

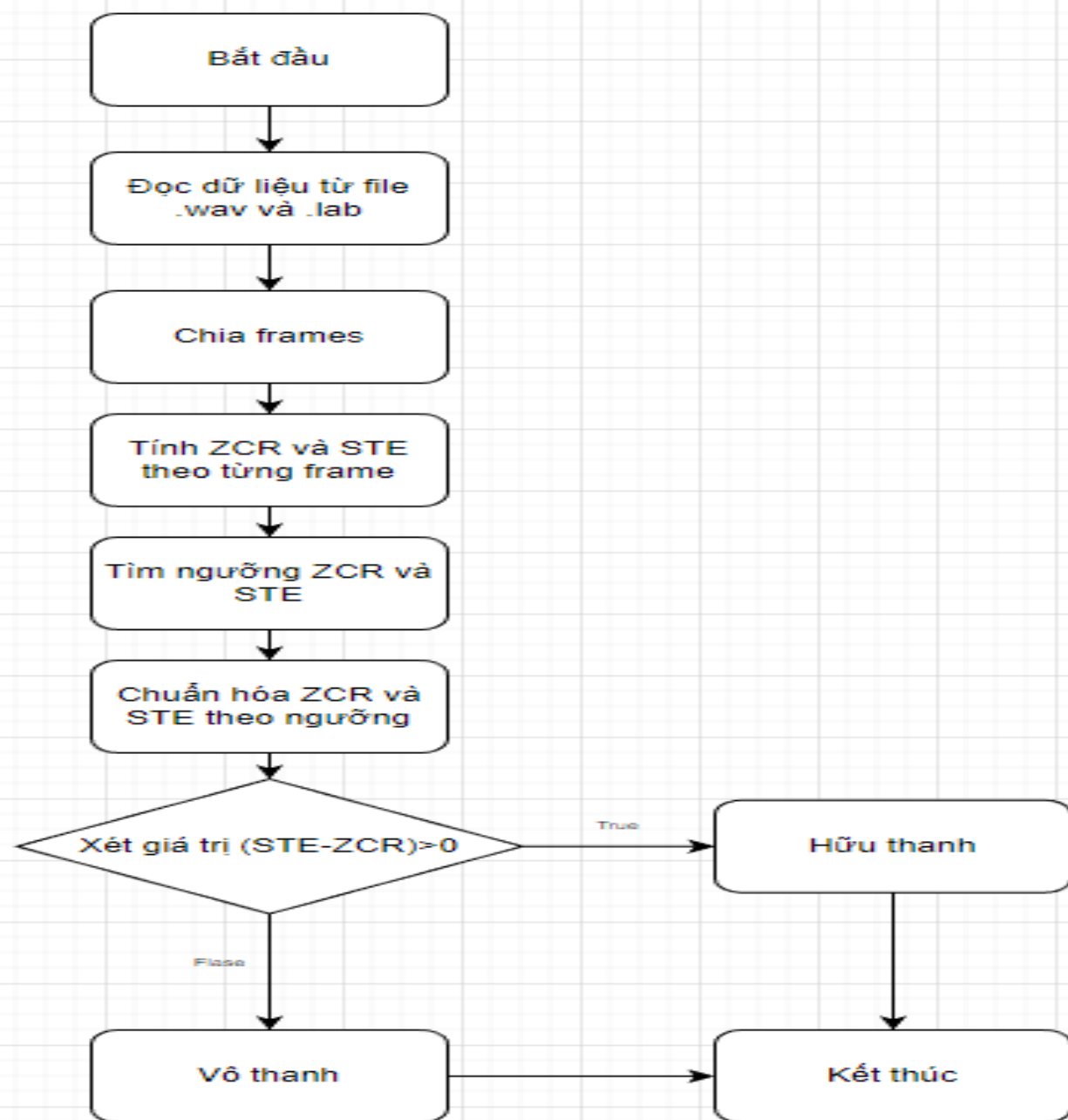
BÁO CÁO XỬ LÝ TÍN HIỆU SỐ

Phân đoạn voice vs unvoice

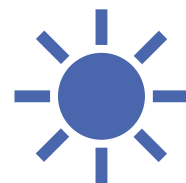
Giáo viên : Ninh Khánh Duy

Sinh viên : Trương Thành Đạt

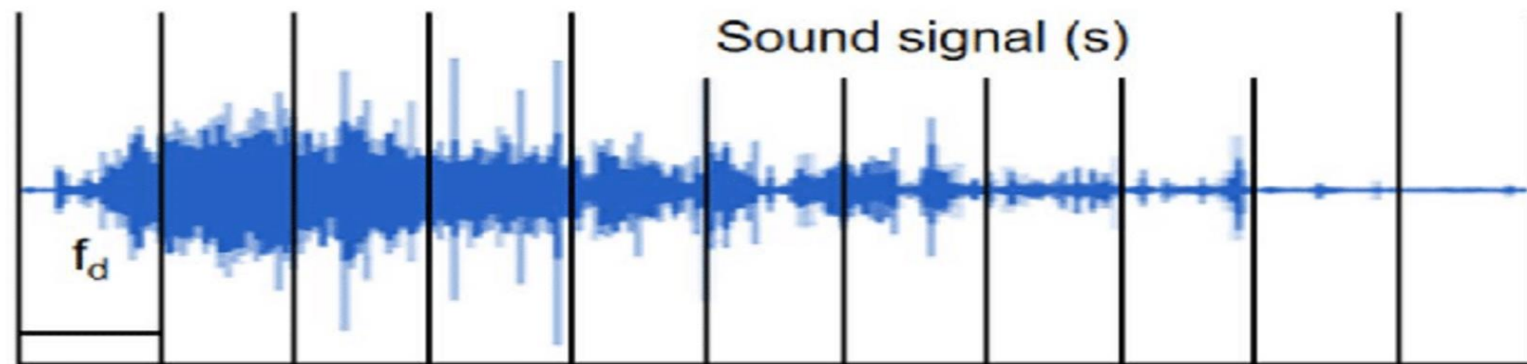
MSSV: 102200205



Đọc dữ liệu vào



- ✓ Đọc tín hiệu vào bằng hàm `waveread()` thu được tín hiệu đầu vào (signal)
- ✓ Đọc file lab tín hiệu
- ✓ Chuẩn hóa Signal đưa về đoạn $[0;1]$
- ✓ Chia dữ liệu đã chuẩn hóa thành nhiều frames , mỗi frame có độ dài 20ms



Tính STE



Tính STE và ZCR theo từng frame ,theo công thức trong CS425 :

