

NGHIÊN CỨU HỆ THỐNG QUAN TRẮC ÂM THANH ĐA CẢM BIẾN PHỤC VỤ MỤC ĐÍCH QUÂN SỰ

RESEARCH SOUND MONITORING SYSTEM USING MULTI SENSOR FOR MILITARY PURPOSES

Hoàng Minh Sáng¹, Bùi Thị Duyên^{2,*}

DOI: <https://doi.org/10.57001/huih5804.2023.104>

TÓM TẮT

Định vị mục tiêu là nhiệm vụ cực kỳ quan trọng trong hầu hết các hoạt động quân sự. Bài báo giới thiệu hệ thống quan trắc, thu thập xử lý đa cảm biến âm thanh trên công nghệ vi mạch điện tử, xử lý phổ âm thanh đa cảm biến trên cơ sở biến đổi Fourier xác định đặc trưng của nguồn phát xạ âm thanh. Từ đó kết hợp thông tin để định hướng nguồn phát xạ nhằm xác định các mục tiêu phát ra tiếng động, cung cấp thêm khả năng định vị mục tiêu thụ động từ tín hiệu âm thanh thu được cho các tổ nhóm tuần tra, nhóm trinh sát, trạm gác, trạm giám sát, nhóm tác chiến... phục vụ mục đích quân sự, an ninh quốc phòng.

Từ khóa: Đa cảm biến, biến đổi Fourier, nguồn âm thanh.

ABSTRACT

Target positioning is an extremely important task in most military operations. This paper introduces the sound monitoring, collecting and multi-sensor processing system on electronic circuit technology, processing multi-sensor audio spectrum on the basis of Fourier transform to identify the characteristics of the sound source system. Then, the information is combined to detect the emission source direction to determine the targets that emit noise, providing additional passive target positioning from the received audio signal for patrol groups, reconnaissance groups, monitoring stations, guard stations, monitoring stations, combat groups... for military purposes, security and defense.

Keywords: Multi-sensor, Fourier transform, sound source.

¹Viện Tự động hóa Kỹ thuật quân sự, Viện Khoa học và Công nghệ Quân sự

²Khoa Điều khiển và Tự động hóa, Trường Đại học Điện Lực

*Email: duyenbt@epu.edu.vn

Ngày nhận bài: 29/3/2023

Ngày nhận bài sửa sau phản biện: 10/5/2023

Ngày chấp nhận đăng: 15/6/2023

1. ĐẶT VẤN ĐỀ

Âm thanh là một đặc trưng của nhiều loại hình hoạt động khác nhau trong đó có hoạt động quân sự. Việc sử dụng âm thanh để trinh sát, phát hiện các đối tượng tác chiến trong các môi trường tác chiến khác nhau đã và đang được nhiều nước trên thế giới nghiên cứu và ứng dụng. Các ứng dụng phổ biến của âm thanh phải nhắc đến là các hệ thống sonar trên các tàu mặt nước, tàu ngầm để xác định các đối tượng trong môi trường nước (xác định vật cản, xác định luồng cá...), hệ thống trinh sát thông qua các cảm

biến âm thanh cũng như nhờ sự phân tích của trắc thủ đã phát huy hiệu quả trên các tàu ngầm, trên bộ các ứng dụng phát hiện tiếng động đã được ứng dụng vào cuộc chiến Việt Nam của quân đội Mỹ trong hàng rào điện tử Macnamara dọc vĩ tuyến 17... Ngày nay việc sử dụng âm thanh đã có sự phát triển vượt bậc ví dụ như ứng dụng cho nhà thông minh, nhận dạng giọng nói trên các thiết bị điện tử, các hệ thống an ninh chống xâm nhập ...

Từ đó thấy được việc nghiên cứu, xử lý tín hiệu âm thanh luôn có tính thời sự và khoa học. Trong bài báo này chúng tôi giới thiệu hệ thống quan trắc, thu thập xử lý tín hiệu âm thanh đa cảm biến trên nền tảng công nghệ vi mạch điện tử để định hướng nguồn phát ra âm thanh (âm thanh trong dải tần số 10Hz đến 20.000Hz) nhằm phục vụ cho các hoạt động đặc thù quân sự.

2. CÁC ĐẶC TRƯNG CƠ BẢN TÍN HIỆU ÂM THANH

Âm thanh là một loại sóng vật lý, giống như nhiều loại sóng khác, được đặc trưng bởi tần số, bước sóng, chu kỳ, biên độ và vận tốc lan truyền [1, 2, 3]. Lan truyền âm thanh là quá trình làm lan truyền dao động sóng với các tính chất như: sóng âm truyền từ môi trường này sang môi trường khác thì tần số và chu kỳ của sóng không đổi; với môi trường đồng nhất thì âm truyền đi với tốc độ cố định; tốc độ truyền âm phụ thuộc vào tính chất của môi trường (tính đàn hồi, mật độ, nhiệt độ...), mật độ càng cao, tính đàn hồi lớn thì vận tốc truyền âm càng nhanh. Về cơ bản vận tốc truyền âm trong vật rắn lớn hơn vận tốc trong nước và trong không khí.

