

TẦN SỐ CƠ BẢN TRÊN MIỀN THỜI GIAN

Võ Đức Hùng Sơn, Nguyễn Thái Quyên, Nguyễn Duy Nhân, Nguyễn Phú Tân

Nhóm 19, lớp HP : 16.11

Điểm	Bảng phân chia nhiệm vụ		Chữ ký của SV
	Võ Đức Hùng Sơn	Cài đặt thuật toán tính F0 trên miền thời gian bằng hàm tự tương quan (báo cáo từ trang 3-11 trừ lọc trung vị)	
	Nguyễn Thái Quyên	Cửa sổ Hamming (), báo cáo phần Đặt vấn đề, Mục lục, Kết luận, format chung cho báo cáo	
	Nguyễn Duy Nhân	Cài đặt thuật toán lọc trung vị (trang 6), báo cáo phần Bài học đạt được, Tài liệu tham khảo	
	Nguyễn Phú Tân	Cài đặt thuật toán tính F0 trên miền tần số (trang 12-18 trừ phần của sổ Hamming)	

Lời cam đoan: Chúng tôi, gồm các sinh viên có chữ ký ở trên, cam đoan rằng báo cáo này là do chúng tôi tự viết dựa trên các tài liệu tham khảo ghi rõ trong phần VII. Các số liệu thực nghiệm và mã nguồn chương trình nếu không chỉ dẫn nguồn tham khảo đều do chúng tôi tự làm. Nếu vi phạm thì chúng tôi xin chịu trách nhiệm và tuân theo xử lý của giáo viên hướng dẫn.

TÓM TẮT—Báo cáo này báo cáo lại quá trình thực hiện việc cài đặt thuật toán tìm tần số cơ bản của tín hiệu âm thanh với những đoạn ghi âm nguyên âm chính của mỗi sinh viên, trên miền thời gian và trên miền tần số. Sử dụng hàm tự tương quan để cài đặt thuật toán tìm F0 trên miền thời gian. Để làm trơn kết quả tính F0, và thu được F0 mong muốn là giới hạn tần số đồng thời sử dụng lọc trung vị để thu kết quả F0 tốt nhất. Còn với thuật toán tìm tần số cơ bản F0 trên miền tần số bằng cách biến đổi Fourier. Để làm tăng độ chính xác của thuật toán cần lựa chọn tín hiệu tốt không bị nhiễu. Khi tín hiệu đưa vào ổn định thì việc tính F0 sẽ chính xác hơn và sai số cũng nhỏ hơn. Kiểm tra lại kết quả với cửa sổ Hamming.

Từ khóa: Tần số cơ bản, chu kỳ, hàm tự tương quan, lọc trung vị, tín hiệu tiếng nói.

Mục lục

A. ĐẶT VẤN ĐỀ.....2

B. *TÍNH FO TRÊN MIỀN THỜI GIAN*

 I. CƠ SỞ LÝ THUYẾT.....3

 1. Hàm tự tương quan3

 2. Sơ đồ thuật toán3

 3. Các vấn đề phát sinh và hướng khắc phục3

 II. CÀI ĐẶT THUẬT TOÁN.....7

 III. KẾT QUẢ THỰC NGHIỆM.....10

 1. Hình ảnh của kết quả tốt nhất, trung bình, xấu nhất10

 2. Sai số RMSE.....11

 3. Ưu, nhược điểm thuật toán11

 4. Các yếu tố ảnh hưởng đến độ chính xác thuật toán11

C. *TÍNH FO TRÊN MIỀN TẦN SỐ*

 I. CƠ SỞ LÝ THUYẾT.....12

 1. Cơ sở lý thuyết12

 2. Sơ đồ thuật toán12

 3. Các vấn đề phát sinh và hướng khắc phục13

 II. CÀI ĐẶT THUẬT TOÁN.....14

 III. KẾT QUẢ THỰC NGHIỆM.....16

 1. Hình ảnh của kết quả tốt nhất, trung bình, xấu nhất16

 2. Sai số RMSE.....18

 3. Ưu, nhược điểm thuật toán18

 4. Các yếu tố ảnh hưởng đến độ chính xác thuật toán18

D. KẾT LUẬN.....18

E. NHỮNG ĐIỀU ĐÃ HỌC ĐƯỢC.....18

F. TÀI LIỆU THAM KHẢO.....18

A. ĐẶT VẤN ĐỀ

Tiếng nói mà con người giao tiếp hằng ngày có bản chất là sóng âm lan truyền trong không khí, hay gọi cách khoa học là tín hiệu âm thanh tiếng nói là tín hiệu biến thiên liên tục về thời gian. Mỗi một tín hiệu tiếng nói có một dải tần riêng, nó trải rộng từ 20 đến 20000Hz. Và trong khoảng 300 đến 3400Hz là ta có thể nghe rõ nhất. Mỗi một âm do con người phát ra nó mang một tần số cơ bản riêng, đặc trưng cho mỗi tín hiệu. [1]

Tín hiệu tiếng nói là tiếng nói của con người dưới dạng tín hiệu. Tín hiệu tiếng nói thường được thể hiện dưới dạng số, tức là được "số hóa", và do đó xử lý tiếng nói có thể được coi là giao của xử lý tín hiệu số.[1]

Vậy quá trình đi tìm tần số cơ bản của mỗi tín hiệu thực hiện ra sao. Bài báo cáo này sẽ trình bày các cơ sở lý thuyết, các thuật toán, thực nghiệm để tìm F_0 .

Trước hết ta tìm hiểu một số khái niệm để phục vụ cho việc cài đặt thuật toán.

Tần số cơ bản (còn gọi là F_0 hoặc cao độ) của một tín hiệu tuần hoàn bằng nghịch đảo của chu kỳ tín hiệu đó[1], nó đặc trưng cơ bản cho 1 tính chất của tín hiệu, tùy theo mỗi tín hiệu mà có tính chất khác nhau.[2]

Chu kỳ được xác định bằng khoảng thời gian ngắn nhất mà tín hiệu lặp lại trên miền thời gian.[3]

Thuật toán tìm cao độ của tín hiệu dùng hàm tự tương quan là một trong những thuật toán thông dụng, sử dụng miền thời gian để tìm ra chu kỳ của tín hiệu. Thuật toán này tìm giá trị lớn nhất của hàm tự tương quan trên một mẫu tín hiệu xác định. Thuật toán này tìm giá trị lớn nhất của hàm tự tương quan trên một mẫu tín hiệu xác định.

Lọc trung vị là một kỹ thuật lọc số phi tuyến, thường được sử dụng để loại bỏ nhiễu khỏi hình ảnh hoặc tín hiệu. Giảm tiếng ồn như vậy là một bước xử lý trước điển hình để cải thiện kết quả xử lý sau này.[4]

Thuật toán tìm tần số cơ bản trên miền tần số dùng biến đổi Fourier [5] là phép biến đổi một hàm số hoặc một tín hiệu theo miền thời gian sang miền tần số. Biến đổi Fourier thường được nghĩ đến như sự chuyển đổi tín hiệu thành các thành phần biên độ và tần số.

Trong phân tích tín hiệu tiếng nói, thay vì sử dụng trực tiếp tín hiệu tiếng nói trong miền thời gian, người ta thường hay sử dụng các đặc trưng phổ của tiếng nói. Phổ biên độ của tín hiệu tiếng nói được định nghĩa là biến đổi Fourier của khung tín hiệu với khoảng thời gian phân tích cố định.[6]

Bài báo có bố cục như sau. Phần A trình bày phương pháp tìm F_0 trên miền thời gian, Phần B trình bày phương pháp tìm F_0 trên miền tần số. Trong mỗi phần A, B sẽ trình bày về tổng quan về cơ sở lý thuyết, sơ đồ thuật toán (kể cả những vấn đề phát sinh và đề xuất các giải pháp khắc phục), cụ thể các thuật toán và kết quả thực nghiệm. Phần C trình bày bài học đạt được. Cuối cùng là phân tài liệu tham khảo dùng trong báo cáo này

B. TÍNH FO TRÊN MIỀN THỜI GIAN:

I. CƠ SỞ LÝ THUYẾT XỬ LÝ TÍN HIỆU TIẾNG NÓI VÀ CÁC THUẬT TOÁN:

1. Hàm tự tương quan:

Hàm tự tương quan của tín hiệu $x(n)$

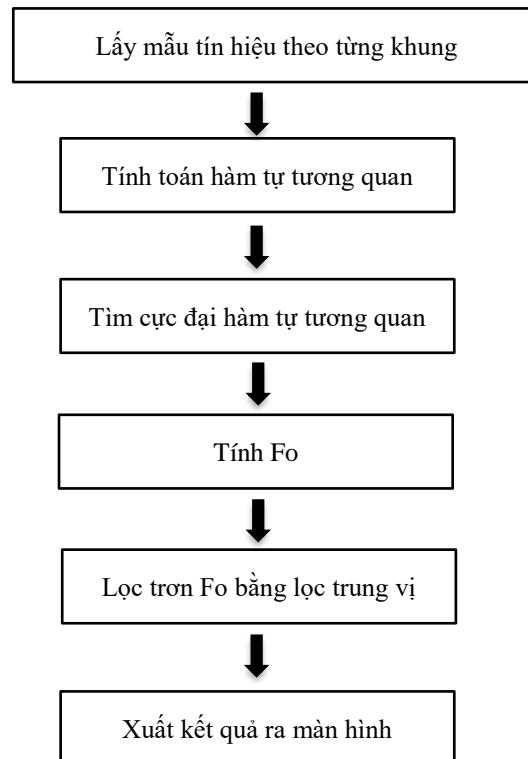
$$r(k) = \sum_{m=-\infty}^{\infty} x(n).x(n+k) \quad [7]$$