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

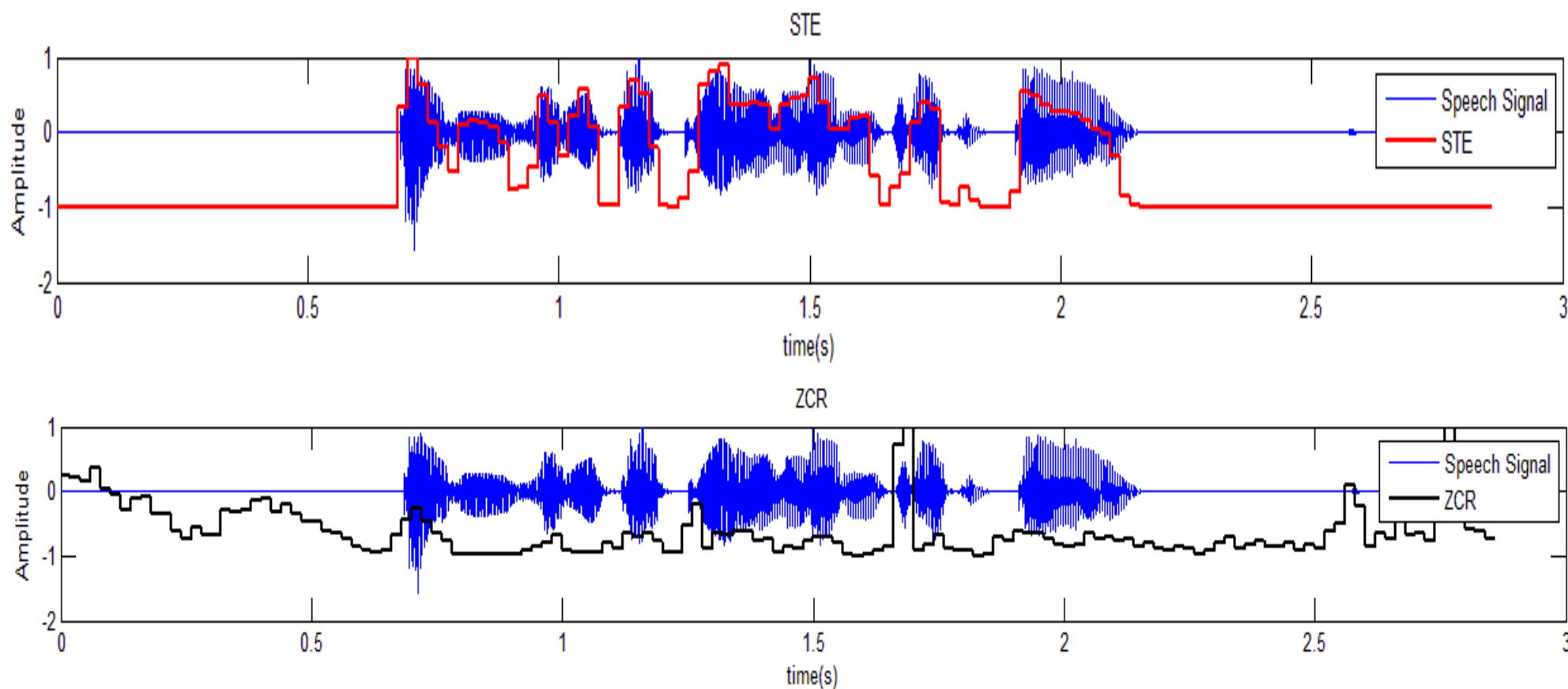
$$\text{ZCR}[n] = \sum_{m=-\infty}^{\infty} y[m]w[n-m] ,$$

Tính ngưỡng STE và ZCR



- ✓ Tính ngưỡng T (threshold) của mỗi tín hiệu huấn luyện ,dựa trên file lab voice và unvoiced theo thuật toán tìm kiếm nhị phân trong tài liệu cs425
- ✓ Với STE và ZCR của 4 tín hiệu huấn luyện , tính trung bình cộng ta được $T_{ste} \approx 1.664$, $T_{zcr} \approx 46.5215$
- ✓ Chuẩn hóa theo ngưỡng $T_{STE} \approx$ và T_{ZCR} tìm được, đưa về đoạn $[-1;1]$

Kết quả file kiểm thử :



