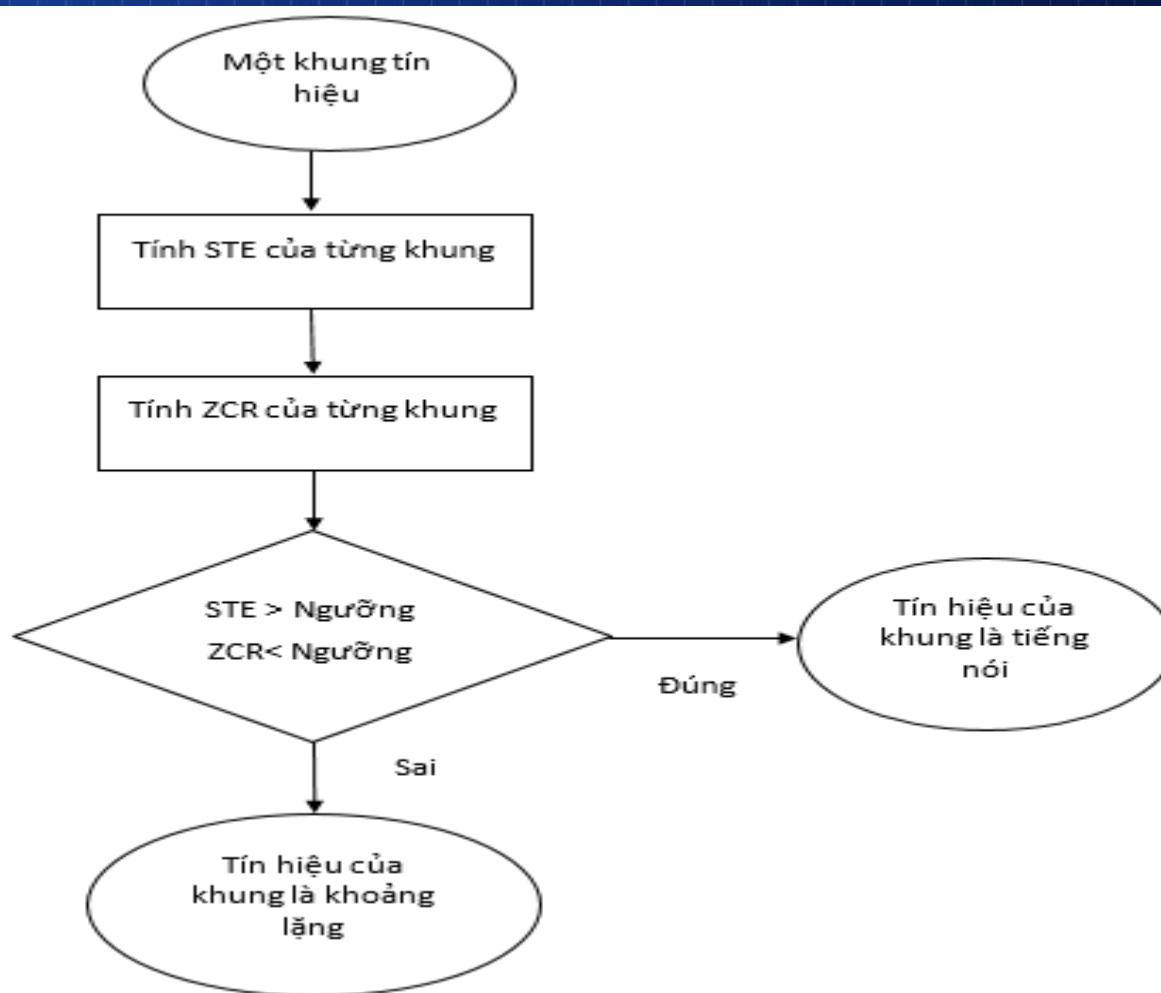
2. Thuật toán kết hợp Zero Crossing Rate và Short-Time Energy