- Một số tính chất của hàm tự tương quan:

* $r(k)$ là hàm chẵn

* $r(k)$ đạt cực đại tại những điểm tương ứng là bội số của chu kỳ cơ bản của tín hiệu. [8]

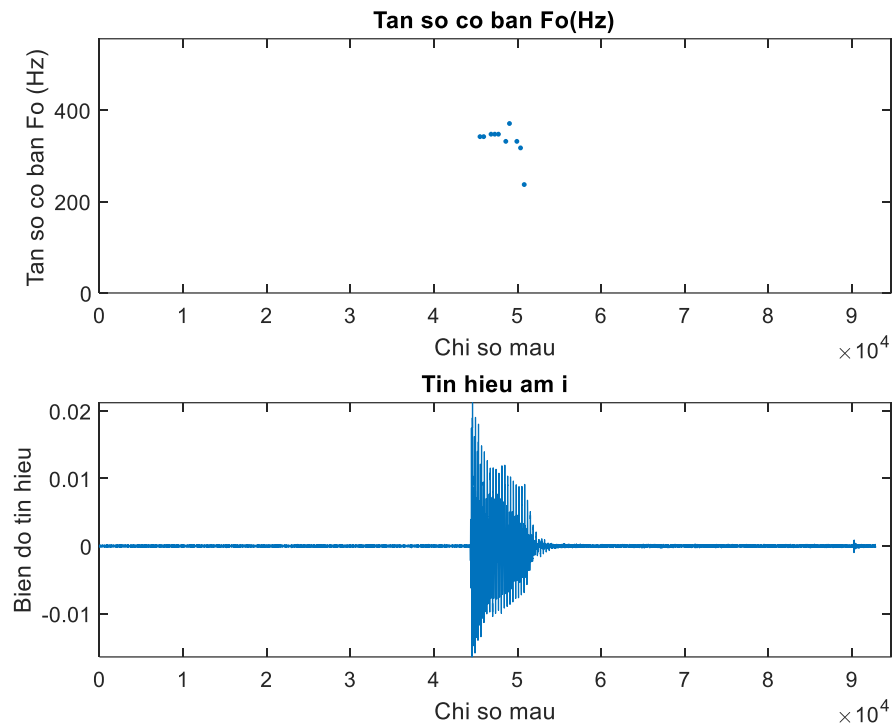
2. Sơ đồ thuật toán:



3. Các vấn đề phát sinh và hướng khắc phục:

3.1. Chọn 2 cực đại kề nhau không chính xác

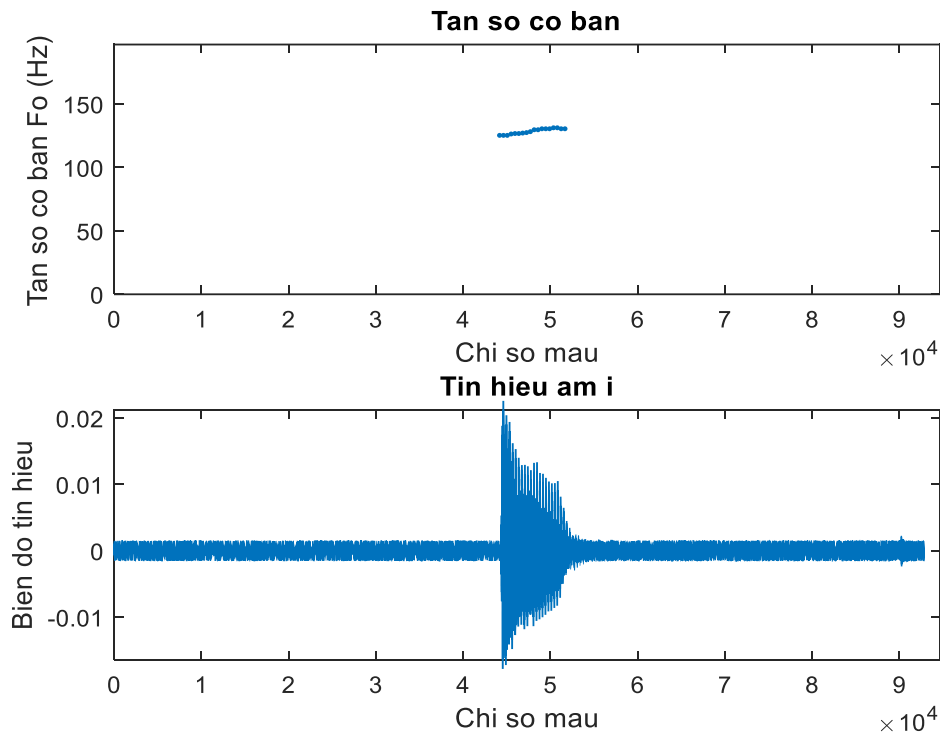
3.1.1. Vấn đề : Hai cực đại được chọn để tính độ trễ không chính xác, trùng cực đại của nhiễu



Hình 1. Kết quả F0 khi chọn sai cực đại

3.1.2. Hướng khắc phục:

Ta sẽ chọn 2 cực đại gồm: max của hàm tự tương quan và 1 cực đại có giá lớn nhất và kề với max chính nhất, khi đó sẽ không chọn nhầm cực đại của nhiễu, độ trễ cần tìm được chính xác.

