NÂNG CAO CHẤT LƯỢNG ƯỚC LƯỢNG HIỆU THỜI GIAN ĐẾN TRONG BÀI TOÁN ĐỊNH VỊ NGUỒN ÂM THEO NGUYÊN LÝ TDOA

Trần Công Thìn[1*], Bùi Ngọc Mỹ[2], Nguyễn Huy Hoàng[3], Phạm Văn Hòa[1]

Tóm tắt: *Bài báo trình bày một số kết quả nghiên cứu giải pháp nâng cao chất lượng định vị nguồn âm sử dụng nguyên lý TDOA, trong trường hợp loại nguồn âm cần định vị được xác định trước. Trong đó đề xuất giải pháp cải thiện chất lượng ước tính hiệu thời gian đến dựa trên thuật toán GCC-PHAT-β với hệ số β thích nghi theo tín hiệu đưa vào ước tính. Trên cơ sở phân tích lý thuyết và tiến hành mô phỏng Monte-Carlo trên Matlab, kết hợp với dữ liệu được thu thập trong điều kiện thực tế bài báo sẽ đánh giá hiệu quả của giải pháp được đề xuất, qua đó cho phép nâng cao chất lượng định vị nguồn âm theo nguyên lý TDOA.*

Từ khóa: Định vị nguồn âm; TDOA; GCC-PHAT; GCC-PHAT- β.

1. MỞ ĐẦU

Trong các hệ thống định vị nguồn âm thanh sử dụng nguyên lý TDOA (Time Differences of Arrival) ước lượng hiệu thời gian đến là bước quan trọng nhất, quyết định chất lượng của hệ thống định vị nguồn âm [4]. Đặc biệt, với các hệ thống định vị nguồn âm ngoài trời, khoảng cách định vị xa, nhiễu và tạp âm tác động lớn tới độ chính xác số ước lượng hiệu thời gian đến, qua đó làm giảm độ chính xác định vị nguồn âm [5]. Hiện nay, việc ước tính trễ thời gian đến phần lớn dựa trên các thuật toán sử dụng thuật toán tương quan chéo tổng quát, trong đó hiệu quả nhất là phương pháp ước tính dựa trên phương pháp biến đổi pha GCC-PHAT (Generalized Cross Correlation - Phase Transform). Trong điều kiện tỉ số tín hiệu trên tạp âm (SNR - Singal to Noise Ratio) lớn, phương pháp GCC-PHAT mang lại hiệu quả cao, tuy nhiên, khi SNR giảm sai số của phương pháp GCC-PHAT là tương đối lớn [9]. Bài báo sẽ trình bày một số kết quả nghiên cứu cải thiện hiệu quả của phương pháp GCC-PHAT trong điều kiện SNR nhỏ bằng cách sử dụng biến thể GCC-PHAT-β, trong đó, hệ số β được tính toán thích nghi theo tín hiệu thu được trên các cảm biến. Hiệu quả của giải pháp góp phần nâng cao chất lượng định vị nguồn âm sử dụng nguyên lý TDOA.

2. ƯỚC LƯỢNG HIỆU THỜI GIAN ĐẾN

2.1. Định vị nguồn âm theo nguyên lý TDOA

Phương pháp định vị nguồn âm theo nguyên lý TDOA được xây dựng trên cơ sở ước tính hiệu thời gian đến τ_{ij} của tín hiệu âm thanh thu được trên các cảm biến.

$$\tau_{ij} = \tau_i - \tau_j \tag{1}$$

Trong đó: τ_i và τ_j lần lượt là thời gian tín hiệu truyền từ nguồn âm tới cảm biến thứ i và j. Triển khai công thức hiệu thời gian đến theo khoảng cách r_i, r_j từ nguồn âm tới các cảm biến và vận tốc lan truyền của âm thanh trong không khí v, thu được phương trình biểu diễn một hyperboloid trong không gian 3 chiều:

$$\tau_{ij} = \frac{r_i}{v} - \frac{r_j}{v} = \frac{\| \mathbf{x}_s - \mathbf{m}_i \| - \| \mathbf{x}_s - \mathbf{m}_j \|}{v} \tag{2}$$

Với $\mathbf{x}_s = [x_s, y_s, z_s]^T$, $\boldsymbol{m_i} = [x_i, y_i, z_i]^T$, $\boldsymbol{m_j} = [x_j, y_j, z_j]^T$ lần lượt là tọa độ của nguồn âm và tọa độ của hai cảm biến thứ i, và thứ j trong không gian.