Hình 2. Ví dụ thuật toán kết hợp ZCR trên file lab_male.wav

II. LÝ THUYẾT XỬ LÝ TÍN HIỆU TIẾNG NÓI VÀ CÁC THUẬT TOÁN

2. Thuật toán kết hợp Zero Crossing Rate và Short-Time Energy



II. LÝ THUYẾT XỬ LÝ TÍN HIỆU TIẾNG NÓI VÀ CÁC THUẬT TOÁN

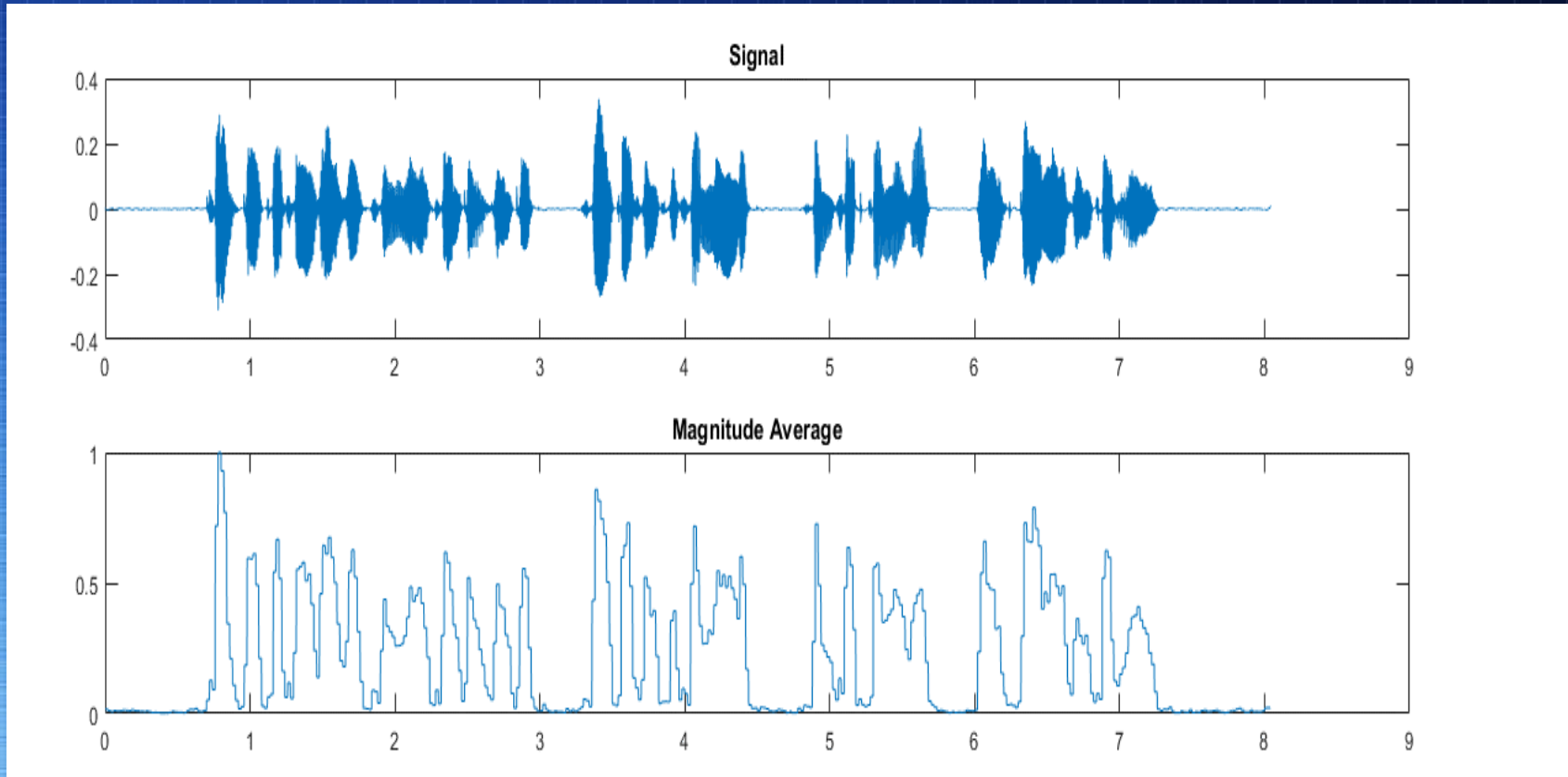
3. Thuật toán Magnitude Average

Thuật toán Magnitude Average được định nghĩa như sau:

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

II. LÝ THUYẾT XỬ LÝ TÍN HIỆU TIẾNG NÓI VÀ CÁC THUẬT TOÁN

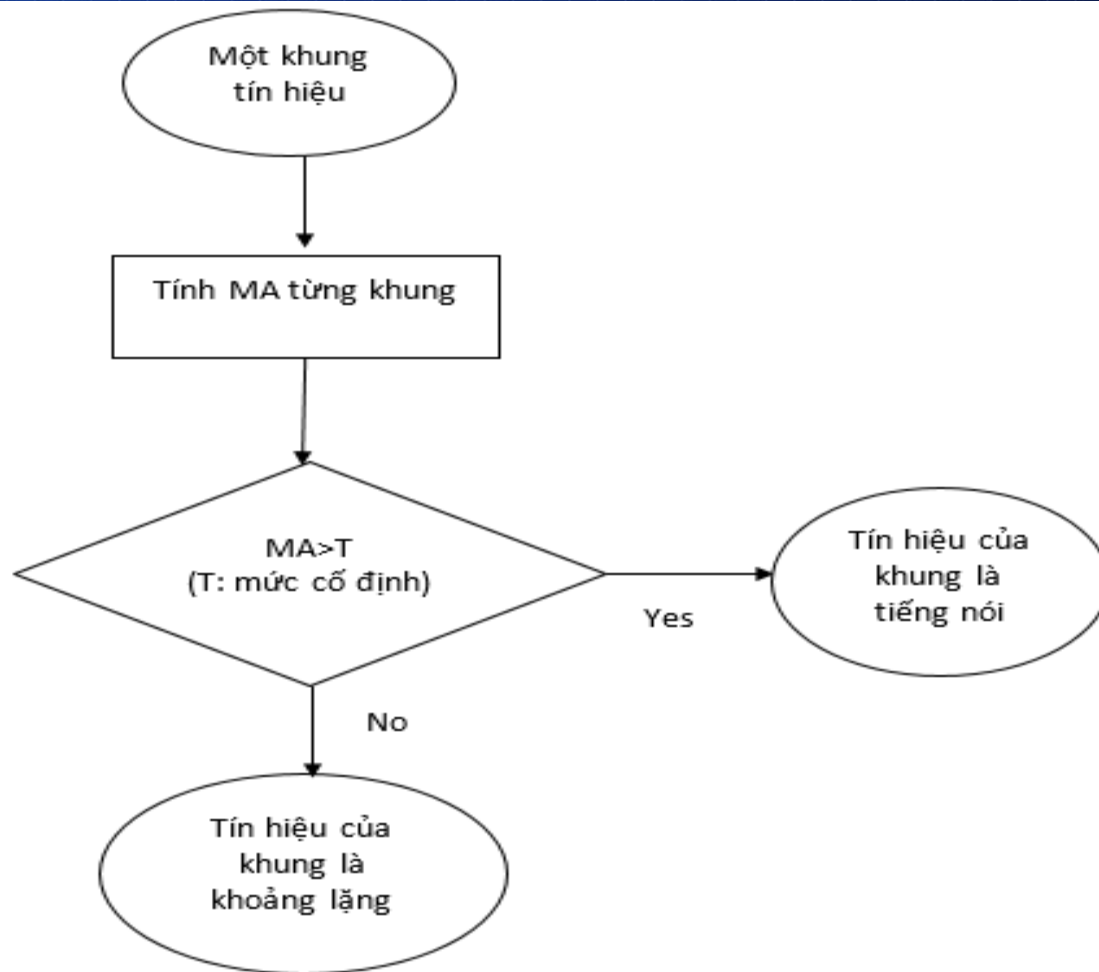
3. Thuật toán Magnitude Average



Hình 3 .Ví dụ về sử dụng hàm MA(Magnitude Average)

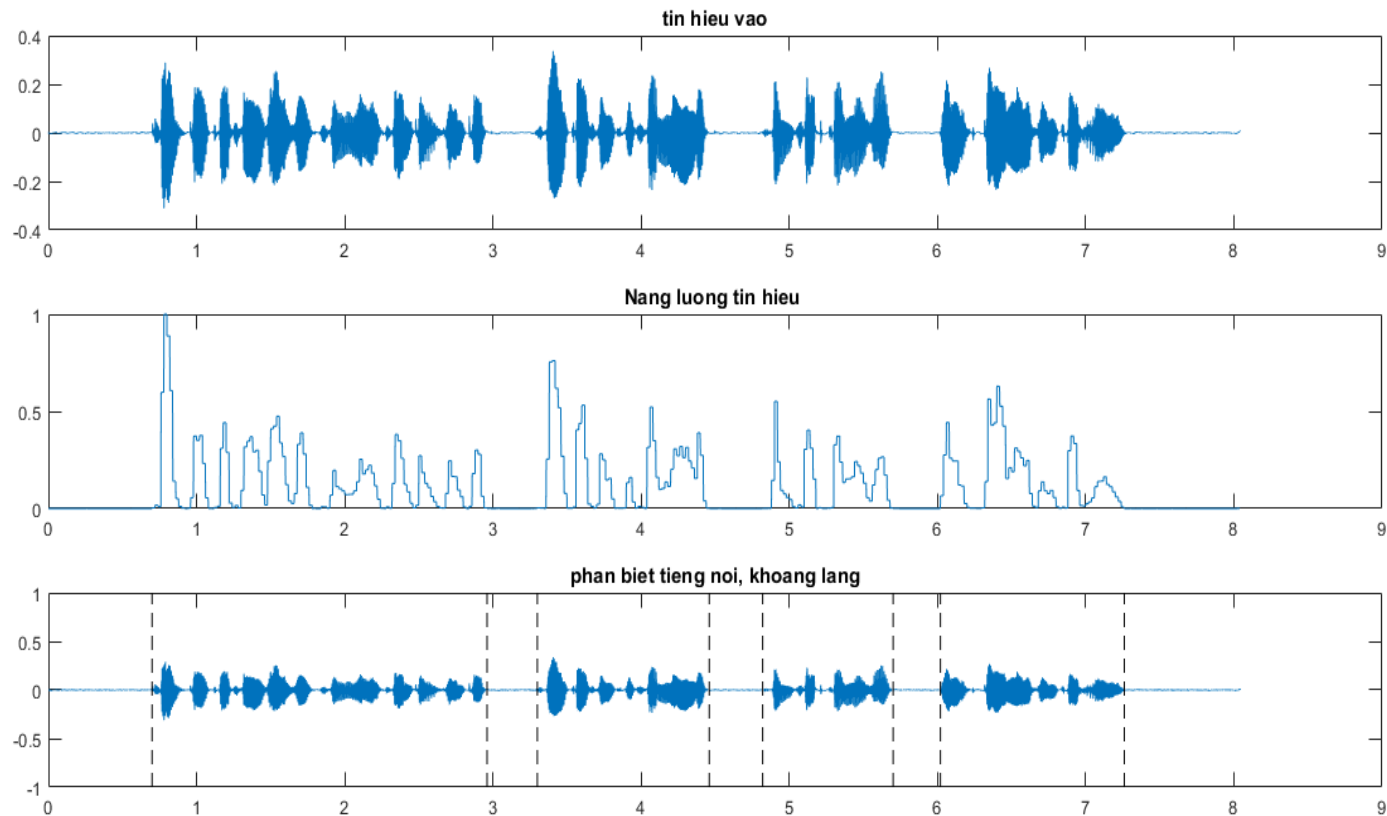
II. LÝ THUYẾT XỬ LÝ TÍN HIỆU TIẾNG NÓI VÀ CÁC THUẬT TOÁN