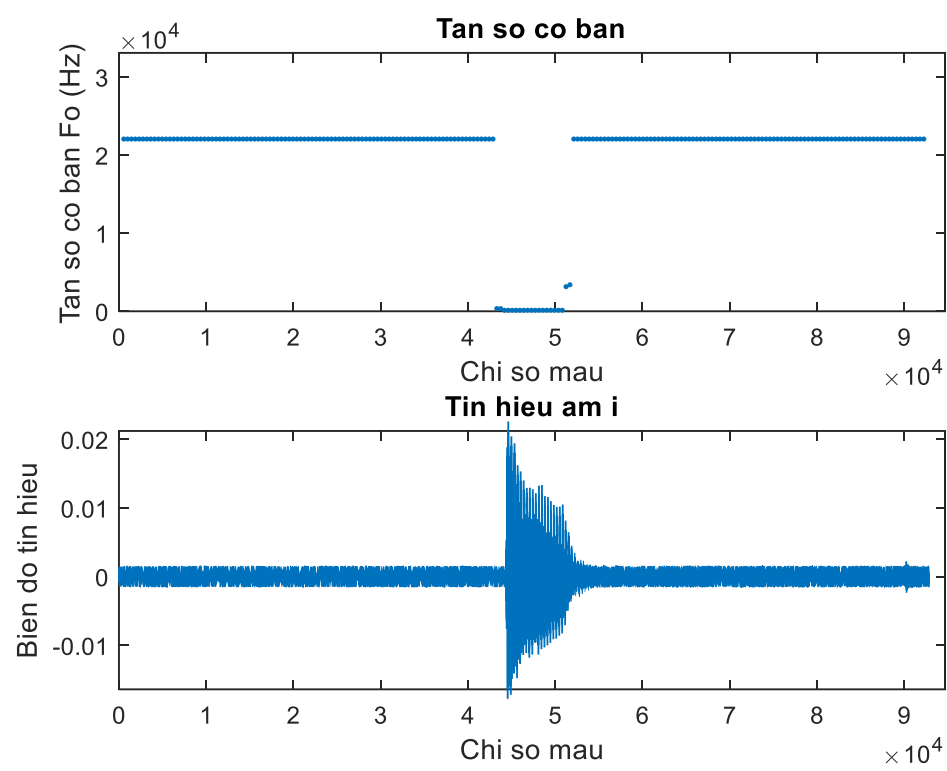


Hình 2. Kết quả F0 khi chọn đúng cực đại

3.2. Xuất hiện nhiều giá trị tần số F0 rất không chính xác:

3.2.1. Vấn đề :

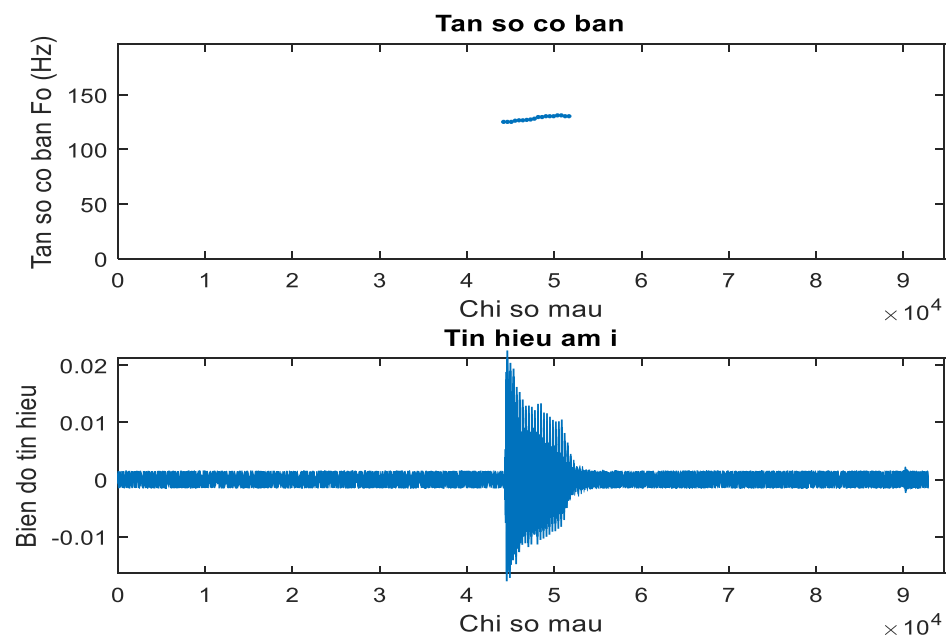
- Xuất hiện những giá trị Fo rất không chính xác, có biên độ rất cao



Hình 3. Kết quả xuất hiện nhiều F0 rất không chính xác.

3.2.2. Hướng khắc phục:

- Giới hạn tần số F0, vì tần số cơ bản của con người dao động từ 80Hz đến 400Hz [9] nên ta chỉ chấp nhận những tần số tính được trong đoạn giới hạn trên .

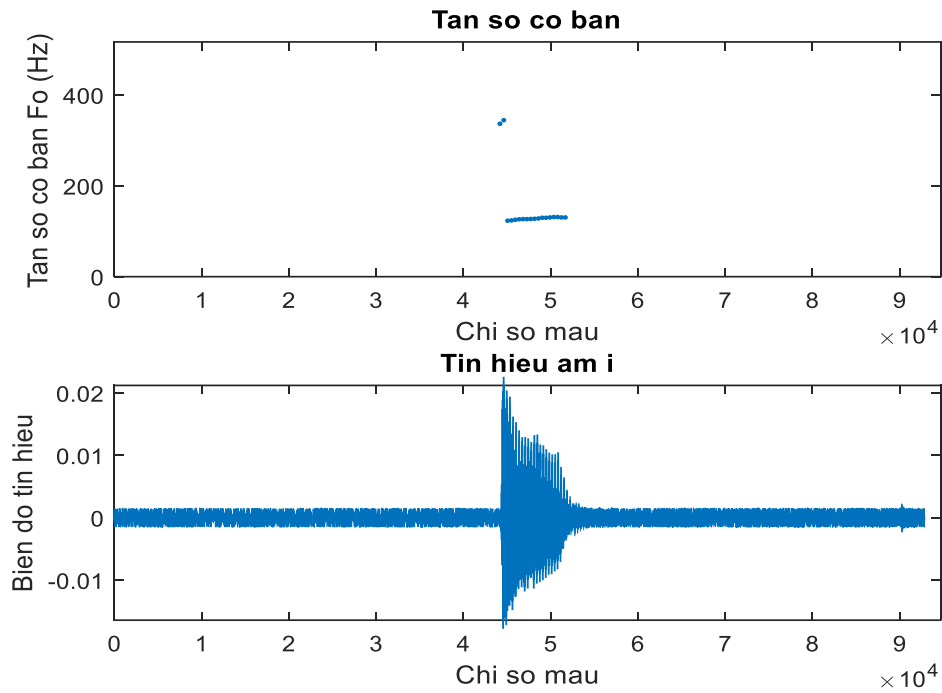


Hình 4. Kết quả F0 sau khi đã giới hạn.

3.3. Giá trị đột biến của Fo (vẫn nằm trong miền giới hạn):

3.3.1. Vấn đề :

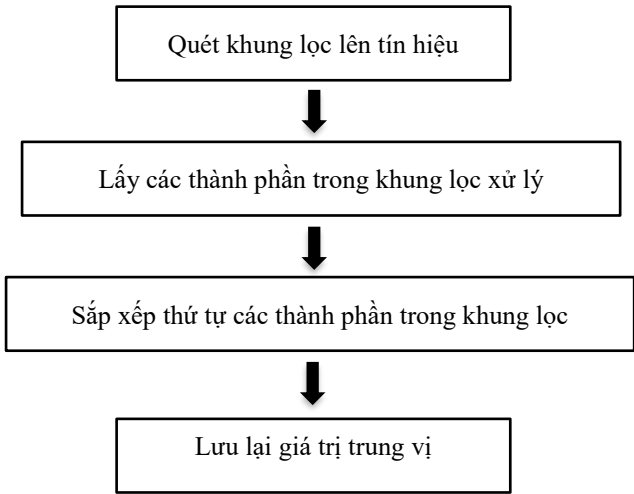
- Kết quả tính ra chứa một vài giá trị Fo cá biệt (quá cao hay thấp so với các giá trị bình thường nhưng vẫn nằm trong miền giới hạn)



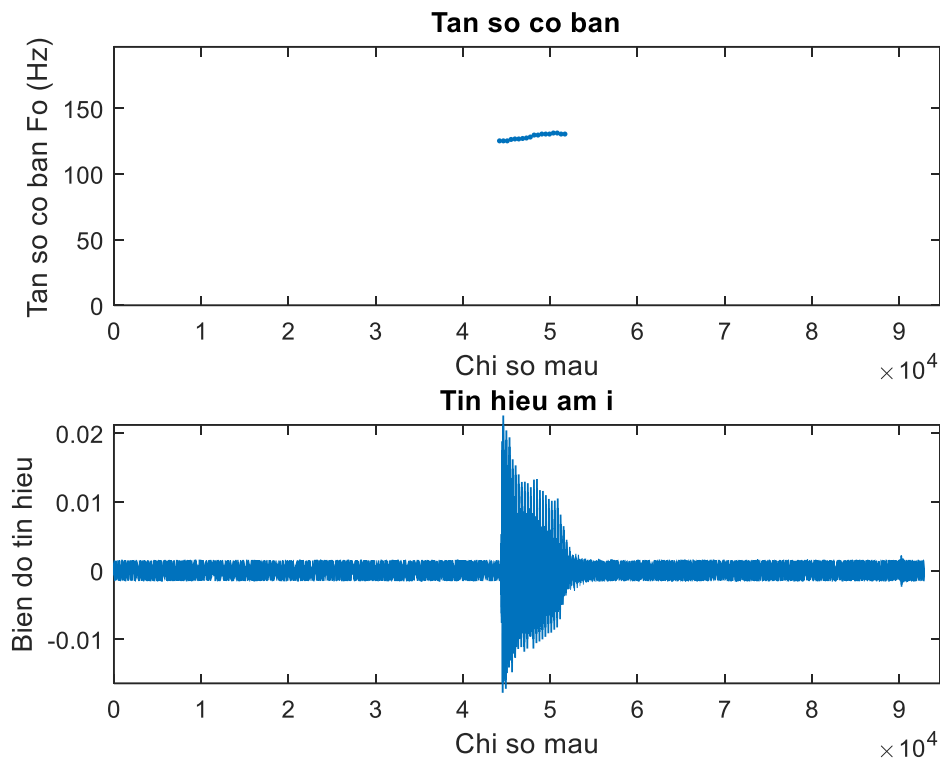
Hình 5. Kết quả F0 khi chưa lọc những đột biến tần số.

3.3.2. Hướng khắc phục:

- Dùng lọc trung vị để loại bỏ các giá trị F0 tăng, giảm đột biến.
- a. Cơ sở lý thuyết:
- Ta sử dụng một khung lọc (kích thước 3x1) quét qua lần lượt từng điểm tín hiệu của tín hiệu đầu vào.
 - Tại vị trí mỗi điểm tín hiệu lấy giá trị của các điểm tín hiệu tương ứng trong vùng 3x1 của tín hiệu gốc “lấp” vào ma trận lọc.
 - Sắp xếp các điểm tín hiệu trong cửa sổ này theo thứ tự (tăng hoặc giảm dần tùy ý).
 - Cuối cùng gán điểm tín hiệu nằm chính giữa (trung vị) cho giá trị điểm tín hiệu đang xét của tín hiệu ra.
- b. Sơ đồ thuật toán:



c. Kết quả



Hình 6. Kết quả F0 khi đã lọc những đột biến tần số.