Để xác định tọa độ nguồn âm $\mathbf{x}_s$ trong không gian 3 chiều, cần giao hội của ít nhất 3 hyperboloid, tương ứng cần ít nhất 4 cảm biến âm thanh (từ 1 - 4) để tạo thành 3 cặp cảm biến độc lập. Khi đó, tọa độ $\mathbf{x}_s$ của nguồn âm chính là nghiệm của hệ phương trình [1, 2]:

$$\begin{cases} v\tau_{12} = \sqrt{(x_1 - x_S)^2 + (y_1 - z_S)^2 + (z_1 - z_S)^2} - \sqrt{(x_2 - x_S)^2 + (y_2 - y_S)^2 + (z_2 - z_S)^2} \\ v\tau_{13} = \sqrt{(x_1 - x_S)^2 + (y_1 - z_S)^2 + (z_1 - z_S)^2} - \sqrt{(x_3 - x_S)^2 + (y_3 - y_S)^2 + (z_3 - z_S)^2} \\ v\tau_{14} = \sqrt{(x_1 - x_S)^2 + (y_1 - z_S)^2 + (z_1 - z_S)^2} - \sqrt{(x_4 - x_S)^2 + (y_4 - y_S)^2 + (z_4 - z_S)^2} \end{cases} \quad (3)$$

Như vậy, để định vị nguồn âm theo nguyên lý TDOA, có hai bước quan trọng cần thực hiện: Ước tính hiệu thời gian đến τ_{ij} của sự kiện âm thanh trên ít nhất 3 cặp cảm biến; Giải hệ phương trình (3) để xác định tọa độ của nguồn âm. Trong hai bước trên ước tính hiệu thời gian đến có vai trò quan trọng, thiết lập hệ số của phương trình định vị nguồn âm, quyết định độ chính xác định vị [4].

2.2. Ước tính hiệu thời gian đến

Để ước tính hiệu thời gian đến của tín hiệu thu được trên hai cảm biến, các nghiên cứu trong những năm gần đây chỉ rằng phương pháp sử dụng tương quan chéo tổng quát mang lại hiệu quả cao nhất [3-5]. Trong phương pháp này, hiệu thời gian đến được ước tính theo công thức [5]:

$$\hat{\tau}_{ij} = \arg\max\ r_{x_i x_j}(\tau) \quad (4)$$

Với $r_{x_i x_j}$ là hàm tương quan chéo tổng quát [5]:

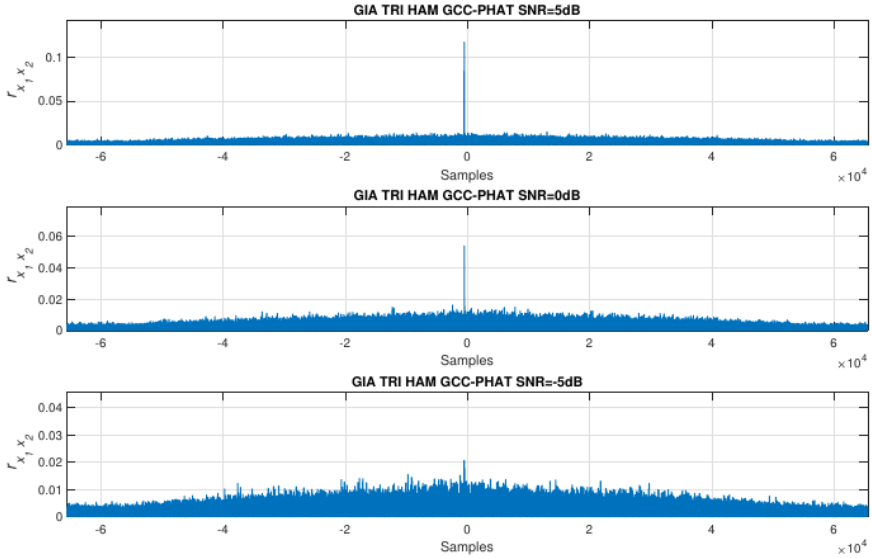
$$r_{x_1 x_2}(\tau) \triangleq \mathcal{F}^{-1}\psi_{x_i x_j}(f) = \int_{-\infty}^{+\infty} \psi_{x_i x_j}(f) e^{i2\pi f\tau} df = \int_{-\infty}^{+\infty} \Phi(f)\mathbf{S}_{x_i x_j}(f) e^{i2\pi f\tau} df \quad (5)$$

Trong đó:

$\mathbf{S}_{x_i x_j}(f) \triangleq \mathbf{E}[\mathbf{X}_i(f)\mathbf{X}_j^*(f)]$ là hàm Cross Spectrum;

$\Phi(f)$ là hàm trọng số trong miền tần số.