Quyết định voice và unvoice

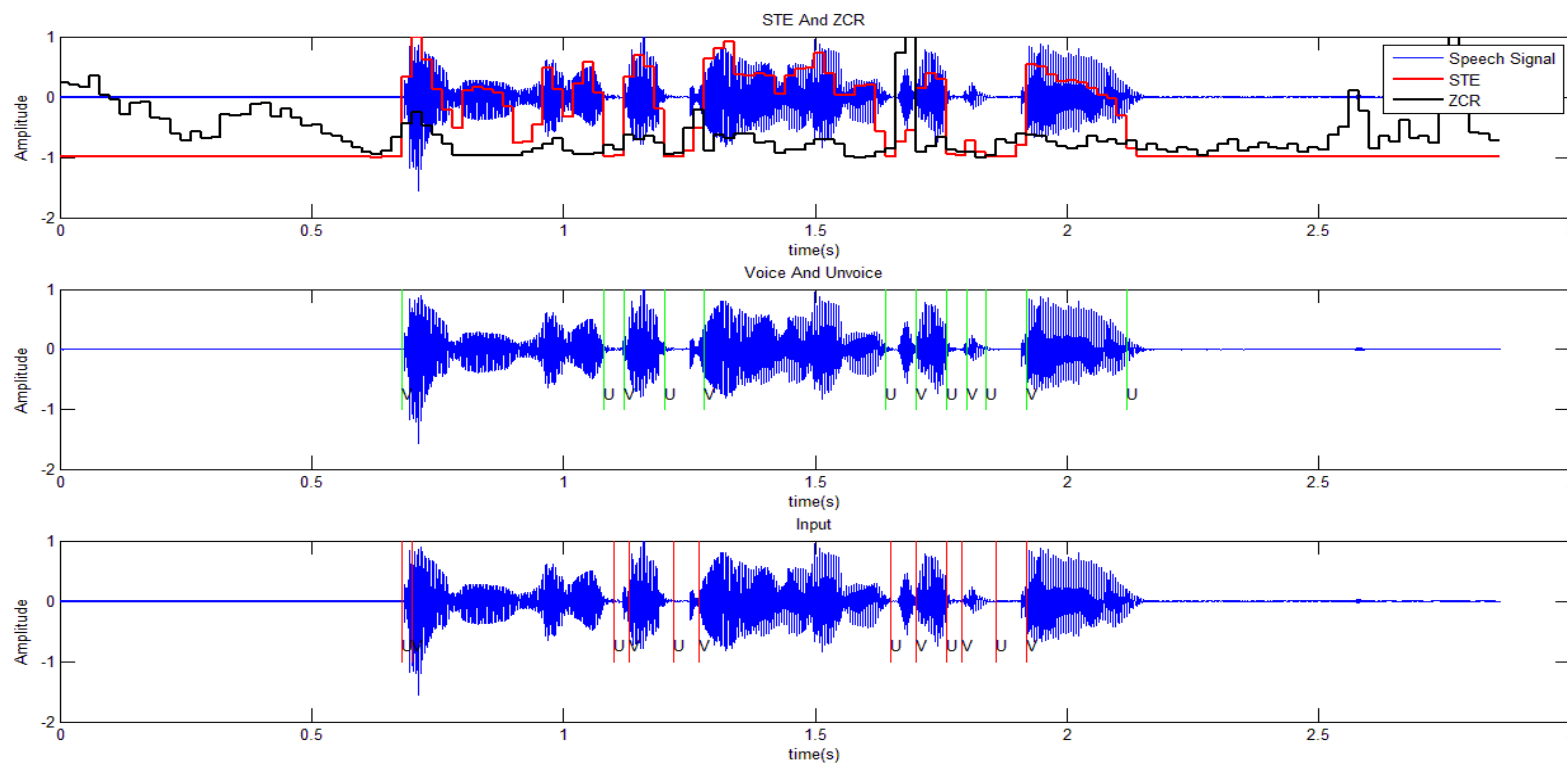


- ✓ $P = \text{ste_wave} - \text{zcr_wave}$; tạo mảng `vauv`
- ✓ Nếu : $P(i) \geq 0$ xác định là voice và lấy giá trị thời gian
 $\text{vauv}(i) = i/\text{fs}$
- ✓ Nếu : $P(i) < 0$ xác định đó là unvoice đánh dấu
 $\text{vauv}(i) = -1$

Kết quả chương trình.