II. CÀI ĐẶT THUẬT TOÁN:

```
clear();
```

```
% ***** LẤY TÍN HIỆU
```

```
[a, fs] = audioread('..\signal\son/i.wav'); % đọc file ghi âm
```

```
% ***** MỘT SỐ HẰNG SỐ
```

```
a_len = length(a); % chiều dài tín hiệu
```

```
frame_len = floor(0.025 * fs); % chiều dài mỗi khung, mỗi khung dài 25ms (thời gian mà đặc tính giọng nói không đổi)
```

```
number_frame = floor((a_len-frame_len)/(0.4*frame_len))+1; % số lượng khung, overlap 60% (15ms so với 25ms)
```

```
index = 1; % chỉ số của mảng lưu fo từng khung
```

```
number_filter = 5; % số lượng mẫu trong 1 khung lọc
```

```
max_period = floor(fs/80); % max số lượng mẫu trong 1 chu kì cơ bản
```

```
min_period = floor(fs/400); % min số lượng mẫu trong 1 chu kì cơ bản
```

```
% ***** VỀ TÍN HIỆU VÀO
```

```
subplot(2,1,2);
```

```
plot(a);
```

```
title('Tín hiệu âm i');
```

```
xlabel('Chỉ số mẫu');
```

```
ylabel('Biên độ tín hiệu');
```

```
axis([0 a_len*1.02 -inf inf]);
```

```
% ***** DUYỆT VÀ TÍNH FO TRONG TỪNG KHUNG
```

```
for k = 1 : number_frame
```

```
    % lấy mẫu tín hiệu theo từng khung
```

```
    % tính toán hàm tự tương quan
```



```

% tìm cực đại hàm tự tương quan và tính fo
% lưu fo vào mảng kết quả, lưu chỉ số mẫu chính giữa của khung, tăng chỉ số của mảng kết quả

    frame = getFrame(k, a, frame_len);
    fr_xcorr = xcorr(frame, frame);
    [fo, loc] = getFo(k, frame_len, fr_xcorr, min_period, max_period, fs);

    if(fo ~= 0)
        fo_vecto(index) = fo; % lưu Fo của khung thứ k vào mảng kết quả
        loc_vecto(index) = loc; %
        index = index + 1;
    end
end

% ***** LỌC VÀ XUẤT KẾT QUẢ
if(fo_vecto ~= 0) % kiểm tra mảng fo_vecto có tồn tại

    % làm trơn mảng kết quả fo bằng lọc trung vị
    % xuất kết quả ra màn hình

    fo_vecto = MyFunctionFilter(fo_vecto, number_filter);

    subplot(2,1,1);
    plot(loc_vecto, fo_vecto, '.');
    title('tần số cơ bản');
    axis([0 a_len*1.02 0 max(fo_vecto)*1.5]);
    xlabel('chỉ số mẫu');
    ylabel('tần số cơ bản fo (Hz)');
end

```

TRONG ĐÓ:

- ***getFrame(k, a, frame_len)***: hàm lấy khung thứ k với độ dài là frame_len của tín hiệu a

```

function [frame] = getFrame(k, a, frame_len)
%***** Tham số
    % k là vị trí khung cần lấy
    % a là tín hiệu vào
    % frame_len là chiều dài mỗi khung
%***** Kết quả trả về
    % frame: khung thứ k của tín hiệu
%***** Code
    % xác định vị trí khung k, nạp chồng 60% (15ms so với 25ms)
    range = floor((k - 1)*0.4*frame_len) + 1 : floor((k - 1)*0.4*frame_len) + frame_len;
    % lấy khung thứ k
    frame = a(range);
end

```

- ***getFo(k, frame_len, fr_xcorr, min_period, max_period, fs)***: hàm tính Fo của khung thứ k

```

function [fo, loc] = getFo(k, frame_len, fr_xcorr, min_period, max_period, fs)
%***** Tham số
    %k là vị trí của khung
    %frame_len là chiều dài mỗi khung

```

```

%fr_xcorr là hàm tự tương quan của khung thứ k
%min_period là số lượng mẫu nhỏ nhất trong 1 chu kì cơ bản
%max_period là số lượng mẫu lớn nhất trong 1 chu kì cơ bản
%***** Kết quả trả về

% fo là tần số cơ bản của khung thứ k
% loc chỉ số mẫu chính giữa của khung thứ k

%***** Code
fo = 0;
loc = 0;

% Tìm cực đại cục bộ hàm tự tương quan
[pks,lag] = findpeaks(fr_xcorr);

% Tính Fo
if(~isempty(pks))

    % Tìm độ trễ của 2 cực đại kề nhau của hàm tự tương quan
    % Hai cực đại gồm: max hàm tự tương quan và cực đại lớn nhất và kề max nhất

    first_loc_max = find(pks == max(pks)); % tìm chỉ số max hàm tự tương quan
    if(first_loc_max > 1) % kiểm tra chỉ số của max có lớn hơn 1

        second_loc_max = 1; % chỉ số của max nửa đầu (kề max chính nhất) hàm tự tương quan
        second_max = pks(first_loc_max - 1); % giá trị max nửa đầu hàm tự tương quan kề max
        chính nhất
        % Tìm max và chỉ số max của nửa đầu hàm tự tương quan
        % Tính độ trễ của 2 cực đại đã tìm
        % Tính fo, chỉ số (kiểm tra delta có thỏa điều kiện max, min số lượng mẫu trong 1 chu kì)

        for i = 1 : first_loc_max - 1
            if(pks(i) >= second_max)
                second_loc_max = i;
                second_max = pks(i);
            end
        end

        delta = lag(first_loc_max) - lag(second_loc_max);

        if(delta >= min_period && delta <= max_period)
            fo = fs/delta;
            loc = floor((k-1)*0.4*frame_len + frame_len/2);
        end
    end
end
end

• MyFunctionFilter(fo_vecto, number_filter): loại bỏ những giá trị Fo tăng, giảm đột biến trong mảng kết quả

function y = FunctionMedianFilter(y,m) % y tín hiệu cần lọc, m là kích thước của khung lọc
khung = 1:m;
for j = 1:length(y)+1-m % quét khung lọc lên tín hiệu vào

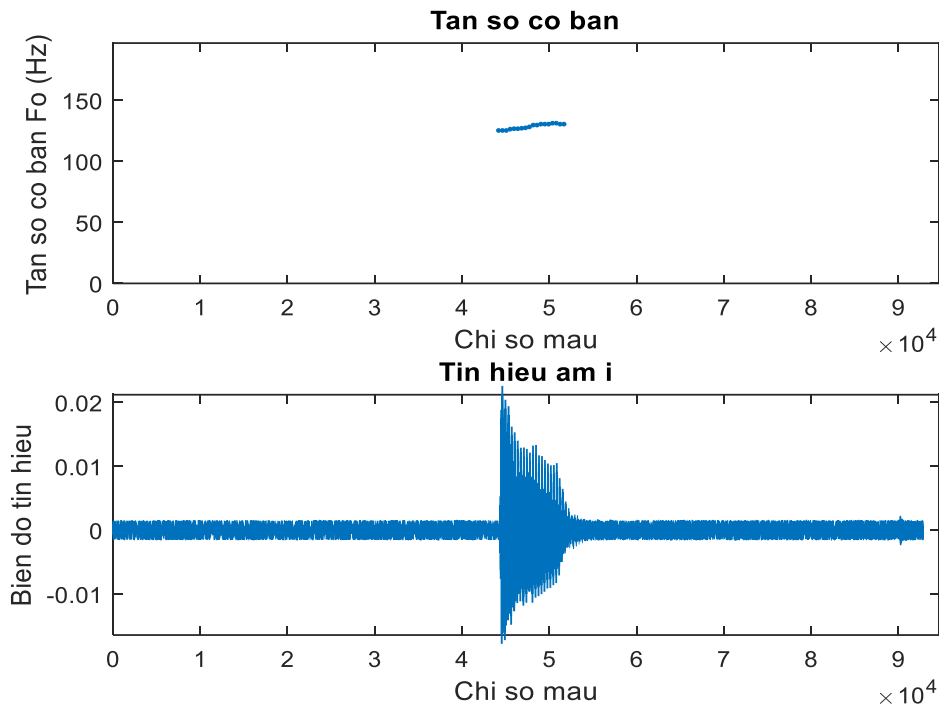
```

```
for i = 1:m      % gán các điểm giá trị của tín hiệu vào khung lọc
    khung(i) = y(j-1+i);
end
khung = sort(khung); % sắp xếp các điểm giá trị trong khung lọc
y(j) = khung((length(khung)+1)/2); % gán giá trị trung vị cho tín hiệu ra
end
end
```