3. Thuật toán Magnitude Average



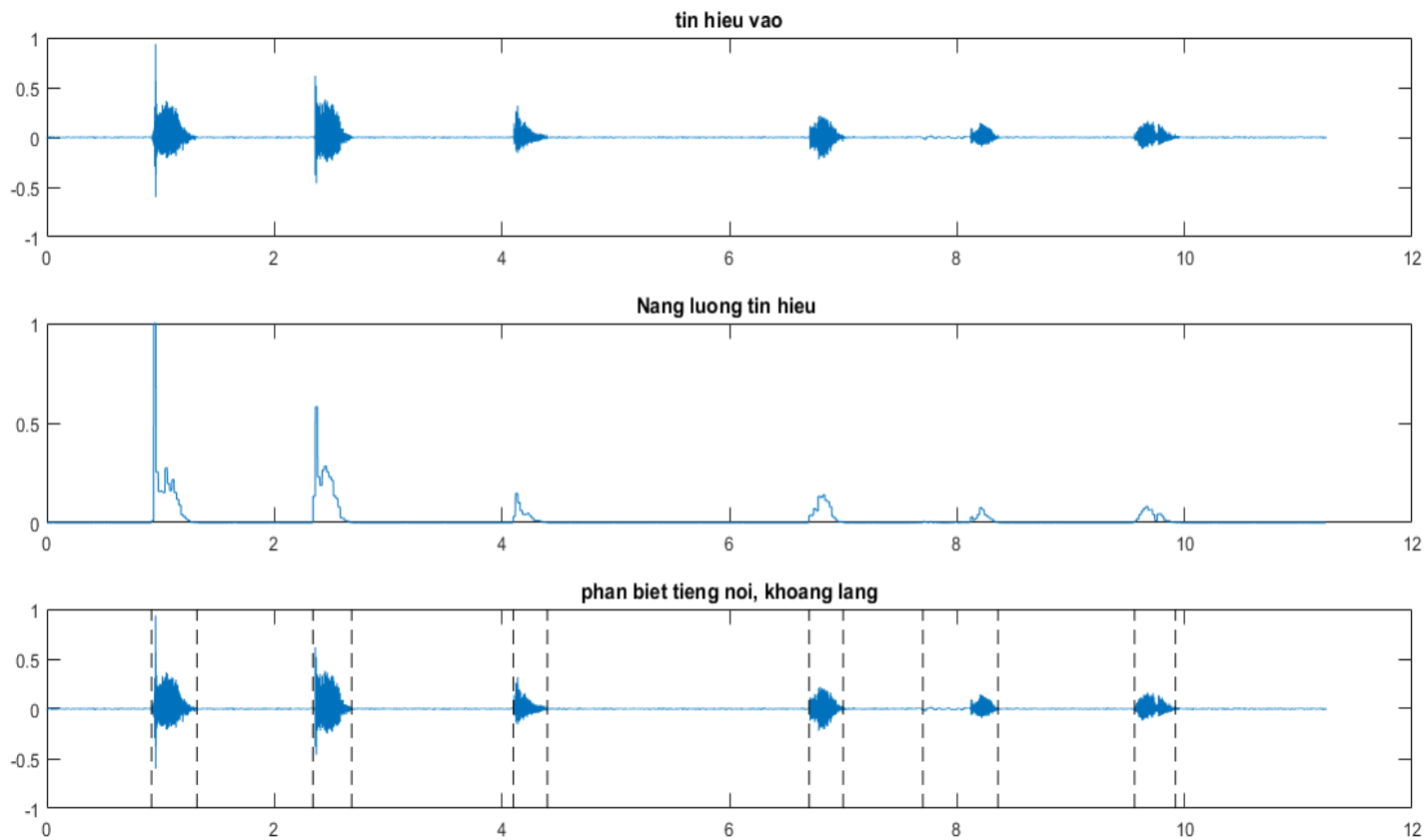
III. KẾT QUẢ THỰC NGHIỆM

1. Thuật toán Short-Time Energy