Studio_F1

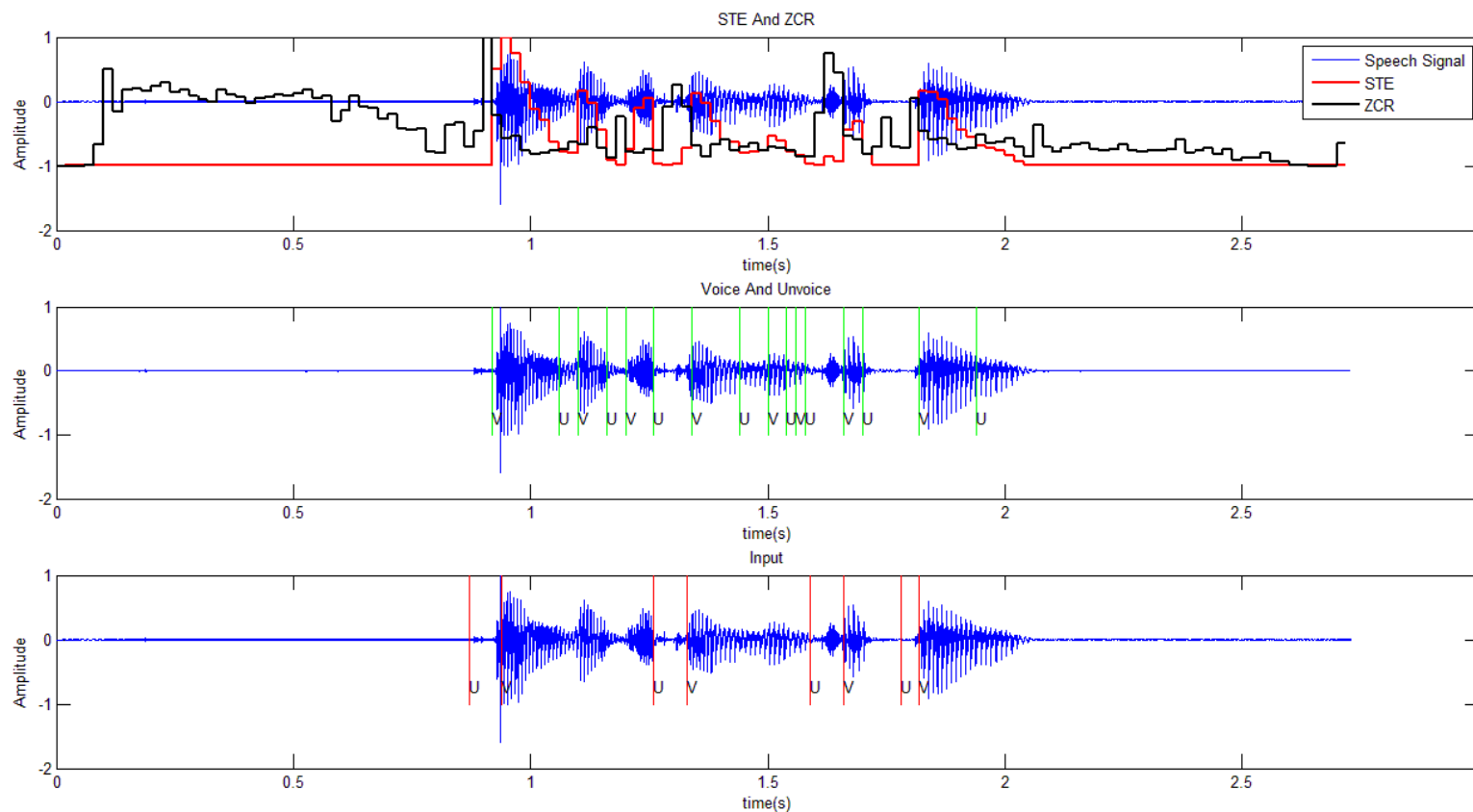


-Độ chính xác so với file của thầy là 90% .
-Sai do độ chính xác của thuật toán chưa hoàn hảo .

Kết quả chương trình.