Nếu xét các môi trường rắn, lỏng, khí là các môi trường đồng nhất thì vận tốc truyền âm tương ứng sẽ là: trong chất khí có vận tốc khoảng 340m/s, chất lỏng khoảng 1.500m/s, chất rắn khoảng 6.100m/s.

Bên cạnh các chất truyền âm thì cũng có các chất liệu triệt tiêu âm thanh có thể chia thành hai nhóm đó là: các chất liệu không cho sóng âm truyền qua hay khả năng truyền qua rất ít gọi là chất cách âm; các chất liệu mà sóng âm truyền qua được nhưng một phần sóng âm bị tiêu hao được gọi là chất tiêu âm.

Việc nghiên cứu xử lý tín hiệu âm thanh nói chung đều dựa trên các đặc trưng cơ bản của âm thanh nói trên cũng như môi trường hoạt động.

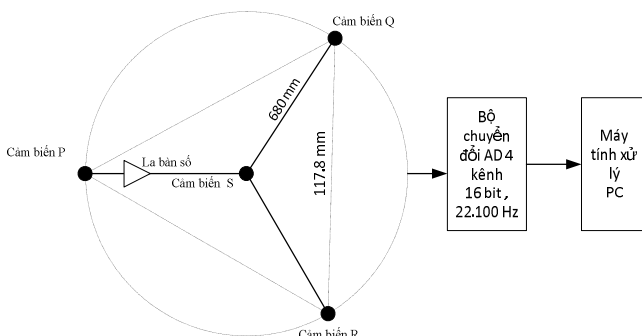
3. XÂY DỰNG MÔ HÌNH, PHẦN MỀM XỬ LÝ ĐỊNH HƯỚNG NGUỒN ÂM THANH

3.1. Nguyên lý xác định nguồn âm của con người

Để có thể xác định nguồn âm thanh trước hết ta tìm hiểu về khả năng định hướng nguồn âm thanh của con người. Trong thực tế khả năng định hướng nguồn âm thanh của con người là một quá trình gọi là “nghe bằng hai tai” (binaural hearing). Tai người được cấu tạo, bố trí đối xứng hai bên đầu, đầu là vật cản tự nhiên đối với các nguồn âm, chính vì thế âm thanh từ các nguồn âm phát ra nghe bởi mỗi tai sẽ thay đổi theo thời gian, âm lượng, pha và tần số. Những sai khác này là các dấu hiệu để não bộ giải mã vị trí (hướng, khoảng cách) của nguồn âm. Đây là một quá trình phức tạp thông qua thực tế não bộ được huấn luyện để đưa ra các dự báo gần đúng nhất. Để xây dựng một hệ thống tương đồng cần nhiều thời gian và công sức, hiện tại tập trung xây dựng hệ thống xác định hướng của nguồn âm phát ra, còn việc xác định khoảng cách nguồn âm sẽ được thực hiện trong nghiên cứu tiếp theo. Để xác định hướng của nguồn âm, não bộ con người có thể thông qua sự khác biệt cơ bản sau của tín hiệu nhận được từ hai tai: trước hết là sự khác biệt về thời gian, khi một nguồn âm đến từ phía trái thì tai trái sẽ nhận được sớm hơn tai phải khoảng 1ms (mili giây); tiếp đến là khác biệt về âm lượng với nguồn âm trên thì tai trái nhận âm lượng lớn hơn tai phải; sau nữa là khác biệt về tần số các tần số cao sẽ bị triệt tiêu, bị cản bởi đầu nên tai phải nhận được ít âm thanh có tần số cao hơn tai trái. Khi nguồn âm thanh đến từ hướng trực diện thì thời gian, âm lượng, và tần số nhận được ở hai tai sẽ như nhau. Theo phản xạ bản năng sẽ xoay đầu một góc cho đến khi nhận được sự khác biệt ở hai bên tai đủ để bộ não giải mã được hướng của nguồn âm thanh.

3.2. Xây dựng mô hình hệ thống

Căn cứ vào những phân tích trên tiến hành xây dựng mô hình hệ thống quan trắc âm thanh gồm: 04 cảm biến (03 cảm biến định hướng và 01 cảm biến đa hướng); bộ chuyển đổi dữ liệu âm thanh sang dữ liệu số 4 kênh AD 16 bit dữ liệu, tần số lấy mẫu lớn nhất 22100Hz; la bàn số; máy tính xử lý tín hiệu. Mô hình hệ thống được mô tả trên hình 1.