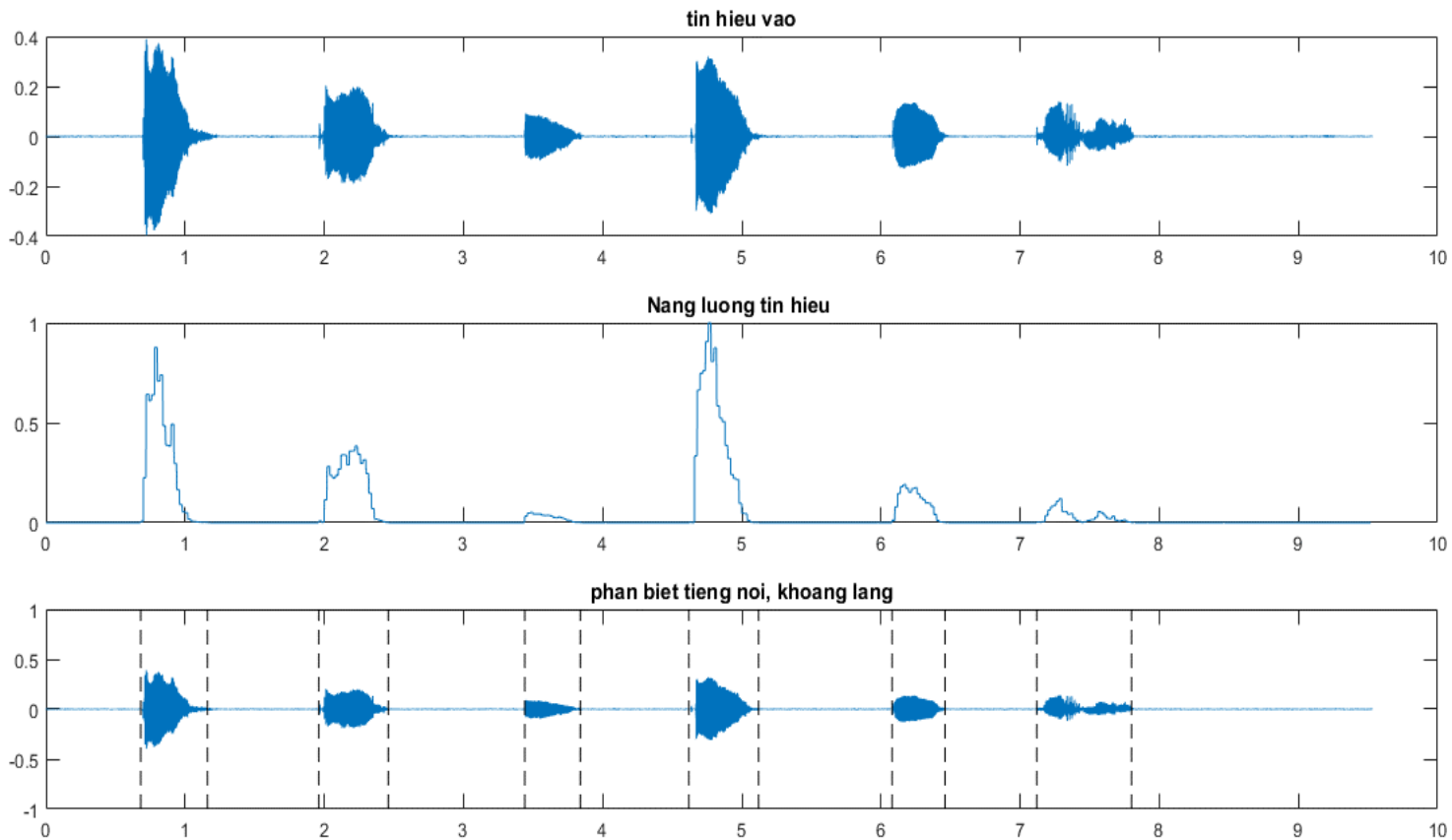
Hình 4. Studio_male

1. Thuật toán Short-Time Energy



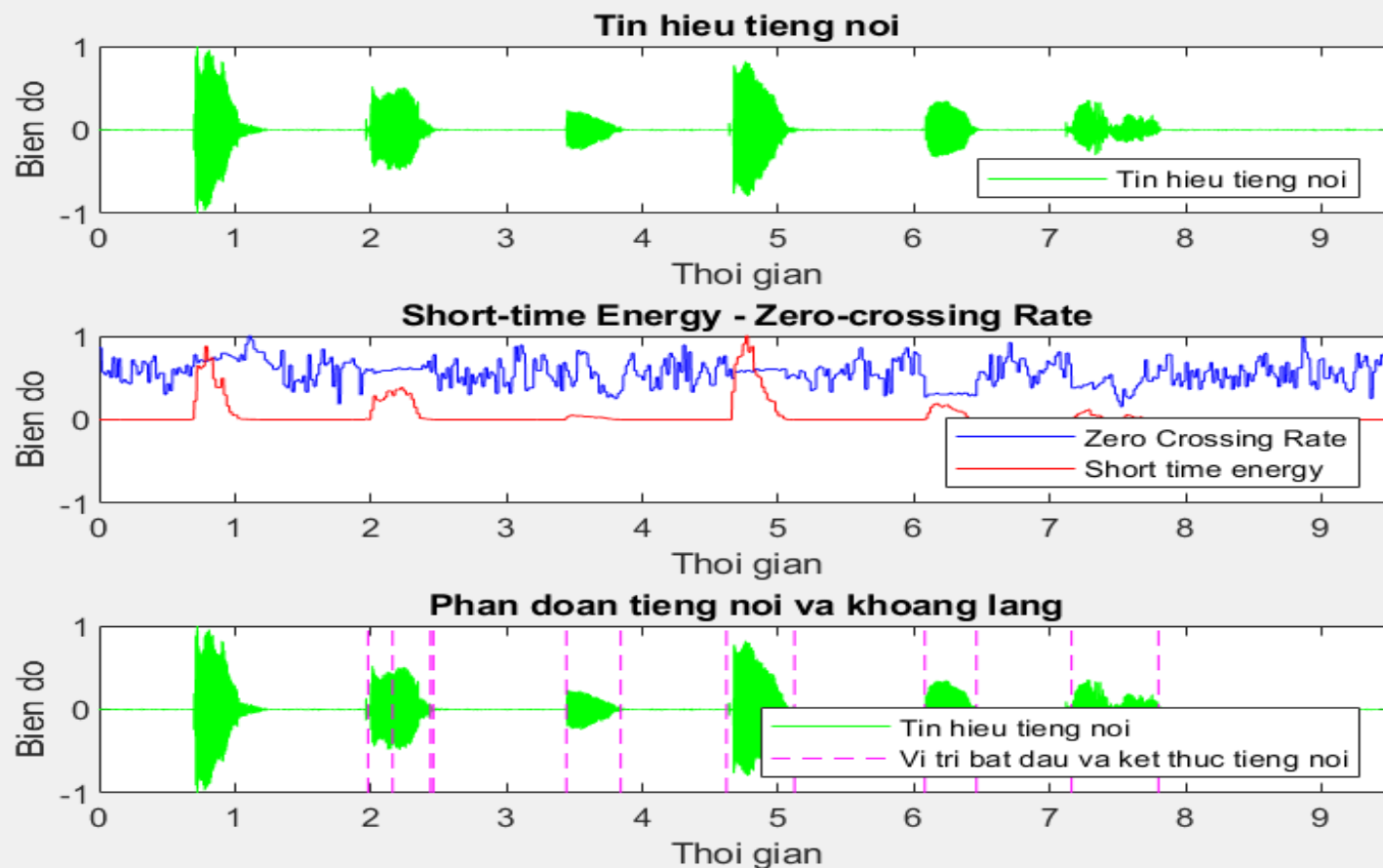