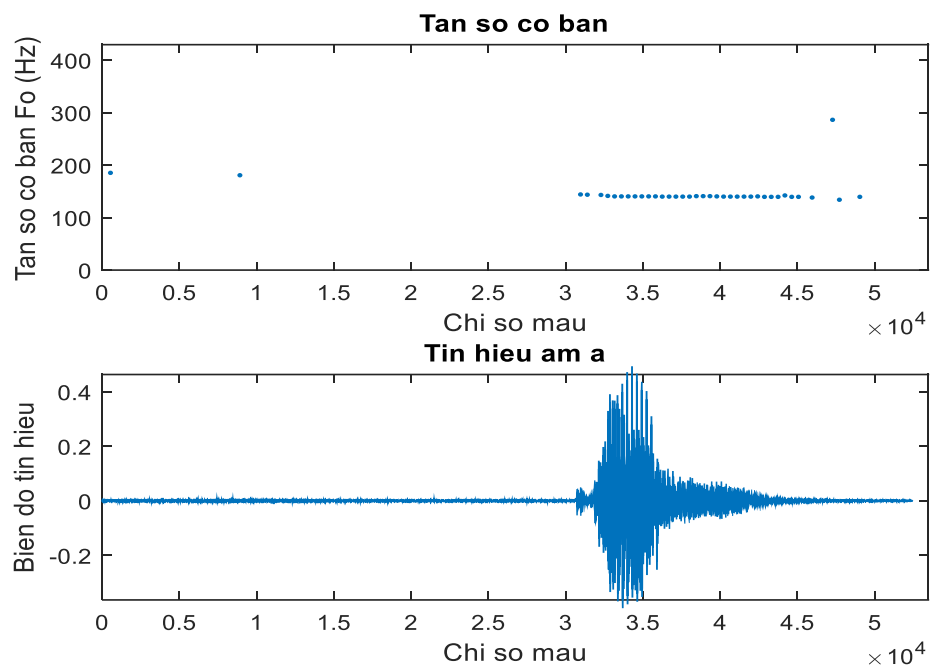
- *xcorr(x,x): tìm hàm tự tương quan tín hiệu $x(n)$ – hàm có sẵn trong Matlab*
- *findpeaks(x): tìm cực đại của tín hiệu $x(n)$ – hàm có sẵn trong Matlab*
- *max(x): tìm giá trị max của $x(n)$ – hàm có sẵn trong Matlab*
- *find(x = k): trả về vị trí trong x mà tại đó x có giá trị bằng k – hàm có sẵn trong Matlab*

III. KẾT QUẢ THỰC NGHIỆM:

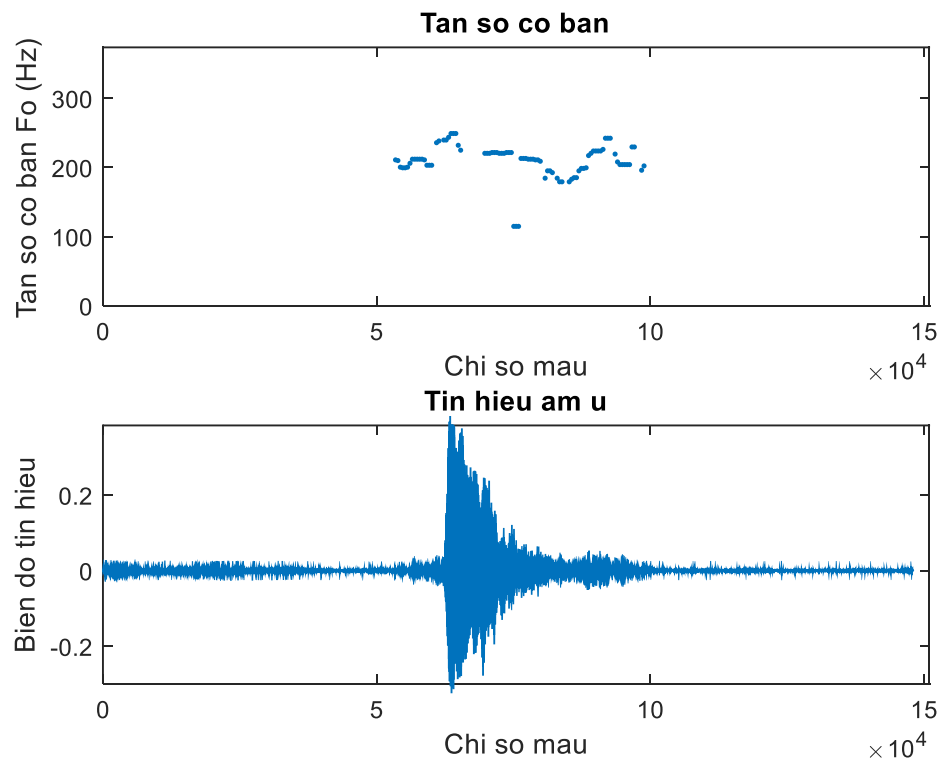
1. Hình ảnh tốt nhất, trung bình và xấu nhất khi tính F0 với thuật toán trên :



Hình 7. Kết quả tính F0 tốt nhất.



Hình 8. Kết quả tính F_0 trung bình



Hình 9. Kết quả tính F_0 xấu nhất.

2. Sai số RMSE

Bảng 1. Bảng sai số RMSE của 20 âm

	a		e		i		o		u	
	Fo chuẩn	Sai số	Fo chuẩn	Sai số	Fo chuẩn	Sai số	Fo chuẩn	Sai số	Fo chuẩn	Sai số
Son	146.4	24.9	142.86	0.8	128.48	2.14	125.5	0.56	145.6	4.22
Nhân	158.85	3.37	147.06	2.86	156.25	1.89	153.85	25.72	149.25	7.12
Quyên	235.83	18.07	230.89	8	233.33	4.4	234.57	15.27	252	20.19
Tân	137.7	3.26	142.2	4.03	152	5.67	147.1	4.7	158	9.8

3.Ưu điểm, nhược điểm của thuật toán:

- a) Ưu điểm: tính được Fo trên miền thời gian, tương đối chính xác khi tín hiệu ít nhiễu
- b) Nhược điểm: phụ thuộc vào tín hiệu vào, gây sai số lớn khi tín hiệu nhiễu nhiều

4. Các yếu tố ảnh hưởng đến kết quả tính Fo:

- Tín hiệu vào ít nhiễu thì sẽ cho kết quả tính Fo chính xác hơn
- Độ dài mỗi khung nằm trong khoảng 25ms – 30ms sẽ cho kết quả chính xác hơn (thời gian mà được tính giọng nói không đổi)
- Giới hạn đúng miền giá trị của Fo, để lấy những giá trị được chính xác hơn

C. TÍNH FO TRÊN MIỀN TẦN SỐ

I. CƠ SỞ LÝ THUYẾT

1. Cơ sở lý thuyết

- a) Phép biến đổi Fourier DTFT (tín hiệu rời rạc vô hạn và không tuần hoàn)
- Biểu diễn tín hiệu rời rạc bằng cách lấy mẫu tín hiệu tương tự dưới dạng:

$$x_s(t) = \sum_{k=-\infty}^{\infty} x(kT)\delta(t-kT)$$

- Biến đổi Fourier cho tín hiệu:

+ Tính biến đổi Fourier của $\delta(t - kt)$.

+ Sử dụng nguyên lý xếp chồng, tìm biến đổi Fourier của $(x_s(t))$

$$x_s(t) \leftrightarrow \sum_{n=-\infty}^{\infty} x(nT)e^{-jn\omega T}$$

Đặt $x(nT)=x[n]$ và thay biến $\Omega = \omega T$ (lưu ý đơn vị của Ω [rad] và ω [rad/s]), ta được:

$$\text{DTFT} : X(\Omega) = \sum_{n=-\infty}^{\infty} x[n]e^{-j\Omega n}$$

- Tín hiệu rời rạc trong miền thời gian nhưng DTFT lại liên tục và tuần hoàn trong miền tần số.
- DTFT chính là hàm phức theo biến tần số thực. Ta gọi DTFT là phổ của tín hiệu rời rạc.
- Phổ của tín hiệu gồm có hai phần: phổ biên độ (magnitude spectrum) và phổ pha (phase spectrum).

$$X(\Omega) = |X(\Omega)|e^{j\theta(\Omega)}$$

Phổ biên độ: $|X(\Omega)|$

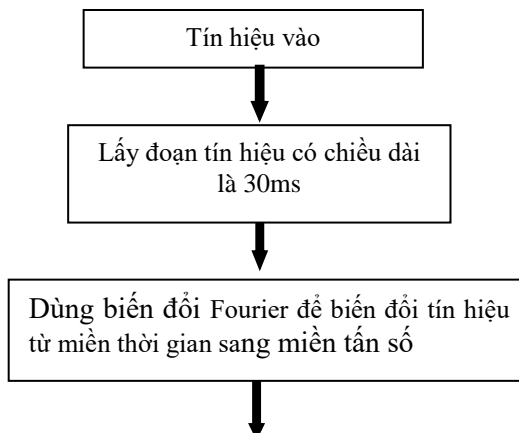
Phổ pha: $\theta(\Omega)$.