Có nhiều cách lựa chọn hàm trọng số $\Phi(f)$ khác nhau, tương ứng với các phương pháp tính toán hiệu thời gian đến khác nhau như, tương quan chéo kinh điển (CCC – Classiscal Cross Correlation), biến đổi làm trơn tín hiệu (SCOT – Smoothed Coherence Transform), biến đổi pha (PHAT – Phase Transform).



Hình 1. *Giá trị của hàm GCC-PHAT với các mức SNR khác nhau.*

Trong các phương pháp nêu trên thì GCC-PHAT là phương pháp có hiệu quả cao nhất, hoạt động hiệu quả và có độ chính xác cao [6]. Trong phương pháp này hàm trọng số trong miên tần số được tính theo công thức [6]:

$$\Phi(f) = \frac{1}{\left|S_{x_i x_j}(f)\right|} = \frac{1}{\left|X_i(f)\left[X_j(f)\right]^*\right|} \tag{6}$$

Với hàm trọng số được tính toán theo công thức trên thì toàn bộ thành phần biên độ được loại bỏ, khi đó việc ước lượng hiệu thời gian đến phụ thuộc hoàn toàn vào pha của tín hiệu.

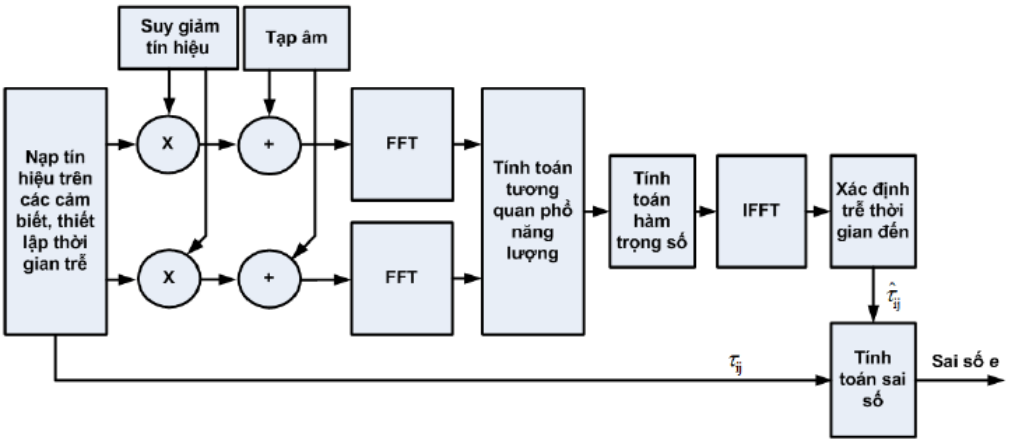
Tuy nhiên, khi phương pháp GCC-PHAT được dùng để ước tính hiệu thời gian đến của những tín hiệu có dải tần rộng, trong đó tạp âm mạnh chiếm ưu thế trong phần lớn băng tần của tín hiệu khiến cho hiệu suất của phương pháp suy giảm [9]. Trên hình 1 thể hiện giá trị của hàm GCC-PHAT tương ứng tại các giá trị SNR khác nhau, có thể nhận thấy khi SNR lớn đỉnh của hàm $r_{x_i x_j}$ nổi bật trên nền nhiễu, tuy nhiên, khi SNR suy giảm, đỉnh này hòa lẫn trên nền nhiễu lớn, gây ra khó khăn khi áp dụng công thức (4). Điều này dẫn tới sai số lớn khi ước tính hiệu thời gian đến, đặc biệt khi ước tính hiệu thời gian đến trong các hệ thống định vị nguồn âm ngoài trời, khoảng cách định vị xa, SNR tại các cảm biến âm thanh nhỏ. Do đó, đối với lớp ứng dụng này, đòi hỏi phải có phương pháp phù hợp nâng cao độ chính xác ước tính hiệu thời gian đến.