Studio_M1

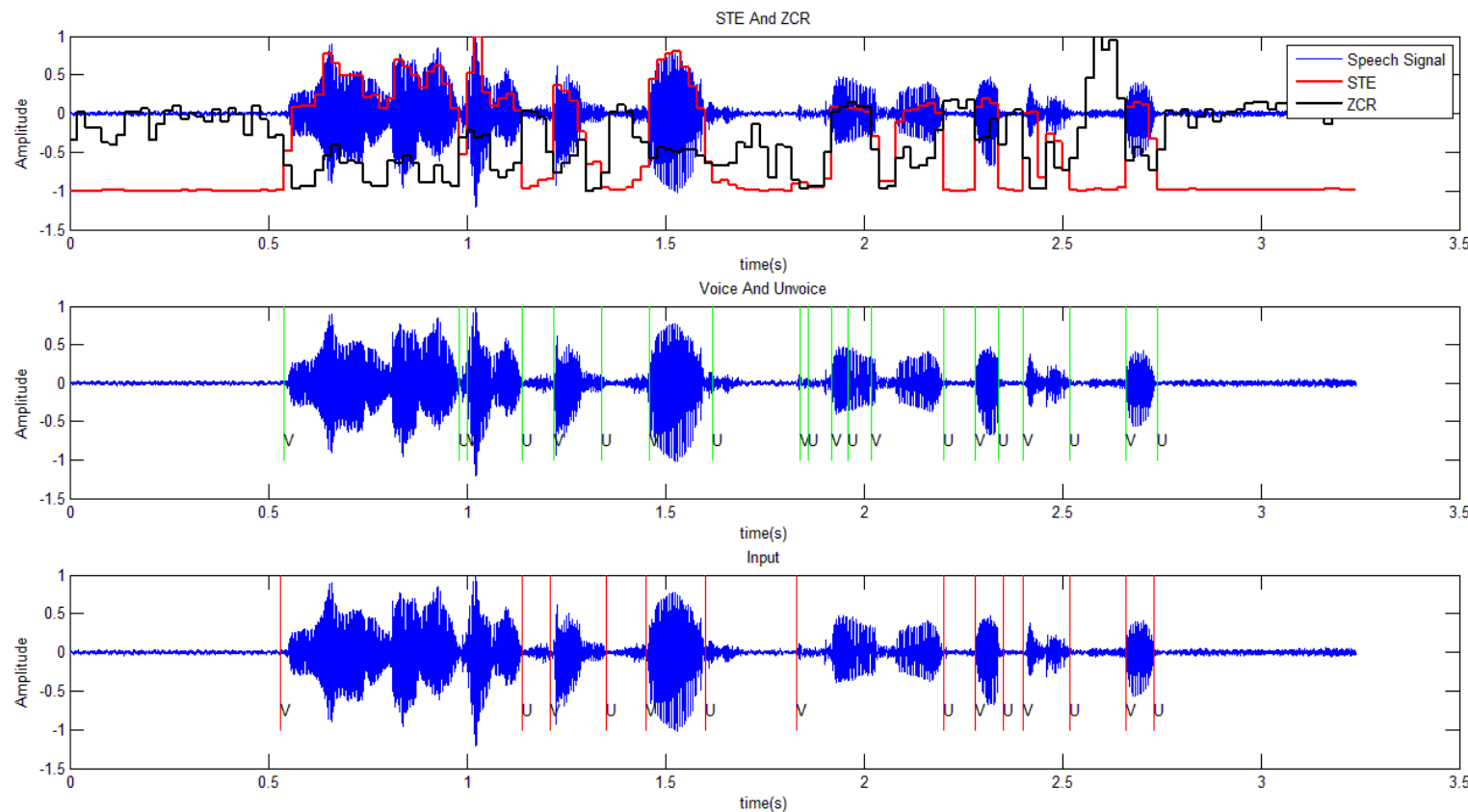


-Giống với file của thầy 70%
-Đánh sai vạch voice , unvoice ở một số vị trí từ giây 1.33- > giây 1.59.
-Người đọc đọc nhanh, lướt , không rõ âm hữu thanh và âm vô thanh

Kết quả chương trình.