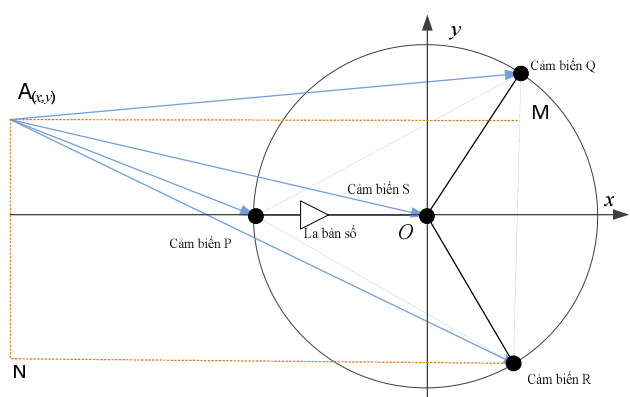


Hình 1. Mô hình hệ thống quan trắc âm thanh đa cảm biến

Các cảm biến âm thanh định hướng P, Q, R có nhiệm vụ nhận tín hiệu âm thanh từ các hướng truyền đến, chúng được bố trí lệch nhau một góc 120° nằm trên đường tròn có bán kính 680mm. Cảm biến đa hướng S được bố trí tại

tâm hình tròn. Cảm biến này được sử dụng để xác định các mẫu tín hiệu cần xử lý và làm cảm biến tham chiếu khi tính toán. La bàn số được lắp có trục tham chiếu trùng với đường thẳng PS kết nối giữa cảm biến P đến cảm biến S, với mục đích tham chiếu hướng nguồn âm thanh với hướng ngoài thực địa. Ngoài ra trong hệ thống còn tích hợp bộ chuyển đổi dữ liệu AD cho phép thu nhận dữ liệu âm thanh từ bốn cảm biến, số hóa dữ liệu và truyền về cho máy tính xử lý.

Giả sử có nguồn âm thanh $A_{(x,y)}$ chưa xác định lan truyền đến hệ thống, các cảm biến thu nhận tín hiệu âm thanh số hóa dữ liệu và truyền về cho máy tính xử lý tính toán, phần mềm sẽ xử lý dữ liệu để đưa ra hướng nguồn âm. Hệ tọa độ xOy được định nghĩa trục Ox là đường thẳng PO, trục Oy vuông góc với trục Ox.



Hình 2. Hệ tọa độ sử dụng trong hệ thống

3.3. Xây dựng phần mềm xử lý tính toán

Dựa trên yêu cầu bài toán đặt ra, phần mềm xử lý tính toán sẽ có các chức năng chủ yếu sau:

- Thu nhận mẫu và xử lý đồng bộ tín hiệu;
- Xác định các tín hiệu âm thanh cần xử lý định hướng thông qua cảm biến trung tâm, tạo mẫu so sánh cho các bước tiếp theo;
- Tính toán độ lệch pha, thời gian và cường độ âm thanh thông qua các hàm tương quan chéo giữa 3 cảm biến P, R, Q với cảm biến S;
- Giải bài toán xác định vị trí nguồn âm.

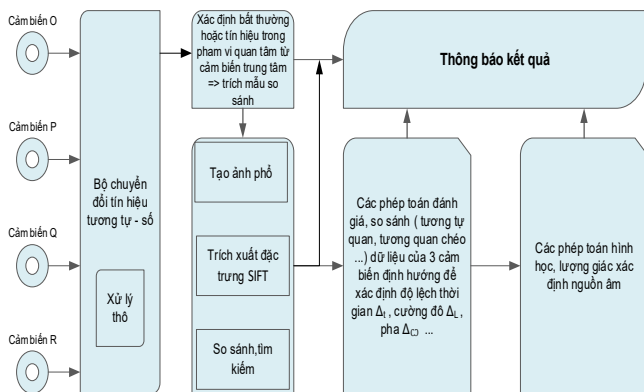
Sơ đồ khối chức năng chính của phần mềm được thể hiện trên hình 3.