3. NÂNG CAO CHẤT LƯỢNG ƯỚC TÍNH HIỆU THỜI GIA ĐẾN

3.1. Ước tính hiệu thời gian đến dựa trên thuật toán GCC-PHAT-β

Để nâng cao chất lượng ước tính hiệu thời gian đến khi SNR nhỏ, một vài nghiên cứu đề xuất thuật toán GCC-PHAT- β, là một biến thể của GCC-PHAT [7]. Trong thuật toán này, hiệu thời gian đến được ước tính theo công thức [7];

$$\hat{\tau}_{ij} = \arg\max\left[\mathcal{F}^{-1}\left(\left[\frac{X_i(f)\left[X_j(f)\right]^*}{\left|X_i(f)\left[X_j(f)\right]^*\right|^{\beta}}\right]\right)\right] \tag{7}$$



Hình 2. *Mô phỏng ước tính hiệu thời gian đến.*

Thuật toán GCC-PHAT-β kiểm soát mức độ loại bỏ thành phần biên độ trong hàm trương quan chéo tổng quát thông qua hệ số β, khi $\beta=1$ toàn bộ thành phần biên độ bị loại bỏ, việc ước lượng chỉ sử dụng pha của tín hiệu tương đương với thuật toán GCC-PHAT [7]. Khi β=0, toàn bộ thành phần biên độ được giữ lại tương đương với phương pháp tương quan chéo kinh điển. Việc

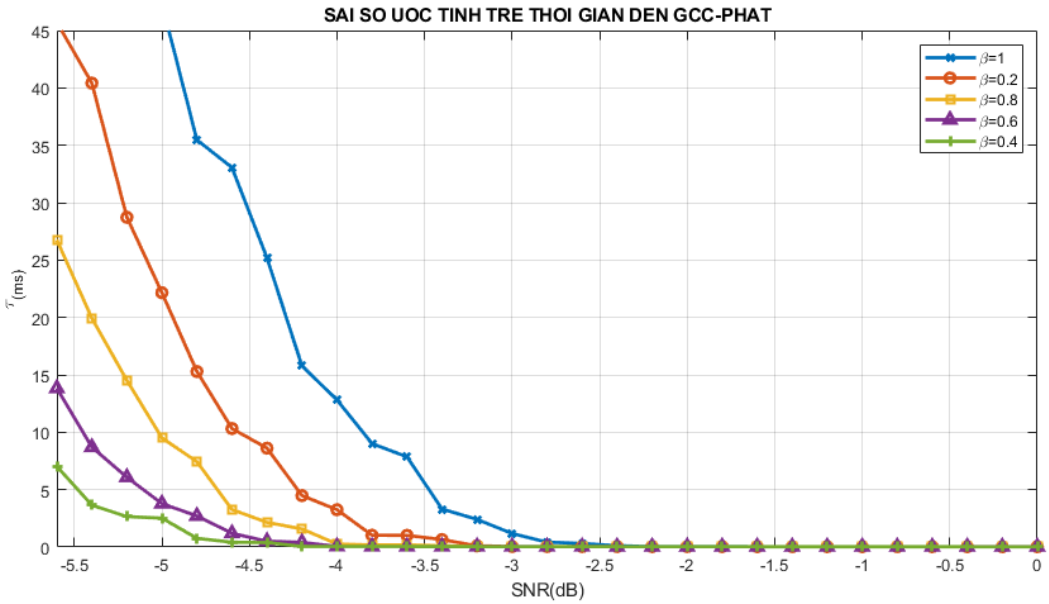
lựa chọn giá trị tham số β tối ưu phụ thuộc vào nhiều yếu tố như bản thân loại tín hiệu cần định vị, SNR của tín hiệu thu được trên các cảm biến. Để đánh giá ảnh hưởng của tham số β tới sai số ước tính hiệu thời gian đến, một chương trình mô phỏng được xây dựng trên Matlab.

Trên hình 2 thể hiện sơ đồ khối các bước thực hiện của chương trình mô phỏng. Trước hết tín hiệu có tồn tại sự kiện âm thanh trên cảm biến thứ i và j được tạo ra từ tín hiệu gốc, trong đó, hiệu thời gian đến τ_{ij} được thiết lập. Tín hiệu sau đó được bổ xung thành phần tạp âm và hệ số suy giảm theo công thức (1). Để ước tính hiệu thời gian đến của tín hiệu trên hai cảm biến trước hết cần biến đổi tín hiệu sang miền tần số bằng thuật toán FFT (Fast Fourier Transform), sau đó tính tương quan phổ năng lượng, bổ xung hàm trọng số trong miền tần số theo công thức (7). Biến đổi các giá trị thu được về miền thời gian theo thuật toán IFFT (Inverse Fast Fourier Transform), và xác định trễ thời gian đến τ_{ij} của hai tín hiệu theo công thức (4). Cuối cùng thực tính toán sai trễ thời gian đến theo công thức:

$$\varepsilon_\tau = \left\| \tau_{ij} - \hat{\tau}_{ij} \right\| \tag{8}$$

Tiến hành mô phỏng với SNR khác nhau nằm trong khoảng [-5.5÷0] dB, bước nhảy 0.2 dB, với mỗi giá trị SNR thực hiện 1000 mẫu thử, trong đó tín hiệu âm thanh được đưa vào mô phỏng là tiếng nổ đầu nòng của súng AK47 được lấy mẫu với tần số *f=20 kHz*, kết quả thu được sau đó được đưa vào đánh giá sai số. Kết quả trên hình 3 thể hiện sai số trễ thời gian đến trong các điều kiện SNR khác nhau với từng hệ số β. Có thể nhận thấy sai số hiệu thời gian đến không chỉ phụ thuộc vào tỉ số SNR mà còn phụ thuộc vào hệ số β, trong mỗi thời điểm với một giá trị SNR nhất định sẽ tồn tại một giá trị hệ số β tối ưu. Việc lựa chọn hệ số này trước đây dựa vào kinh nghiệm, hệ số thường cố định với các mức SNR khác nhau và không tối ưu trong mọi thời điểm.

Do đó, bài toán lựa chọn giá trị β tối ưu được đưa ra nhằm ước tính chính xác giá trị β thay đổi để sai số ước tính hiệu thời gian đến nhỏ nhất được cải thiện so với trường hợp sử dụng thuật toán GCC-PHAT.



Hình 3. *Sai số ước tính trễ thời gian đến với các hệ số β.*

3.2. Giải pháp lựa chọn hệ số β thích nghi theo tín hiệu thu được trên cảm biến

Như đã trình bày, việc lựa chọn hệ số β chính là lựa chọn tỉ lệ thành phần biên độ bị loại bỏ trong việc ước tính hiệu thời gian đến. Để xác định hệ số β thích hợp không phải là một việc dễ

dàng, bản thân hệ số này không những phụ thuộc vào SNR mà còn phụ thuộc vào bản thân tín hiệu cần định vị.

Tuy nhiên, trong bài toán định vị nguồn âm, trong đó, nguồn âm cần định vị được xác định trước ví dụ như định vị tiếng nổ đầu nòng của súng, pháo,... việc lựa chọn hệ số β có thể được thực hiện thông qua hệ số tương quan giữa tín hiệu thu được trên các cảm biến và tín hiệu mẫu.

Hệ số tương quan, ký hiệu là r được sử dụng để đo lương mức độ phụ thuộc tuyến tính của hai biến ngẫu nhiêu, ở đây, hai biến ngẫu nhiên bao gồm tín hiệu thu được trên cảm biến X_i và tín hiệu mẫu X_s, khi đó, hàm tương quan có dạng [8]:

$$r(m) = \frac{\sum_{n=1}^{N}\left(X_S(n)-\bar{X}_S\right)\left(X_i(n)-\bar{X}_i\right)}{\sqrt{\sum_{n=1}^{N}(X_s(n)-\bar{X}_s)^2}\sqrt{\sum_{n=1}^{N}(X_i(n)-\bar{X}_i)^2}} \tag{9}$$

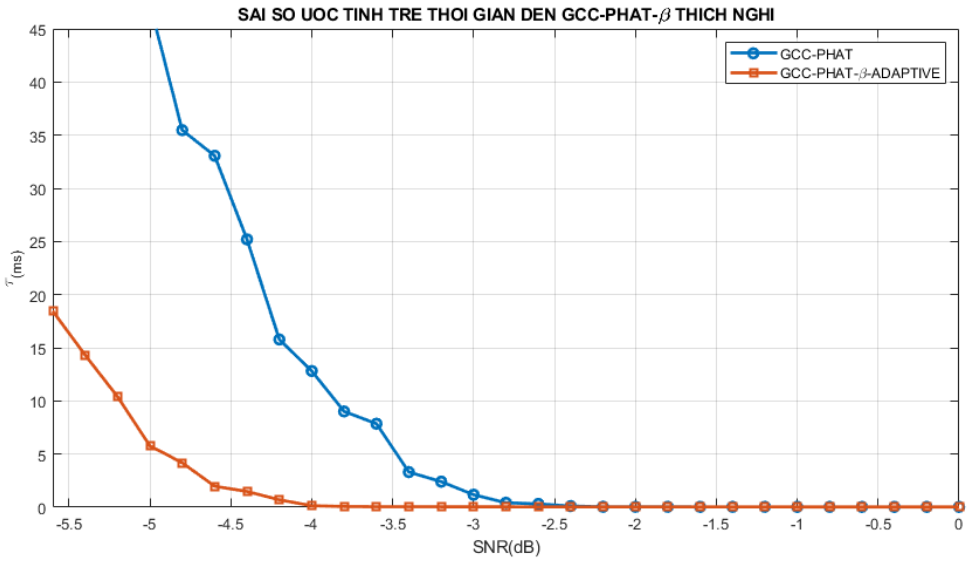
Trong đó, $\bar{X}_S$ và $\bar{X}_i$ lần lượt là giá trị trung bình các của tín hiệu mẫu và tín hiệu tại cảm biến thứ i. Dựa trên công thức (9) giá trị tương quan cực trị được xác định theo công thức [8]:

$$r = \arg\max_m\{r(m)\} \tag{10}$$

Hệ số tương quan r được tính toán có gia trị nằm trong khoảng [0,1], khi $r \to 1$ tín hiệu thu được trên cảm biến có tính tương quan cao với tín hiệu mẫu, từ đó có thể đưa ra kết luận rằng tỉ lệ SNR lớn, mặt khác khi $r \to 0$ tín hiệu thu được có tính tương quan thấp với tín hiệu mẫu tương ứng với SNR nhỏ. Qua đó có thể sử dụng công thức tính toán hệ số β theo công thức:

$$\beta = \frac{r_{is} + r_{js}}{2} \tag{11}$$

Trong đó, r_{is} và r_{js} lần lượt là giá trị tương quan giữa tín hiệu thu được trên cảm biến thứ i và j với tín hiệu mẫu.



Hình 4. *Sai số ước tính trễ thời gian đến với hệ số β thích nghi.*

Trên hình 4 thể hiện kết quả mô phỏng với tỉ lệ SNR khác nhau của phương pháp GCC-PHAT và phương pháp GCC-PHAT- β với giá trị hệ số β được ước tính theo tương quan. Với phương pháp GCC-PHAT khi tỉ số *SNR* > *-2.5 dB* sai số $\varepsilon_\tau \approx 0$, tuy nhiên, khi tỉ số *SNR* < *-2.5*

dB, sai số hiệu thời gian đến ε_τ tăng nhanh. So sánh với phương pháp GCC-PHAT phương pháp GCC-PHAT-β thích nghi cho thấy trong các điều kiện SNR khác nhau, phương pháp cải tiến đều có sai số nhỏ hơn, đặc biệt hiệu trong điều kiện SNR nhỏ.

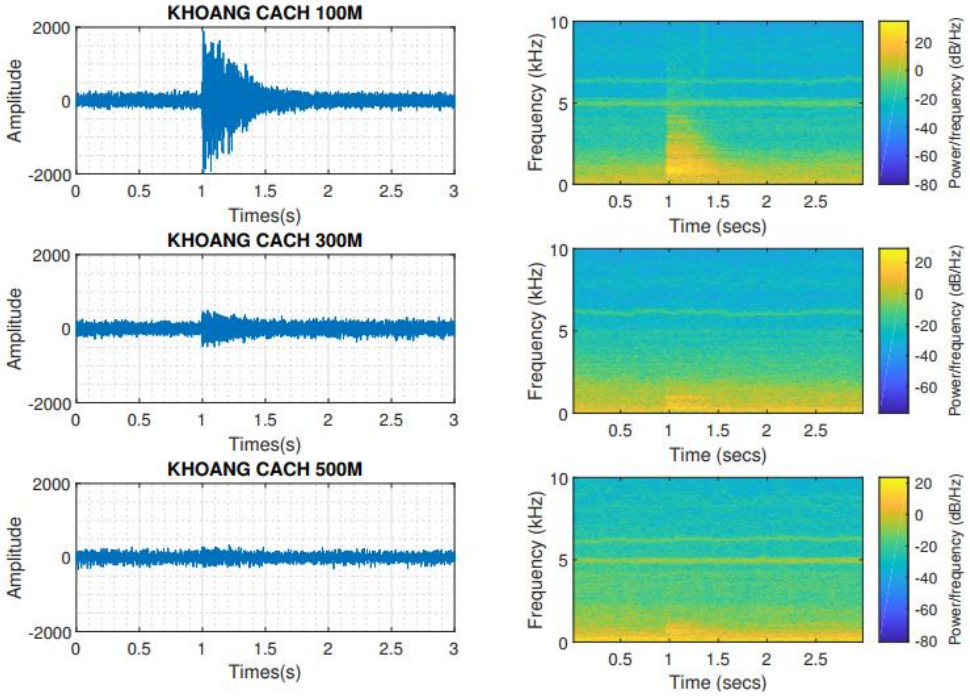
3.3. Một số kết quả thực nghiệm của giải pháp