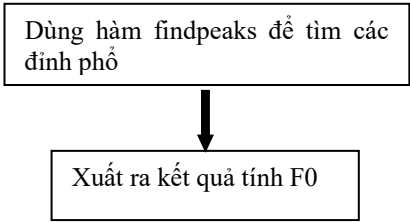
- Trường hợp dãy tín hiệu hữu hạn : dùng phép biến đổi Fourier rời rạc DFT.
 - Các phép tính chuỗi Fourier được thực hiện bằng cách sử dụng phép biến đổi Fourier rời rạc (DFT), do đó được thực hiện bằng cách sử dụng thuật toán hiệu quả được gọi là chuyển đổi Fourier(FFT).
- b) Cửa sổ hamming (kiểm tra kết quả tính được Fo trên miền tần số):
- Biểu diễn tín hiệu rời rạc bằng cách lấy mẫu tín hiệu tương tự dưới dạng:

$$w(n) = 0.52 - 0.46\cos(2\pi\frac{n}{N}) \quad [10]$$
 - Làm ngắn đoạn tín hiệu, trả về tần số gần đúng của tính hiệu

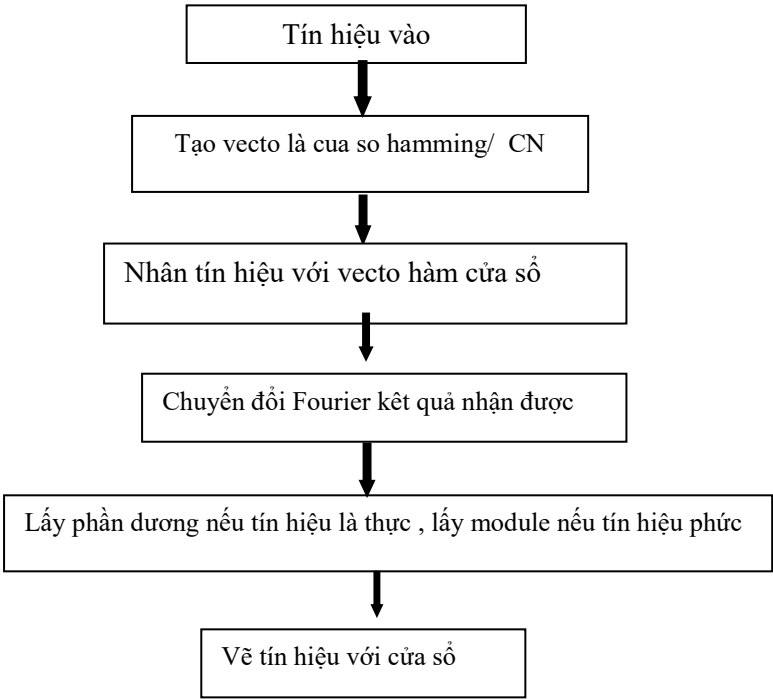
2. Sơ đồ thuật toán :

2.1 Biến đổi Fourier:





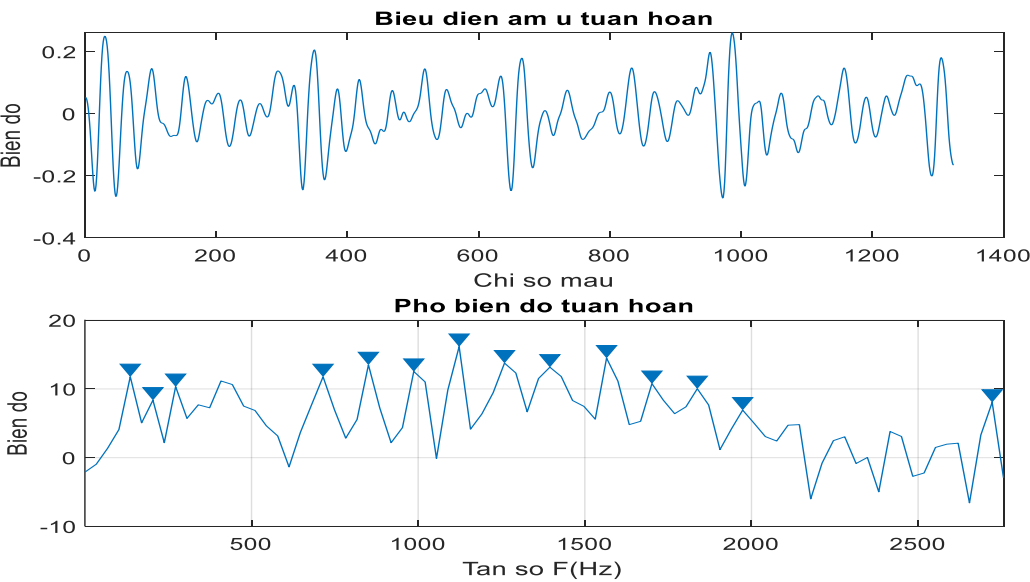
2.2. Hàm sửa số kiểm tra kết quả:



3. Các vấn đề phát sinh và hướng giải quyết

3.1. Vấn đề:

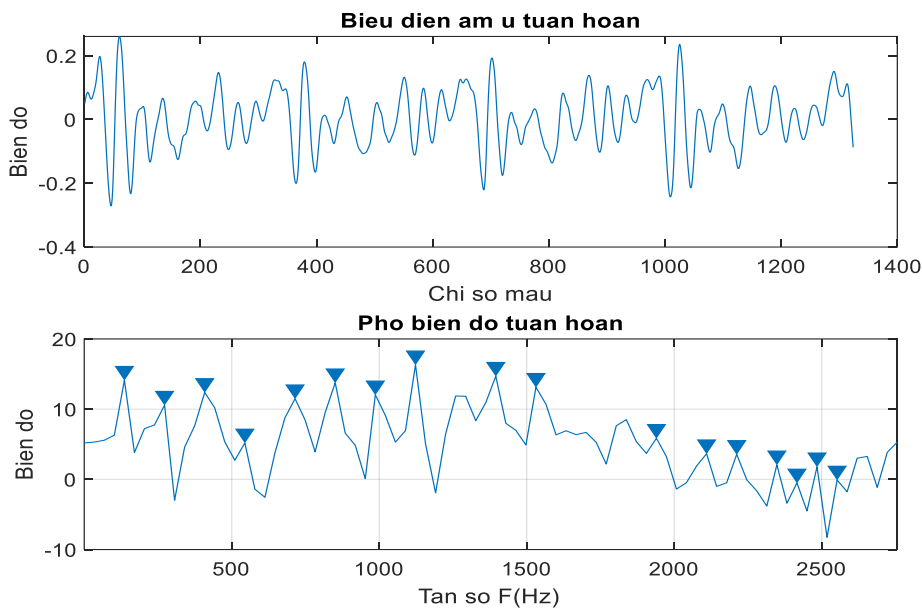
Hàm findpeaks tìm số đỉnh của phổ không chính xác tuyệt đối, từ đó dẫn đến sai số trong quá trình tính tần số cơ bản F0.



Hình 10. Hình ảnh tính F0 khi lọc tín hiệu

3.2. Hướng giải quyết:

Do quá trình thu âm tín hiệu tiếng nói bị nhiễu hay lẫn nhiều hợp âm nên để xác định các đỉnh phổ một cách chính xác thì phải lọc và loại bỏ các tín hiệu dư thừa.



Hình 11. Hình ảnh tính Fo khi đã lọc tín hiệu

II. CÀI ĐẶT CÁC THUẬT TOÁN

```
% tin hieu vao la doan ghi am
[y,Fs]=audioread('D:\Download\Lop1611-Nhom19\code\frequencyDomain\signal\tan\1.wav');
% lay doan tin hieu o vung tuan hoan 30ms
y_segment1 = y(0.57*Fs:0.6*Fs);

subplot(321);
plot(y_segment1);
axis([0 1200 -0.5 0.5 ]);
xlabel('Chi so mau');
ylabel('Bien do');
title('Bieu dien am a tuan hoan');

subplot(323);
% tin hieu vao y2 doan 30ms, tan so
[dft_ylog, tt] = convertFrquency(y_segment1, Fs);
plot(tt,dft_ylog);
xlabel('Tan so F(Hz)');
ylabel('Bien do');
title('Pho bien do tin hieu tuan hoan');

% Ve cac dinh so dinh
[pks,locs] = findpeaks(dft_ylog,tt,'Threshold',1);%tim cac dinh cua tin hieu va vi tri cua chung
findpeaks(dft_ylog,tt,'Threshold',1);% ve cac dinh cua tin hieu
do=length(locs) ; %so dinh

%tim do chenh lech giua hai dinh
for i=2:1:do
    f1(i,1)= locs(i)- locs(i-1);
end
F1=sum(f1)/(length(f1)-1); %tinh F0

subplot(322);
y_segment2 = y(0.42*Fs:0.45*Fs);
```



```

plot(y_segment2);
axis([0 1200 -0.5 0.5 ]);
title('Bieu dien am a khong tuan hoan');
xlabel('Chi so mau');
ylabel('Bien do');

subplot(324);
% tin hieu vao y2 doan 30ms, tan so
[dft_ylog, tt] = convertFrquency(y_segment2, Fs);
plot(tt,dft_ylog);
title('Pho bien do tin hieu khong tuan hoan');
xlabel('Tan so F(Hz)');
ylabel('Bien do');

% Ve cac dinh so dinh
[pks,locs] = findpeaks(dft_ylog,tt,'Threshold',1);%tim cac dinh cua tin hieu va vi tri cua chung
findpeaks(dft_ylog,tt,'Threshold',1);%ve cac dinh cua tin hieu
title('Pho bien do khong tuan hoan');
xlabel('Tan so F(Hz)');
ylabel('Bien do');
do=length(locs) ;    %so dinh

%tim do chenh lech giua hai dinh
for i=2:1:do
    f2(i,1)= locs(i)- locs(i-1);
end
F2=sum(f2)/(length(f2)-1);    %tinh F0

% kiem tra tan so
subplot(325);
N= 2048;
[xabs] = check_hamm(y, N);
plot(xabs(1:N));
xlabel('Tan so (Hz)');
ylabel('Bien do |x|');
title('Check with Hamming Window N = 1024');