Thiết lập điều kiện làm việc của hệ thống gồm tập các tham số:

- Tần số lấy mẫu;
- Kích thước khung (frame size): số điểm mẫu trong mỗi khung;
- Chồng lấp khung (frame overlap): số điểm mẫu chồng lấp giữa các khung liên tiếp;
- Bước khung (hoặc kích thước hop): bằng với kích thước khung trừ đi khoảng chồng lấp;
- Tỷ lệ khung (frame rate): số các khung trên một giây, bằng với tần số mẫu chia cho bước khung;

- Xác định tín hiệu nền: là các tín hiệu xuất hiện liên tục, ổn định mà các cảm biến nhận được khi không có âm thanh. Các bất thường của tín hiệu kéo dài dưới 10ms được coi là nhiễu nền. Tín hiệu nền phụ thuộc vào thiết bị và môi trường vận hành hệ thống;

- Đồng bộ tín hiệu nhận được từ các cảm biến: trong thực tế, với môi trường được coi là đồng đẳng thì các cảm biến thu nhận tín hiệu từ cùng 1 nguồn luôn có sự chênh lệch nhau về thời gian, độ lớn... Cần phải hiệu chuẩn lại hệ thống các cảm biến bằng phần mềm thông qua cách xác định trước các Δ_0 cho các thông số độ trễ về thời gian, cường độ giữa các cảm biến.



Hình 3. Các khối chức năng của phần mềm xử lý tính toán

Xác định nguồn âm thanh mẫu:

Trong hệ thống (hình 2), cảm biến S (đặt ở tâm hình tròn) có nhiệm vụ xác định mẫu nguồn âm cần phân tích. Tín hiệu từ cảm biến này được quét liên tục, khi phát hiện bất thường về biên độ trong khoảng thời gian lớn hơn 10ms phần mềm sẽ thực hiện các bước sau:

- Biến đổi FFT với các tham số đã được thiết lập:
- + Tần số lấy mẫu: $f_s = 20050\text{Hz}$;
- + Kích thước khung: $f_z = \text{frame_size} = f_s * 25 / 1000$;
- + Chồng lấp khung: $f_o = \text{frame_overlap} = f_s * 15 / 1000$;
- + Bước khung: $h_s = \text{hop_size} = f_z - f_o$;
- + Tỷ lệ khung: $f_r = \text{frame_rate} = f_s / h_s$;

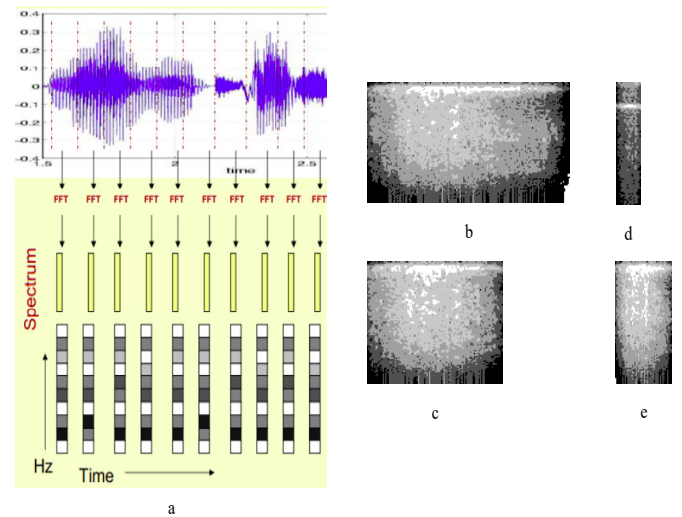
Nếu tần số và cường độ nằm trong phạm vi quan tâm thì thực hiện trích mẫu tín hiệu nguồn âm tại thời điểm ghi nhận bất thường - có sự kiện âm thanh (t_0) và chuyển sang bước tiếp theo.

- Xác định mẫu âm thanh thu nhận được có thuộc các sự kiện cần theo dõi hay không. Để giải quyết nhiệm vụ này, chúng tôi thực hiện biến đổi tín hiệu âm thanh thu được sang dạng ảnh phổ, trích các đặc trưng SIFT (Scale-Invariant Feature Transform) của ảnh phổ để tìm kiếm mẫu tương đồng với mẫu tín hiệu thu được trong cơ sở dữ liệu đã lưu.

Phổ của tiếng nói là một phương pháp biểu diễn tín hiệu âm bằng cách kết hợp thời gian và tần số trong đó một trục, một trục biểu diễn thời gian và giá trị mỗi điểm ảnh là biên độ của các thành phần tần số có trong tín hiệu.

Bản chất của cách biểu diễn này là biểu diễn tín hiệu trên miền tần số nhưng được thực hiện với các đoạn tín hiệu ngắn đủ ngắn để đảm bảo tín hiệu ổn định theo thời gian.

Để tạo ra ảnh phổ, tín hiệu âm thanh sẽ được phân thành các khung có thời gian ngắn nhằm đảm bảo tính ổn định của tín hiệu. Sau đó, các khung dữ liệu được cho qua hàm cửa sổ, thực hiện phân tích FFT. Ghép các hệ số FFT theo thứ tự thời gian sẽ thu được phổ tần số của tín hiệu tiếng nói.



