

ĐIỀU KHIỂN LOẠI BỎ PHẢN HỒI ÂM THANH SỬ DỤNG HAI MICROPHONE KẾT HỢP VỚI TIỀN BỘ LỌC VÀ NHIỀU DÒ

Trần Thị Thục Linh

Học Viện Công Nghệ Bưu Chính Viễn Thông

Tóm tắt: Một trong các vấn đề chính của máy trợ thính là phản hồi âm thanh (acoustic feedback). Phương pháp loại bỏ phản hồi âm thanh phổ biến hiện nay là dùng một bộ lọc thích nghi để ước lượng kênh phản hồi (feedback path). Tuy nhiên việc ước lượng kênh phản hồi thường có sai số sinh ra do sự tương quan giữa tín hiệu đầu vào mic và tín hiệu thu được tại loa. Phương pháp dự đoán lỗi sử dụng tiền bộ lọc kết hợp với máy trợ thính sử dụng 2 mic và 1 loa (PEM-AFC2) được biết đến là một phương pháp hiệu quả trong việc giảm sai số này. Để nâng cao hơn nữa hiệu suất của việc loại bỏ phản hồi âm thanh của phương pháp PEM-AFC2 bài báo này đề xuất một phương pháp mới được xây dựng dựa trên việc tích hợp nhiễu dò (probe noise) vào PEM-AFC2. Phương pháp mới không những cải thiện được hai chỉ số quan trọng là độ lệch chuẩn hóa (MIS) và độ khuếch đại ổn định gia tăng (ASG) mà còn duy trì được chất lượng tốt của âm thanh thu được.

Từ khóa: Loại bỏ phản hồi âm thanh, nhiễu dò, phương pháp tiền bộ lọc, phương pháp tiền bộ lọc kết hợp 2 mic.

I. GIỚI THIỆU

Một trong những vấn đề chính của máy trợ thính là hiện tượng phản hồi âm thanh (acoustic feedback) gây ra bởi tín hiệu từ loa quay trở lại mic khiến cho tín hiệu tại mic lúc này không chỉ là tín hiệu đầu vào mong muốn mà còn cộng thêm tín hiệu phản hồi từ loa. Tín hiệu này đi qua đường xử lý tín hiệu (the forward path $K(q)$) sẽ bị trễ đi và được khuếch đại trước khi tới loa. Quá trình trên lặp đi lặp lại tạo nên một hệ thống vòng kín (closed loop system). Phản hồi âm thanh làm giảm chất lượng của tín hiệu thu được đồng thời hạn chế độ khuếch đại có thể đạt được của máy trợ thính. Trong một số trường hợp nhất định phản hồi âm thanh có thể khiến hệ thống mất ổn định và “tiếng rú” (howling) có thể xuất hiện. Việc giải quyết vấn đề phản hồi âm thanh ngày càng trở nên thách thức khi nhu cầu về máy trợ thính có kích thước nhỏ ngày càng gia tăng.

Trong suốt 60 năm qua có rất nhiều phương pháp loại bỏ phản hồi âm thanh đã được giới thiệu [1]–[3], trong số đó các phương pháp loại bỏ phản hồi âm thanh thích nghi (Adaptive Feedback Cancellation-AFC) được biết đến