```

***Các hàm được sử dụng trong bài**

a. Hàm chuyển đổi

```

% dua tin hieu vao va tan so tin hieu
% dophan gia cu Fs1 Hz thi lay mau 1 lan
% lay phan duong neu la thuc, lay module neu phuc , dong thoi bien doi fourier dfty1=abs(fft(signal));
%tao length(dfty1) diem, khoang cach giua cac diem ; (Fs1/16 -1/Fs1)/ (length(dfty1) - 1);
% dftylog1=10*log10(dfty1);    %chuyen sang mien logarit
function [dft_ylog, tt] = convertFrquency(signal, Fs)
Fs1 = Fs;
dft_y=abs(fft(signal));
dft_y=dft_y(1:(length(dft_y)/16));
tt=linspace(1/Fs1,Fs1/16,length(dft_y));
dft_ylog=10*log10(dft_y);
end

```

b. Hàm cửa sổ kiểm tra tần số tính được:

```

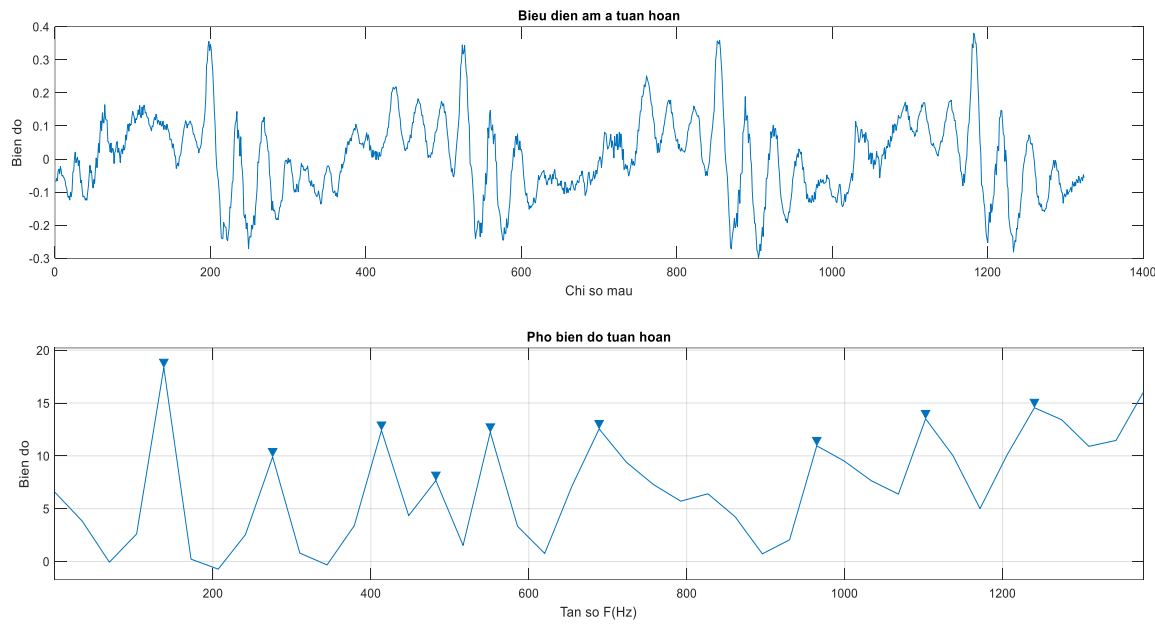
% dua vao tin hieu
% tao vecto cua so hamming
% nhan tin hieu v?i ham cua so hamming

```

```
% chuyển đổi fourier
% lấy phần dương nếu là thực, lấy module nếu phức
function [xabs] = check_hamm(signal, N)
x= signal;
hamm = [hamming(N)' zeros(1, length(x) - N)];
x_hamm = hamm.* x';
dftx = fft(x_hamm);
xabs = abs(dftx);
end
```

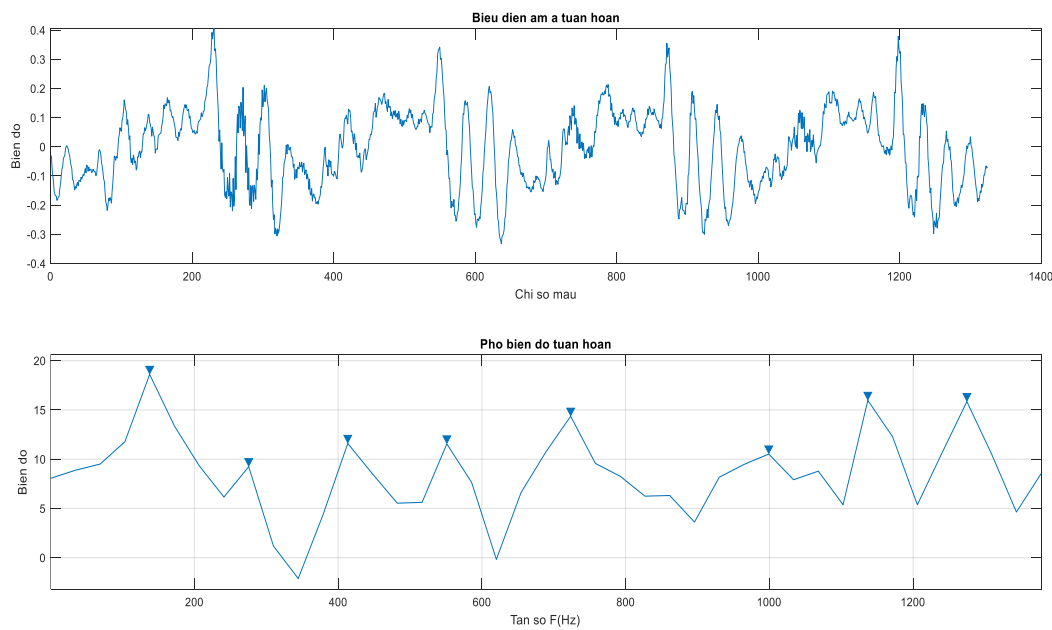
III. KẾT QUẢ THỰC NGHIỆM:

- 1. Hình ảnh tốt nhất, trung bình và xấu nhất khi tính F0 với thuật toán trên :
 - Trong trường hợp F0 cho kết quả tốt nhất khi chọn tín hiệu tại đoạn tuần hoàn thì F0= 142 (Hz)



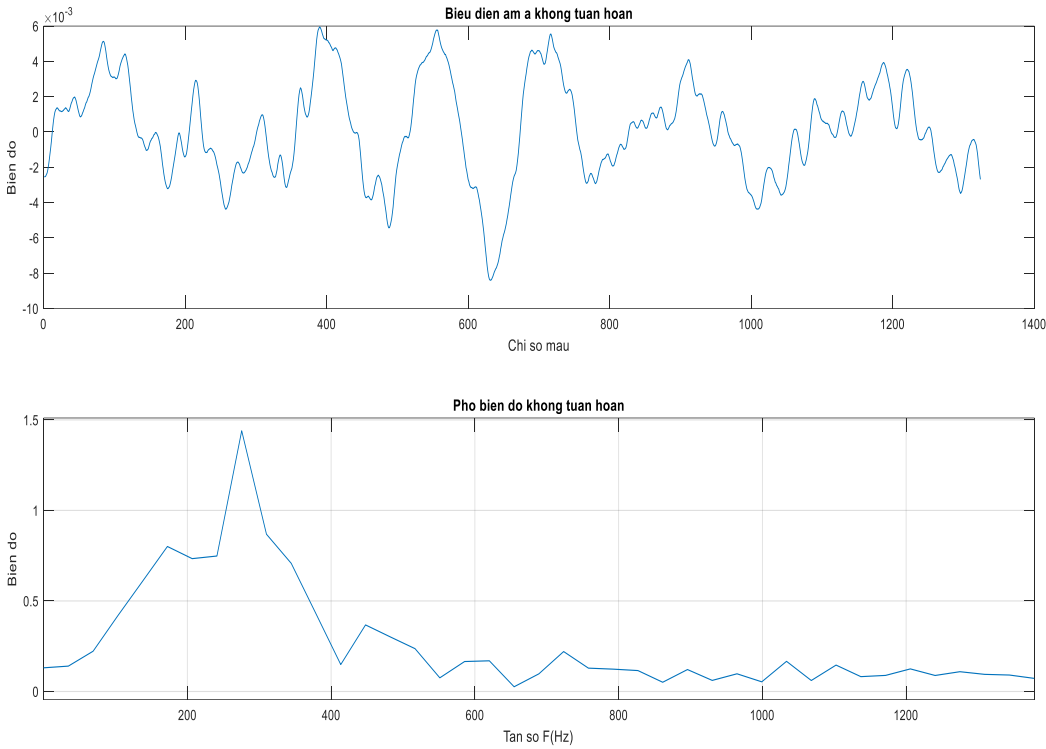
Hình 12: Trường hợp tín hiệu cho kết quả tính F0 chính xác nhất