Hình 4. Sơ đồ trích xuất (a) và ảnh phổ của tiếng súng AK lưu trong CSDL (b), tiếng súng AK ở hiện trường (c), tiếng huyết sáo (d) và tiếng vỗ tay (e)

Việc chuyển tín hiệu âm thanh thành ảnh phổ, trong đó các hàng biểu diễn các khoảng tần số, các cột biểu diễn thời gian, giá trị màu biểu diễn cường độ tín hiệu được thực hiện với giải thuật tổng quát sau:

Tham số vào:

aSignal As Double(): mảng chứa tín hiệu đã được lượng tử hóa $0 \Rightarrow 5V$

Tham số hệ thống:

- FrameSize As Integer: Kích thước khung - số mẫu trong mỗi khung
- FrameOverlap As Integer: Chồng lấp khung - số mẫu chồng lấp giữa các khung liên tiếp;
- sampleRate As Integer: tần số lấy mẫu;
- frequencyInterval As Double: Độ rộng vùng tần số trên phổ công suất (Hz) - độ rộng của Bin. Nếu độ rộng Bin = 200 thì khoảng tần số nghe được của người (20000Hz) được biểu diễn với 100 Bin

Kết quả trả về: ảnh mức xám 256 màu với định dạng mảng byte

Bắt đầu:

B1: Khởi tạo các biến cục bộ

`samplesPerMillisecond As Double = (sampleRate / 1000)`

`samplesPerFrame As Integer = CInt(FrameSize * samplesPerMillisecond)`

```

iStep As Integer = (FrameSize - FrameOverlap)
'in millisecond
samplesCount As Integer = aSignal.Length
jPos As Integer = 0; jCol As Integer = 0
samplesToCalc(samplesPerFrame - 1) As Double
iRows As Integer = bin= (SampleRate/2) /
frequencyInterval
iData(,) As Byte = {} ' ma trận số nguyên 2 chiều lưu các
giá trị màu của các điểm ảnh kết quả. Trong đó:
    • Số hàng (iRows) được xác định theo số lượng bin :
    bin= (SampleRate/2) / frequencyInterval
    • Số cột phụ thuộc vào các tham số FrameSize,
    FrameOverlap, sampleRate, frequencyInterval ...
B2: Tiền xử lý tín hiệu đầu vào: lọc nhiễu, lọc thông thấp,
lọc thông cao
B3: Biểu diễn cường độ các tần số theo thời gian thành
điểm ảnh
While jPos <= samplesCount - samplesPerFrame
- Lấy khung tín hiệu từ vị trí jPos và lấy samplesPerFrame
mẫu
samplesToCalc = GetFrame ( jPos,samplesPerFrame)
- Cho khung tín hiệu qua cửa sổ Hamming
- Chuyển khung tín hiệu từ miền thời gian sang miền
tần số để thu được mảng phổ năng lượng của tín hiệu ở
các vùng tần số (mảng có số phần tử = Bin )
PowerSpectrum = FFT (samplesToCalc)
- Hiệu chỉnh kích thước ảnh:
iData((iRows - 1), jCol)
- Cập nhật giá trị các điểm ảnh cho các hàng (vùng tần
số) tại theo thời điểm
For i = 0 To PowerSpectrum.Length - 1
+ b = PowerSpectrum (i) ' byte giá trị màu của điểm ảnh
+ Chuẩn hóa dữ liệu ảnh nhằm phân hóa các đoạn mức
xám rõ rệt hơn, tùy mức độ yêu cầu ta có thể phân hóa
theo các ngưỡng khác nhau cho phù hợp (trường hợp dưới
đây, sau khi chuẩn hóa ảnh chỉ còn 6 màu)
Select Case b
Case b < 40: b = 0
Case 41 To 80: b = 40
Case 81 To 120: b = 80
Case 121 To 160: b = 160
Case 161 To 200: b = 200
Case b > 200 : b = 255
End Select
+ Cập nhật màu cho điểm ảnh iData(i, jCol) = b
Next
jPos += CInt(iStep * samplesPerMillisecond)

```

```

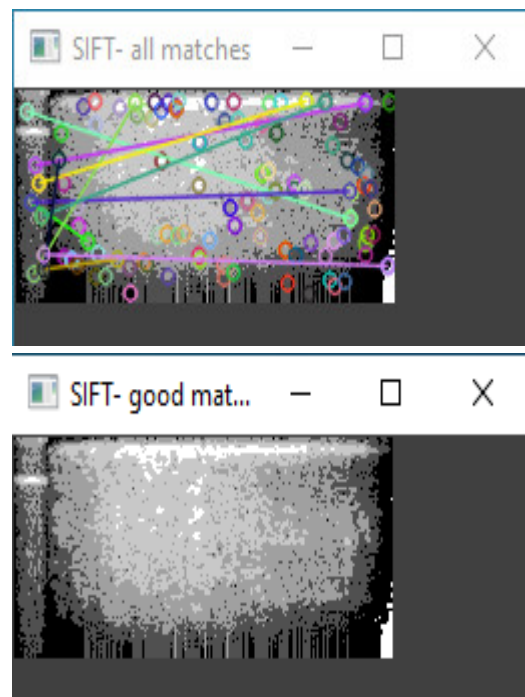
jCol += 1
End While
B4: Return iData