Hình 5. Lab_Male

1. Thuật toán Short-Time Energy



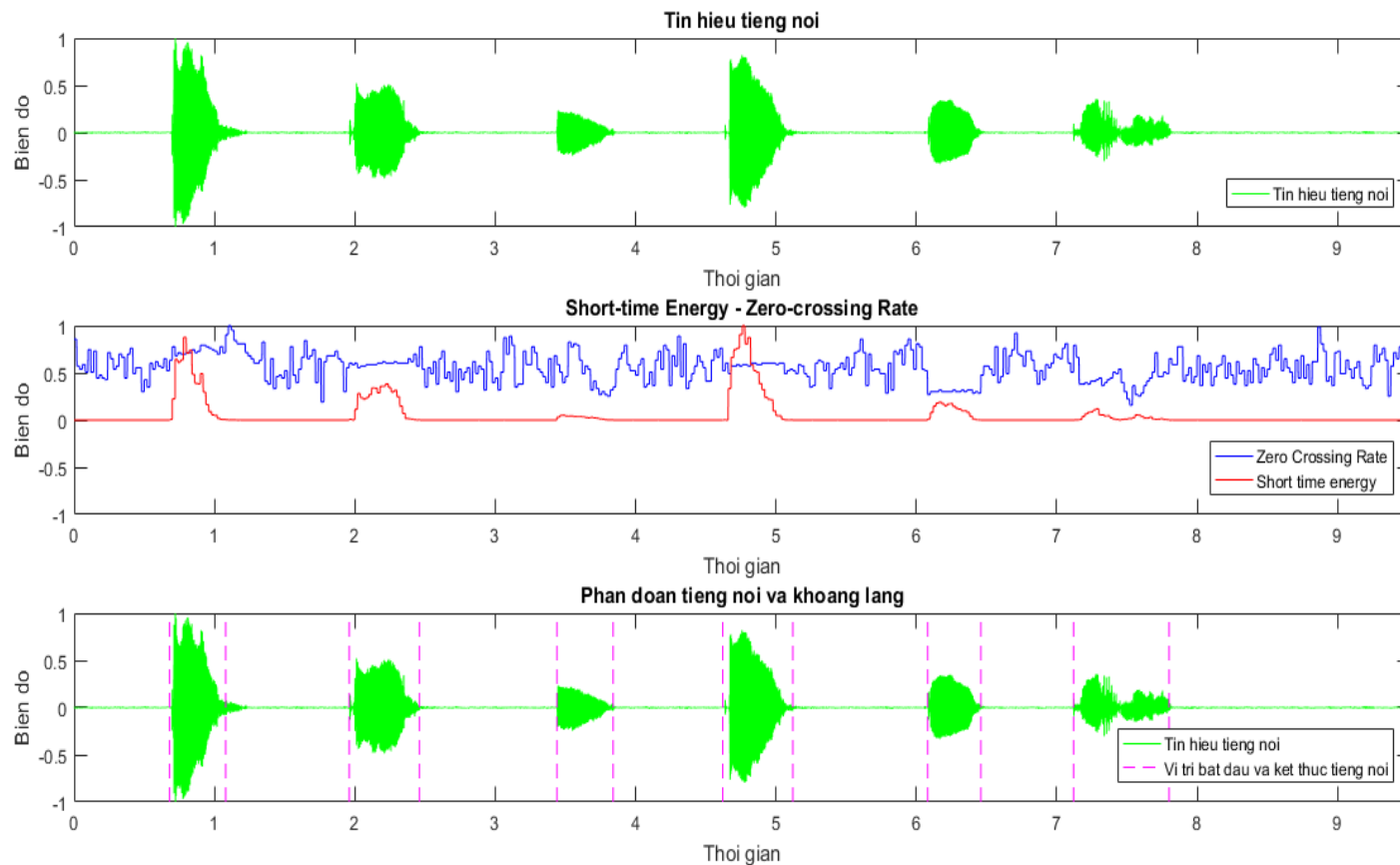
Hình 5. Lab_Female

2. Thuật toán kết hợp Zero