Để đánh giá hiệu quả của của thuật toán GCC-PHAT- β thích nghi, một hệ 08 cảm biến âm thanh INMP401 của hãng InvenSense được sử dụng để thu nhận tín hiệu âm thanh, đây là loại cảm biến vi cơ tiên tiến có độ nhạy cao, kích thước nhỏ gọn. Những cảm biến này được bố trí trên hai hình tứ diện đều lồng nhau có độ dài cạnh *1 m* như trên hình 5, việc bố trí các cảm biến nằm trên tứ diện đều giúp tín hiệu thu được đồng đều trên các hướng tới [3].



Hình 5. *Cụm cảm biến thu tín hiệu âm thanh.*

Tiến hành phát tín hiệu âm thanh tạo giả tiếng nổ đầu nòng của súng AK47 qua một loa nén công suất *30 W* tại các cự ly khác nhau trong điều kiện môi trường thực tế. Tín hiệu sẽ được thu đồng thời trên 08 cảm biến, những tín hiệu này được số hóa và lưu trữ với tần số lấy mẫu *f_s=20 kHz* làm cơ sở cho các bước xử lý tín hiệu và định vị nguồn âm thanh.



Hình 6. *Tín hiệu âm thanh thu được tại các cự ly khác nhau.*

Hình 6 thể hiện dạng và giản đồ spectrogram của tín hiệu trên các cảm biến ở các cự ly khác nhau. Có thể nhận thấy ảnh hưởng của tạp âm lên tín hiệu tại các cự ly khác nhau là rất rõ ràng, tại cự ly nhỏ hơn *300 m* hình ảnh tín hiệu là tương đối rõ ràng và phân biệt với nền, mức độ năng lượng tương ứng với tần số âm thanh có thể nhận thấy tương đối rõ. Tuy nhiên, khi cự ly lớn hơn *300m* việc phân biệt tín hiệu và tạp âm tương đối khó khăn, phổ năng lượng tín hiệu cũng bị hòa lẫn vào năng lượng của tạp âm.

Kết quả ước tính hiệu thời gian đến trên bộ dữ liệu thu được dựa trên các thuật toán GCC-PHAT và GCC-PHAT-β thích nghi được thể hiện trên các bảng 1 và bảng 2.

Bảng 1. *Sai số hiệu thời gian đến tại cự ly 100 m.*

ε_τ (ms)	τ_{12}	τ_{13}	τ_{14}	τ_{15}	τ_{16}	τ_{17}	τ_{18}
GCC-PHAT (ms)	0.023	0.025	0.021	0.023	0.016	0.018	0.016
GCC-PHAT-β-TN (ms)	0.012	0.0011	0.011	0.010	0.009	0.0010	0.009

Có thể nhận thấy, việc sử dụng thuật toán GCC-PHAT-β thích nghi cải thiện sai số hiệu thời gian đến trung bình tại so với thuật toán GCC-PHAT truyền thống, đặc biệt khi cự ly tới nguồn âm xa điều này càng được thể hiện rõ ràng hơn.

Bảng 2. *Sai số hiệu thời gian đến tại cự ly 500 m.*

ε_τ (ms)	τ_{12}	τ_{13}	τ_{14}	τ_{15}	τ_{16}	τ_{17}	τ_{18}
GCC-PHAT (ms)	0.526	0.438	0.481	0.541	0.45	0.585	0.471
GCC-PHAT-β-TN (ms)	0.024	0.022	0.016	0.028	0.024	0.021	0.027

Việc cải thiện sai số hiệu thời gian đến trung bình cho thấy độ chính xác ước lượng hiệu thời gian đến được nâng cao, từ đó xây dựng được các hệ số chính xác cho hệ phương trình định vị nguồn âm 3, làm cơ sở cho việc nâng cao độ chính xác định vị nguồn âm.