```

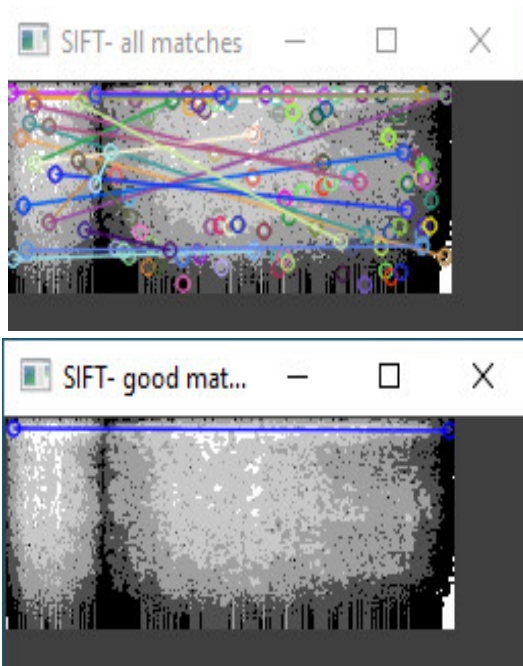
Kết thúc

Đến đây ta thu được ảnh phổ của tín hiệu âm thanh làm dữ liệu đầu vào cho việc nhận dạng với giải thuật SIFT.

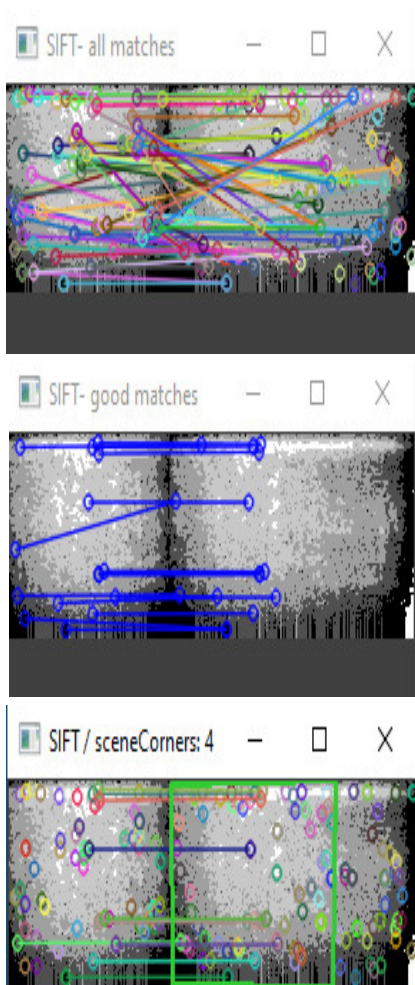
SIFT là một đặc trưng được sử dụng trong lĩnh vực thị giác máy, dùng để nhận dạng và miêu tả những điểm đặc trưng cục bộ trong ảnh được giới thiệu bởi David Lowe năm 1999 [Lowe, 2004] . Với đặc trưng bất biến với phép co giãn và phép xoay (Scale Invariant Feature Transform - SIFT) nó được sử dụng rất hiệu quả trong bài toán nhận dạng đối tượng. Các điểm đặc trưng SIFT thu được từ ảnh phổ tần số của tín hiệu âm thanh là các điểm cực trị trong phổ tần số, nghĩa là tại điểm đó biên độ của thành phần tần số đó là cực đại hoặc cực tiểu tương ứng với âm lượng của thành phần tần số đó là lớn hơn hoặc nhỏ hơn so với các thành phần tần số xung quanh nó. Mặc dù tín hiệu âm thanh phụ thuộc vào ngữ cảnh môi trường nhưng nó luôn vẫn tồn tại những đặc trưng bất biến, vì vậy, luôn tồn tại những điểm bất biến của những đỉnh cộng hưởng tần số trong tín hiệu của một âm thanh. Những đỉnh cộng hưởng này có thể bị tịnh tiến lên xuống do môi trường ngữ cảnh, có thể bị tịnh tiến sang trái, phải do thời gian thu tín hiệu lệch nhau, nhưng xét trong một phạm vi cục bộ thì chúng là bất biến. Vì vậy, SIFT là một đặc trưng phù hợp cho bài toán nhận dạng ảnh phổ âm thanh. Trong bài toán này chúng tôi dùng phương pháp so sánh các đặc trưng SIFT của các ảnh phổ kết hợp với phép tương quan để tìm xác định các mẫu tín hiệu âm thanh có đặc trưng tương tự nhau.