- Trường hợp F0 cho kết quả trung bình khi chọn tín hiệu tại đoạn gần tuần hoàn thì F0= 183 (Hz)

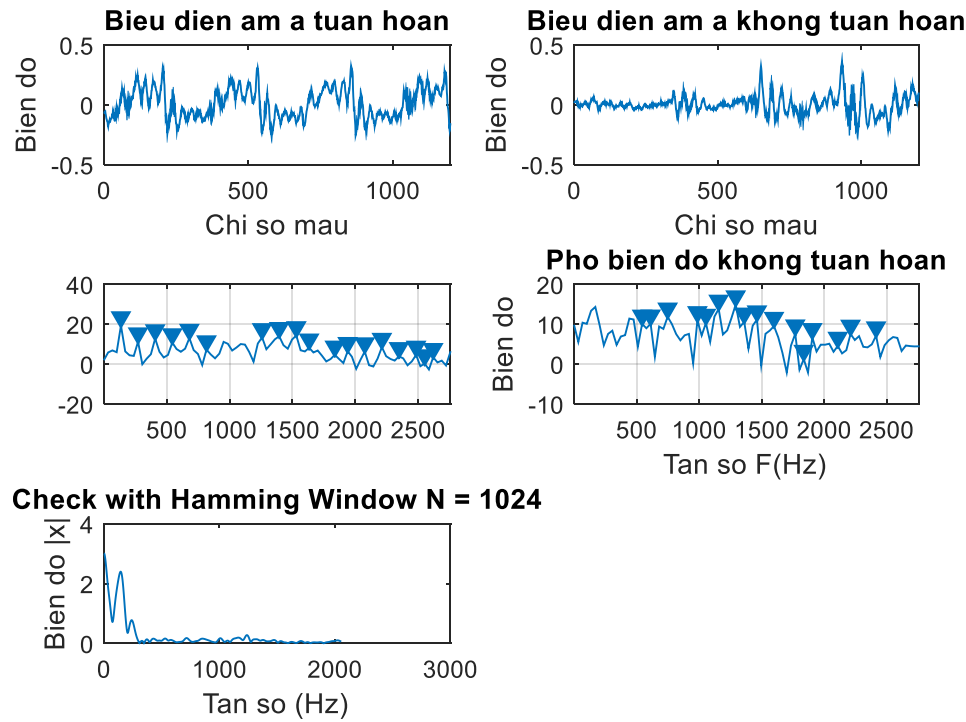


Hình 13: Trường hợp tín hiệu cho kết quả tính F0 trung bình

- Trong trường hợp F0 cho kết quả xấu nhất khi chọn tín hiệu tại đoạn không tuần hoàn thì chương trình báo lỗi vì không xác định được F0.



Hình 14: Trường hợp tín hiệu cho kết quả tính F0 kém chính xác



Hình 15: Hình ảnh dùng của sổ hamming kiểm tra kết quả $F = 140$ (Hz)

2. Sai số RMSE:

Bảng 2 . Bảng sai số RMSE của 20 âm

	a		e		i		o		u	
	Fo chuẩn	Sai số	Fo chuẩn	Sai số	Fo chuẩn	Sai số	Fo chuẩn	Sai số	Fo chuẩn	Sai số
Son	146.4	9.8	142.86	7.5	128.48	4.4	125.5	6.56	145.6	3.72
Nhân	158.85	13.7	147.06	7.31	156.25	10.96	153.85	14.52	149.25	8.24
Quyên	235.83	3.83	230.89	5.22	233.33	6.25	234.57	9.68	252	8.35
Tân	137.7	6.2	142.2	4.38	152	3.85	147.1	6.3	158	7.2

- Thuật toán tính F0 tự động trên miền tần số có độ chính xác tương đối so với kết quả tính thủ công.
- 3. Ưu nhược điểm thuật toán
 - Ưu điểm của thuật toán: việc phân tích phổ giúp tìm được tần số cơ bản của tín hiệu với độ chính xác tương đối.
 - Nhược điểm của thuật toán: khi tín hiệu bị nhiễu hay lẫn nhiều tạp âm thì tần số tìm được với độ chính xác thấp và các đỉnh phổ của tín hiệu được xác định với độ chính xác không tuyệt đối.
- 4. Yếu tố ảnh hưởng đến kết quả tính F0:
 - Khi tính tần số cơ bản F0 trên miền tần số, yếu tố ảnh hưởng đến độ chính xác của thuật toán là tín hiệu vào phải không bị nhiễu và phải chọn được những đoạn tuần hoàn của tín hiệu để xuất ra được kết quả F0 chính xác nhất. Khi chọn những đoạn tín hiệu bị nhiễu hay không tuần hoàn thì sẽ không xác định được các đỉnh của phổ biên độ, từ đó thuật toán tính F0 cũng trở nên sai.
 - Chọn những đoạn tín hiệu tuần hoàn thì kết quả tính F0 sẽ chính xác hơn.

D. KẾT LUẬN

- Bài báo cáo giúp em hiểu rõ hơn về một số kiến thức (cách tìm tần số cơ bản trên miền tần số và thời gian)
- Kỹ năng làm việc nhóm với các bạn được nâng cao
- Tăng khả năng tự tìm và đọc tài liệu để hoàn thành bài báo cáo
- Tăng khả năng viết báo, trình bày slide, thuyết trình
- Cảm ơn thầy đã cung cấp cho em các kiến thức cần thiết để làm bài báo cáo này.

E. NHỮNG ĐIỀU ĐÃ HỌC ĐƯỢC

Bài báo này thực hiện việc cài đặt thuật toán tìm tần số cơ bản của tín hiệu trên miền tần số và miền thời gian. Để làm tăng độ chính xác của thuật toán cần lựa chọn tín hiệu tốt không bị nhiễu. Khi tín hiệu đưa vào ổn định thì việc tính F0 sẽ chính xác hơn và sai số cũng nhỏ hơn. Trong tương lai sẽ thử nghiệm các thuật toán tính tần số cơ bản trên miền tần số cũng như miền thời gian tiên tiến hơn để cải thiện độ chính xác của giá trị F0.

F. TÀI LIỆU THAM KHẢO

[1] <https://viblo.asia/p/xu-ly-tieng-noi-tong-quan-ve-tieng-noi-ngu-dieu-va-mo-hinh-ngu-dieu-trong-xu-ly-tieng-noi-3KbvZ1NLGmWB>

[2], [3] Mục I đặt vấn đề của CẢI TIẾN THUẬT TOÁN TỰ TƯƠNG QUAN TÌM CAO ĐỘ CỦA TÍN HIỆU ĐÀN GHI-TA TRÊN VI XỬ LÝ ARM CORTEX-M4 (Nguyễn Bình Thiên, Ninh Khánh Duy)

[4] <https://www.ieev.org/2009/06/bo-loc-trung-vi-median-filter.html>

[5] <https://ww2.mathworks.cn/help/matlab/ref/fft.html>

[6] <http://tongquanvienthong.blogspot.com/2012/02/pho-tin-hieu.html>

[7] , [8] Sách Xử lý tín hiệu số - Vũ Văn Biên

[9] Wikipedia: Tần số âm cơ bản (https://vi.wikipedia.org/wiki/T%E1%BA%A7n_s%E1%BB%91_%C3%A2m_c%C6%A1_b%E1%BA%A3n)

[10] <https://www.mathworks.com/help/signal/ref/hamming.html>

Phạm Xuân Sự, Lê Xuân Thành, Bài giảng xử Xử ý tín hiệu số - Vũ Văn Biên
lý tiếng nói, Học viện công nghệ bưu chính viễn thông, 2010

(https://vi.wikipedia.org/wiki/T%E1%BA%A7n_s%E1%BB%91_%C3%A2m_c%C6%A1_b%E1%BA%A3n)

Phạm Xuân Sự, Lê Xuân Thành, Bài giảng xử lý tiếng nói, Học viện công nghệ bưu chính viễn thông, 2010

Xử ý tín hiệu số - Vũ Văn Biên

Wikipedia

(https://vi.wikipedia.org/wiki/T%E1%BA%A7n_s%E1%BB%91_%C3%A2m_c%C6%A1_b%E1%BA%A3n)

] Oberhettinger F., "Tables of Fourier Transforms and Fourier Transf
orms of Distributions." Springer, 1990.