Phone_F1



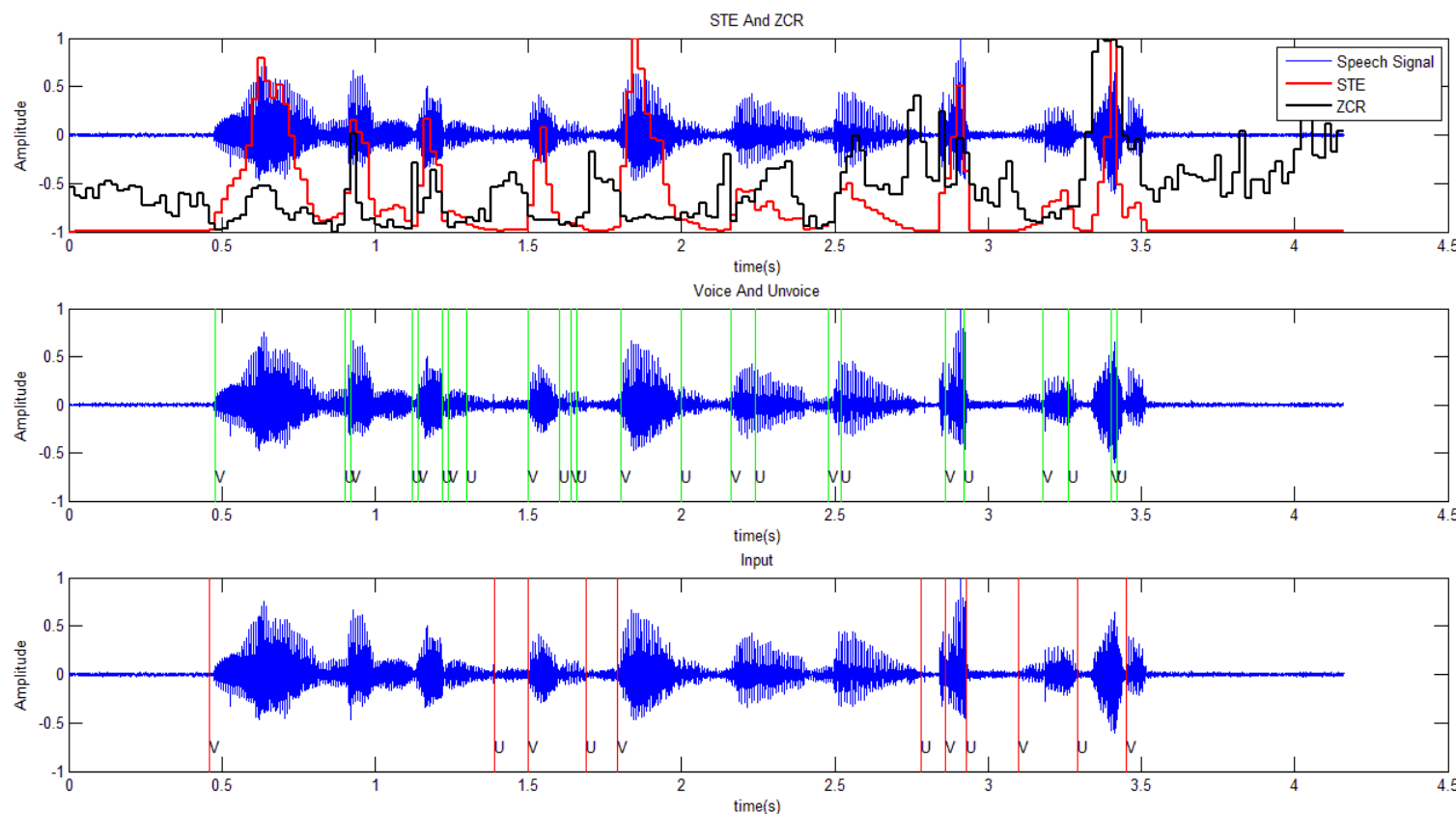
Độ chính xác so với file của thầy là 60%

- Sai giây 1.66
- Một số vị trí bị đánh dấu sai do môi trường thu âm có nhiều nhiễu, do người đọc hơi gió một số vị trí.

Kết quả chương trình.



Phone M1



Độ chính xác so
với file của thầy là
: 40%

Nhiều vị trí bị
đánh dấu sai do
môi trường có
nhiều nhiễu ảnh
hưởng đến giá trị
STE và ZCR dẫn
đến thuật toán sai.

Xin cảm ơn.