Hình 5. Trích rút đặc trưng SIFT và so sánh tương đồng giữa ảnh phổ của tiếng huýt sáo với tiếng súng AK trong CSDL



Hình 6. Trích rút đặc trưng SIFT và so sánh tương đồng giữa ảnh phổ của tiếng vỗ tay với tiếng súng AK trong CSDL



Hình 7. Trích rút đặc trưng SIFT và so sánh tương đồng giữa ảnh phổ của tiếng AK ở hiện trường với tiếng súng AK trong CSDL

Khi 2 ảnh phổ không có điểm tương đồng (hình 5) thì hai tín hiệu âm thanh không có đặc trưng tương tự nhau và không đồng âm.

Khi 2 ảnh phổ có điểm tương đồng (hình 6) thì hai tín hiệu âm thanh có một số đặc trưng tương tự nhau nhưng chưa đủ để kết luận đồng âm.

Khi 2 ảnh phổ có nhiều điểm tương đồng (hình 7) thì hai tín hiệu âm thanh có đủ số đặc trưng tương tự nhau và đồng âm.

Sau khi tìm được mẫu tương đồng sẽ chuyển sang bước tiếp theo.

Xác định nguồn phát theo nguyên lý TDOA (time difference of arrival):

Lập tương tự quan giữa các cảm biến định hướng với cảm biến đa hướng, lập tương quan chéo giữa các cảm biến định hướng, xác định độ lệch tín hiệu giữa các cảm biến. Với mỗi cảm biến định hướng: Biến đổi FFT tại thời điểm t_0 , thực hiện các phép lọc IIR, FIR; thực hiện các phép so sánh tương tự quan chéo giữa 3 cảm biến nhằm xác định các chênh lệch về thời gian của tín hiệu giữa cảm biến S với các cảm biến P, Q, R.

Hệ thống được thiết lập để xác định hướng của nguồn âm xuất phát từ ngoài hình tròn nên ta chỉ xét đến chiều âm thanh có hướng $P \rightarrow S, Q \rightarrow S, R \rightarrow S$. Khi đó luôn có ít nhất 1 cảm biến định hướng nhận được tín hiệu sớm hơn cảm biến trung tâm (theo hướng nguồn âm).

Gọi k là giá trị chênh lệch lớn nhất về vị trí (index) giữa điểm bắt đầu của mẫu trích từ S với điểm bắt đầu của mẫu tương quan trích từ P, Q, R được xác định bằng công thức:

$$k = \text{Tần số lấy mẫu} * \frac{d_0}{\text{Tốc độ âm}} \tag{1}$$

với: d_0 - là khoảng cách giữa cảm biến S đến các cảm biến P, Q, R (bán kính hình tròn).

Gọi n là điểm bắt đầu xuất hiện tín hiệu cần tìm tương quan trong khoảng thời gian được xét của cảm biến S. Ta có các tập tương quan chéo theo công thức:

$$r_{SP}(i) = \sum_{j=(n-k)}^{n+k} S(n). P(n - j) \tag{2}$$

$$r_{SQ}(i) = \sum_{j=(n-k)}^{n+k} S(n). Q(n - j) \tag{3}$$

$$r_{SR}(i) = \sum_{j=(n-k)}^{n+k} S(n). R(n - j) \tag{4}$$

Xác định vị trí đầu tiên có tương quan chéo bằng cách lấy phần tử khác 0 đầu tiên trong các tập $r_{SP}(i), r_{SQ}(i), r_{SR}(i)$ ta có các giá trị i_P, i_Q, i_R là các điểm trên trục thời gian mà tại đó các cảm biến P, Q, R tương ứng bắt đầu nhận được tín hiệu của nguồn âm đang xử lý ($i_P, i_Q, i_R \in Z$). Từ đó ta tính được:

$$N_P = (n - i_P); N_Q = (n - i_Q); N_R = (n - i_R), \tag{5}$$

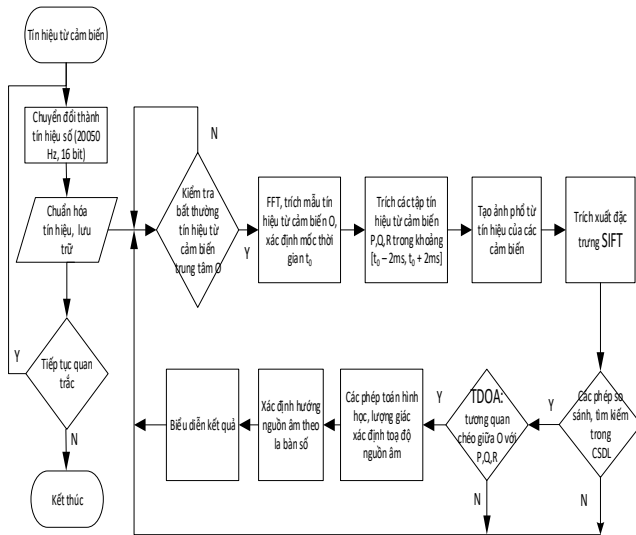
là số điểm chênh lệch trên trục thời gian so với điểm n (cách chia số điểm phụ thuộc vào tần số lấy mẫu). Sau đó tính được độ lệch khoảng cách giữa AS với AP, AQ, AR theo công thức:

$$\begin{aligned} \partial_P &= (N_P / f_s) * 340; \partial_Q = (N_Q / f_s) * 340; \\ \partial_R &= (N_R / f_s) * 340 \end{aligned} \tag{6}$$

Áp dụng tương tự cho cường độ và pha ta có: Chênh lệch cường độ $\Delta_L (P, Q, R)$; Chênh lệch pha $\Delta_w (P, Q, R)$.

Sau khi tính toán được chênh lệch khoảng cách giữa AO với AP; AO với AQ; AO với AR. Bằng các phép tính hình học và lượng giác ta xác định được tọa độ x, y chính là vị trí nguồn âm thanh cần tìm.

Sơ đồ qui trình xử lý tín hiệu, xác định hướng nguồn âm của phần mềm như hình 8.



Hình 8. Sơ đồ qui trình xử lý tín hiệu, xác định hướng nguồn âm của phần mềm

Phần mềm được xây dựng và viết trong môi trường .Net Framework. Thu nhận dữ liệu từ các cảm biến âm thanh, xử lý, tính toán và đưa ra hướng nguồn âm cần xác định. Hiện tiếp tục trong quá trình thử nghiệm, hiệu chỉnh.

4. KẾT LUẬN

Bài báo trình bày về hệ thống xử lý tín hiệu âm thanh đa cảm biến nhằm xác định hướng nguồn âm ứng dụng cho các mục đích quân sự. Trong hệ thống đã thu nhận xử lý, tính toán dữ liệu từ 04 cảm biến âm thanh. Tiến hành phân tích phổ, tìm kiếm tương quan khác biệt giữa các cảm biến, từ đó xác định được hướng nguồn âm mà hệ thống thu nhận được. Trong các nghiên cứu tiếp theo, sẽ nâng cao độ chính xác của việc định hướng nguồn âm, cũng như tiến tới xác định cự ly nguồn âm hay tọa độ nguồn âm.

TÀI LIỆU THAM KHẢO

- [1]. Phạm Thị Ngọc Yên, 2010. *Xu lý tín hiệu: Tín hiệu - hệ thống tuyến tính lọc số và DSP*. Science and Technics Publishing House, Hanoi.
- [2]. Nguyễn Quốc Trung, 1999. *Xu lý tín hiệu và lọc số, tập 1*. Science and Technics Publishing House, Hanoi.
- [3]. Nguyễn Quốc Trung, 2001. *Xu lý tín hiệu và lọc số, tập 2*. Science and Technics Publishing House, Hanoi.
- [4]. Horiuchi T., Mizumachi M., Nakamura S., 2004. *Iterative compensation of microphone array and sound source movements based on minimization of arrival time differences*. IEEE Processing Workshop 2004 Sensor Array and Multichannel Signal, pp. 566-570.
- [5]. Toni Heittola, Annamaria Mesaros, Tuomas Virtanen, Antti Eronen, 2011. *Sound Event Detection in Multisource Environments Using Source Separation*. Machine Listening in Multisource Environments.
- [6]. Chang Yao Tang, Wu Chi Lin, Cheng Hsu Chih, 2014. *International Symposium on Computer, Consumer and Control - Integrated Cross-Correlation and Genetic Algorithm to Improve TDOA-Based Locating System in Preventing Radio Interference*. IEEE 2014 International Symposium on Computer, Consumer and Control, pp. 1078-1081.
- [7]. Despoina Pavlidi, Anthony Griffin, Matthieu Puigt, Athanasios Mouchtaris, 2013. *Real-Time Multiple Sound Source Localization and Counting Using a Circular Microphone Array*. IEEE Transactions on audio, speech, and language processing, Vol. 21, No. 10, pp. 2193-2205.

AUTHORS INFORMATION

Hoang Minh Sang¹, Bui Thi Duyen²

¹Control Automation in Production and Improvement of Technology Institute, Academy of Military Science and Technology, Hanoi, Vietnam

²Faculty of Control and Automation, Electric Power University, Hanoi, Vietnam