4. KẾT LUẬN

Trên cơ sở phân tích lý thuyết và tiến hành mô phỏng trên Matlab, cho thấy phương pháp sử dụng thuật toán GCC-PHAT-β trong đó hệ số β được tính toán thích nghi theo giá trị tương quan giữa tín hiệu thu được trên các cảm biến và tín hiệu mẫu cho thấy tính hiệu quả trong việc nâng cao chất lượng ước tính hiệu thời gian đến, qua đó nâng cao chất lượng định vị nguồn âm sử dụng nguyên lý TDOA. Giải pháp được chứng minh thông qua chương trình mô phỏng trên Matlab, cũng như trên dữ liệu tín hiệu âm thanh thu được trên các cảm biến trong môi trường thực tế. Tuy nhiên, giải pháp nêu trên chỉ được áp dụng trong các bài toán định vị nguồn âm có tồn tại tín hiệu mẫu, đồng nghĩa với âm thanh cần định vị được xác định trước. Do đó, giải pháp đưa ra chỉ phù hợp với một lớp các ứng dụng nhất định, với các nguồn âm phức tạp và biến đổi liên tục như tiếng nói, nguồn tiếng ồn,... giải pháp đưa ra chưa có khả năng ứng dụng hiệu quả.

TÀI LIỆU THAM KHẢO

[1]. Phạm Văn Hòa, Nguyễn Văn Hiếu, Trần Quang Huy, Trần Công Thìn, *"Xây dựng hệ thống phát hiện và định vị hỏa lực bắn tỉa theo nguyên lý TDOA sóng âm"*, Tạp chí nghiên cứu KHCN quân sự, số 10 (2015), tr 88-95.

[2]. Trần Công Thìn, Bùi Ngọc Mỹ, Nguyễn Huy Hoàng, Phạm Văn Hòa, *''Xây dựng giải pháp định vị nguồn âm theo nguyên lý TDOA trong điều kiện vận tốc âm thanh biến đổi"*, Hội thảo Ứng dụng Công nghệ cao vào thực tiễn - 60 năm phát triển Viện KH-CN quân sự, 2020.

[3]. Antonio Canclini, *"A robust and low-complexity source localization algorithm for asynchronous distributed microphone networks"*, IEEE/ACM Transactions on Audio, Speech and Language Processing *(TASLP)*, Vol. 23, No.10 (2015), pp. 1563–1575.

[4]. Maximo Cobos, *"A Survey of Sound Source Localization Methods in Wireless Acoustic Sensor Networks"*, Wireless Communications and Mobile Computing, 2017, pp. 1–24.
[5]. Yiteng Huang, Jacob Benesty, and Jingdong Chen, *"Time Delay Estimation and Source Localization"*, Springer Handbook of Speech Processing, Springer Berlin Heidelberg, 2008, pp. 1043–1063.
[6]. C. Knapp and G. Carter, "*The generalized correlation method for estimation of time delay"*, IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. 24, No.4 (1976), pp. 320–327.
[7]. Kevin D. Donohue, Alvin Agrinsoni, and Jens Hannemann, *"Audio signal delay estimation using partial whitening"*, in: *Proceedings 2007 IEEE SoutheastCon*, 2007, pp. 466–471.
[8]. Volodymyr KHARCHENKO and Maryna MUKHINA (2017), *"Model of geophysical fields representation in problems of complex correlationextreme navigation"*, *Transport Problems*, 10 (3), pp. 35–41.
[9]. Cha Zhang, D. Florencio, and Zhengyou Zhang, *"Why does phat work well in lownoise, reverberative environments?*", 2008 IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 2565–2568, March 2008.

ABSTRACT

IMPROVING QUALITY OF TIME DELAY ESTIMATION
IN SOUND SOURCE LOCALIZATION USING TDOA METHOD

This paper presents several research results that enhance TDOA based sound localization accuracy with the priority of the source of interest. The GCC-PHAT algorithm is used to evaluate the estimated localization error. From that, we propose an adaptive method called GCC-PHAT-β to improve time differences estimated, where the parameter β is adapted to the input parameters in TDOA method. From analisys and discussions are made on that design and using Monte Carlo simulations with the data collected in real environment, the results show the efficiency of our proposed method in TDOA based localization (compared to conventional TDOA based localization).

Keywords: Sound Source Localization; TDOA; GCC-PHAT; GCC-PHAT- β.

Nhận bài ngày 17 tháng 9 năm 2021
Hoàn thiện ngày 20 tháng 10 năm 2021
Chấp nhận đăng ngày 28 tháng 10 năm 2021

Địa chỉ: [1]Viện Điện tử/Viện Khoa học và Công nghệ quân sự;
[2]Phòng Đào tạo, Viện Khoa học và Công nghệ quân sự;
[3]Học viện Kỹ thuật quân sự.
[*]Email: thin.vdt@outlook.com.