Crossing Rate và Short-Time Energy



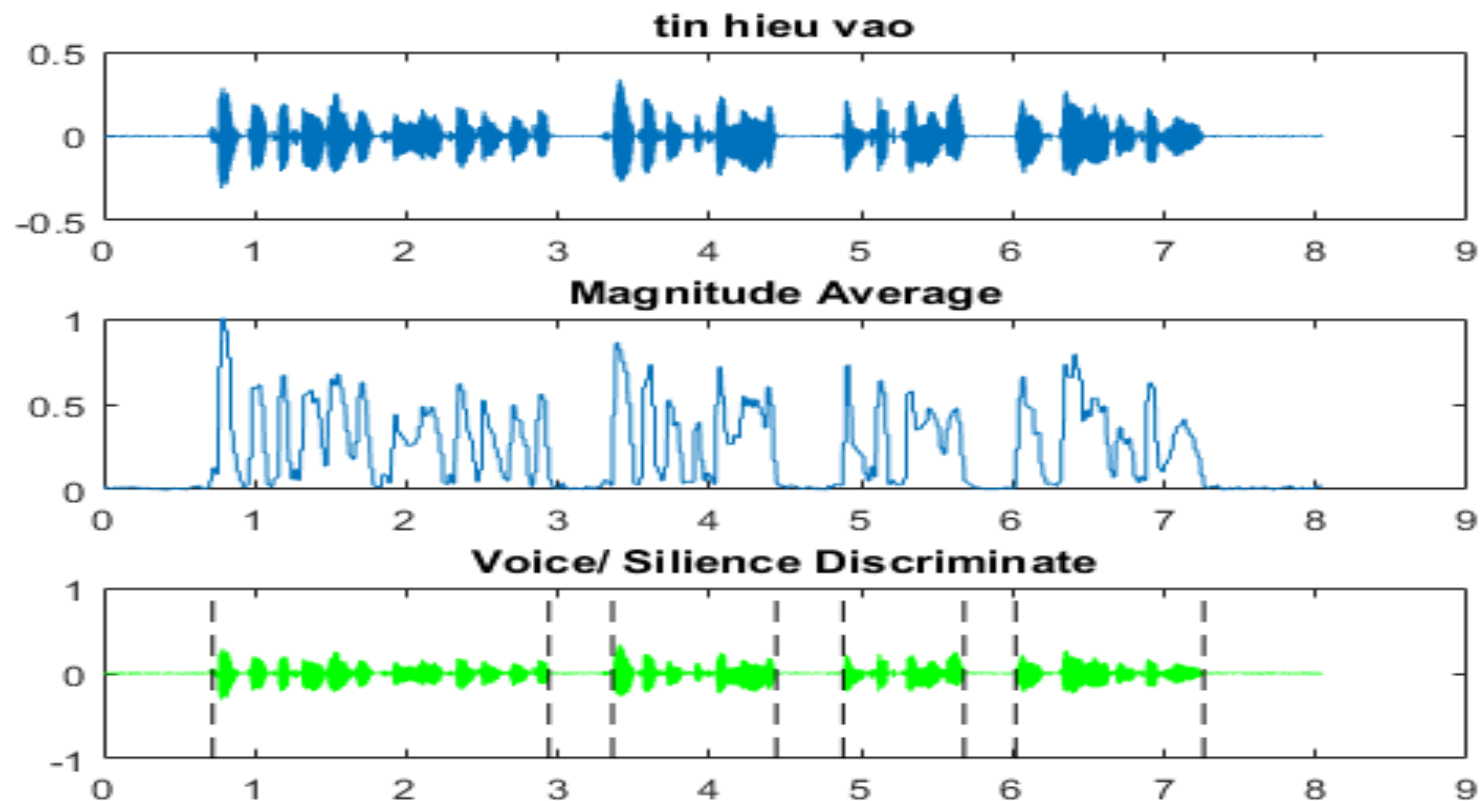
Hình 6. . Kết quả của lab_female khi ZCRct = 0.8 và Ect = 0.001

2. Thuật toán kết hợp Zero Crossing Rate và Short-Time Energy



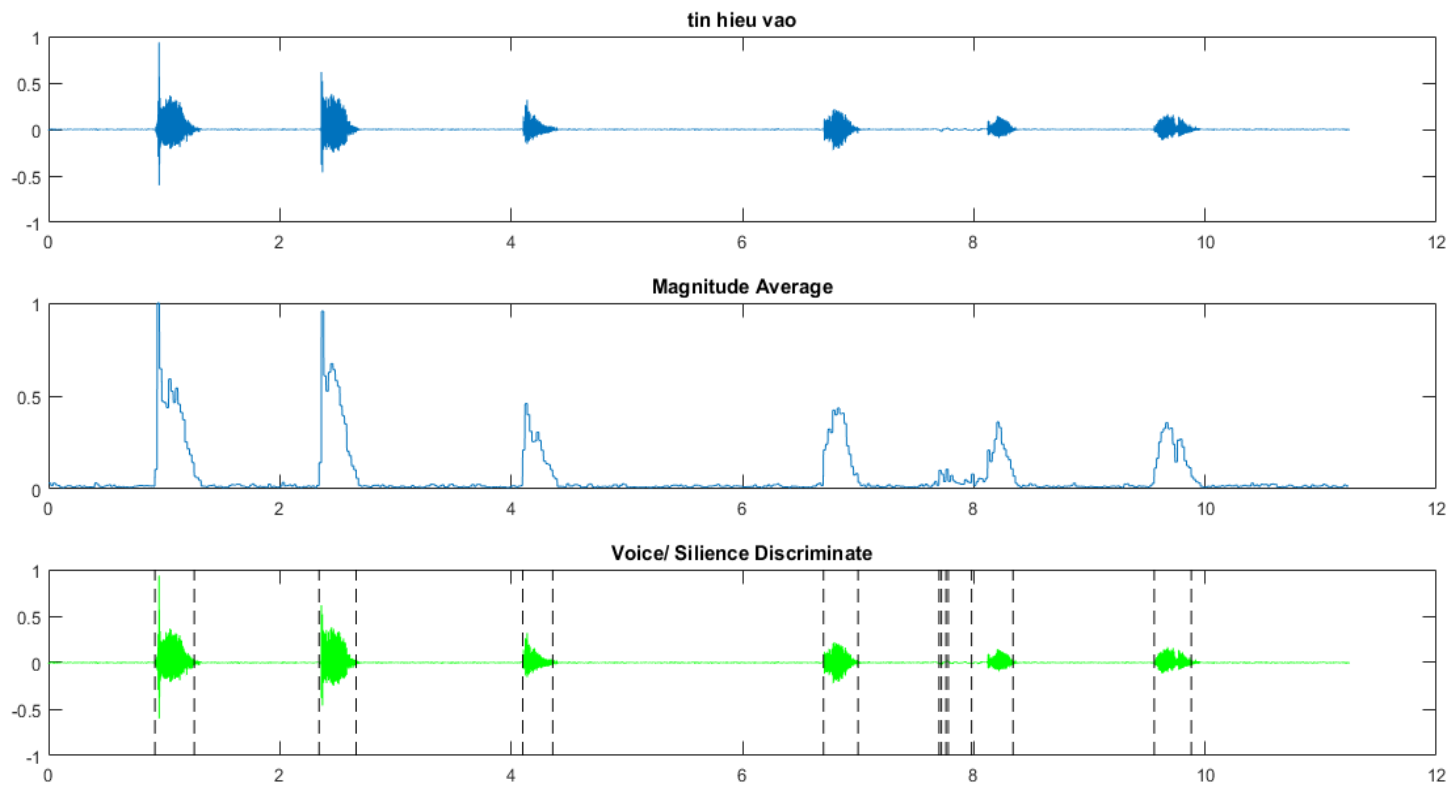
Hình 7. Kết quả của lab_female khi $ZCR_{ct} = 0.6$ và $E_{ct} = 0.001$

3. Thuật toán Magnitude Average



Hình 8. Studio_male

3. Thuật toán Magnitude Average

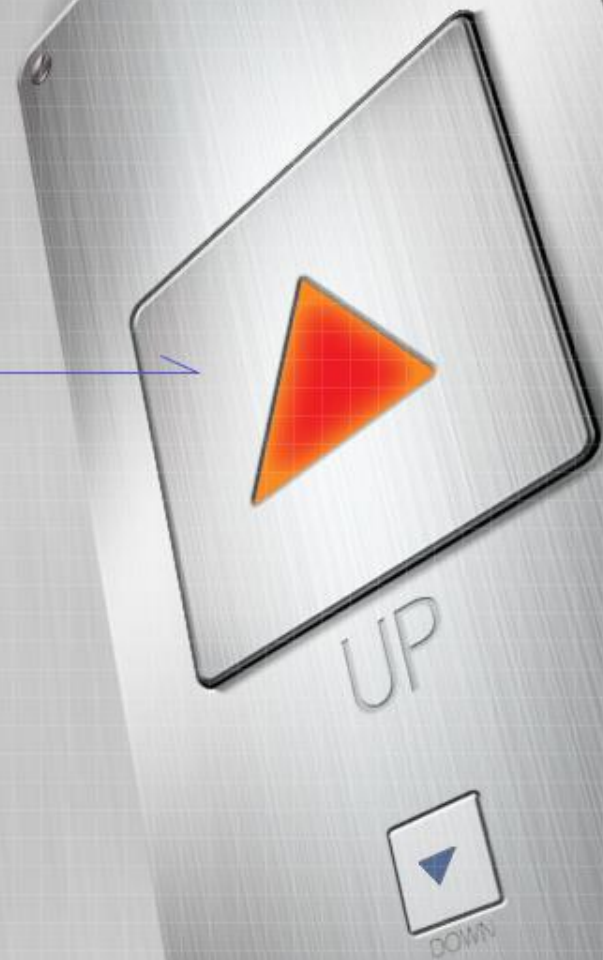


Hình 9. Lab_male

V. KẾT LUẬN

- Thấy được rằng thuật toán dùng STE cho kết quả chính xác hơn thuật toán STE kết hợp với ZCR và thuật toán Magnitude Average.
- Tiếng ồn và tiếng nhiễu,... đã ảnh hưởng đến độ chính xác của thuật toán đạt kết quả không cao.

CẢM ƠN!



LOGO