rộng rãi nhờ tính hiệu quả của nó. Phương pháp AFC chủ yếu dựa trên việc dùng một bộ lọc FIR để ước lượng đường phản hồi âm thanh mà sau đó được sử dụng để tính tín hiệu phản hồi âm thanh. Tín hiệu lỗi được tạo ra bằng cách lấy tín hiệu từ mic trừ đi tín hiệu phản hồi âm thanh. Đường phản hồi âm thanh được ước lượng thông qua việc tối thiểu hóa lỗi bình phương trung bình (MSE). Tuy nhiên việc ước lượng đường phản hồi âm thanh như vậy thường có sai số lớn do có sự tương quan cao giữa tín hiệu tại loa và tín hiệu tới mic (the incoming signal) [1], [2], [4], đặc biệt khi tín hiệu tới mic là tín hiệu phổ màu (spectrally colored) chẳng hạn như tiếng nói, âm nhạc, ... Để giảm sai số này nhiều phương pháp loại bỏ tương quan đã được đề xuất cho AFC chẳng hạn như chèn thêm trễ vào đường xử lý tín hiệu [1], [5], sử dụng dịch tần số [6], [7], điều chế pha [8], chèn nhiễu dò (PN) [9]–[12], hoặc sử dụng tiền bộ lọc [13], [14]. Một số phương pháp khác sử dụng các bộ lọc thích nghi có tốc độ hội tụ nhanh [15]–[17]; hay sử dụng thuật toán biến đổi bước [7], [18]–[20]. Trong [21] một phương pháp AFC dựa trên việc phân tách một bộ lọc thích nghi dài thành tích Kronecker của hai bộ lọc ngắn hơn đã được đề xuất. Các phương pháp AFC lai dựa trên việc chuyển mạch giữa hai thuật toán có tốc độ hội tụ khác nhau nhằm mục đích nâng cao tốc độ hội tụ của phương pháp lai trong khi vẫn duy trì được độ lệch chuẩn hóa thấp, ASG cao và chất lượng âm thanh tốt được giới thiệu trong [22], [23]. Trong các phương pháp trên thì phương pháp dự đoán lỗi (Prediction Error Method) cho AFC [14], [24]–[27], gọi là PEM-AFC, là phương pháp được sử dụng rộng rãi cho máy trợ thính sử dụng một mic và một loa.

Hiện nay nhiều phương pháp loại bỏ phản hồi âm thanh thích nghi sử dụng nhiều mic và một loa đã được nghiên cứu cho các ứng dụng như diễn thuyết (public address) [28], [29] hay máy trợ thính [11], [30]–[32]. Trong các phương pháp này một beamformer được sử dụng để giảm phản hồi âm thanh. Trong [33], [34] một phương pháp mới sử dụng 2 mic cho máy trợ thính (AFC2) được giới thiệu. Việc tích hợp PEM-AFC vào AFC2 được đề xuất trong [35], gọi là PEM-AFC2. Phương pháp này tận dụng được cả ưu điểm của việc sử dụng tiền bộ lọc để làm trắng (pre-whitening) tín hiệu đầu vào của bộ lọc thích nghi trong PEM-AFC và ưu điểm của việc khai thác thông tin không gian do vị trí tương đối giữa 2 mic của phương pháp AFC2. Kết quả là phương pháp PEM-AFC2 có sai số thấp hơn phương pháp AFC2 trong việc ước lượng đường phản hồi

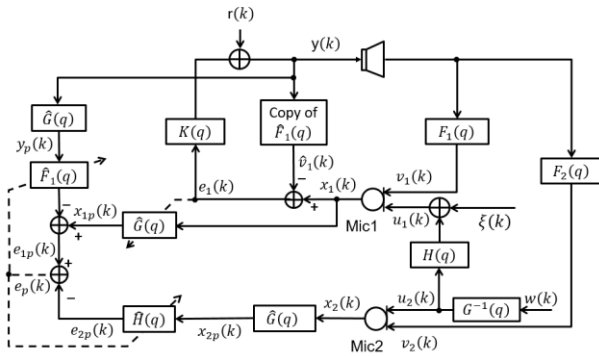
Tác giả liên hệ: Trần Thị Thục Linh,

Email: linhtt@ptit.edu.vn

Đến tòa soạn: 8/8/2022, chỉnh sửa: 10/10/2022, chấp nhận đăng: 22/10/2022.

âm thanh. Ngoài ra phương pháp này cũng có độ lệch (MIS) thấp hơn và độ khuếch đại ổn định gia tăng (ASG) cao hơn so với hai phương pháp thành phần là PEM-AFC và AFC2 đối với tín hiệu đầu vào là âm nhạc hoặc tiếng nói. Hơn nữa phương pháp PEM-AFC2 cũng cho phép duy trì chất lượng tín hiệu tốt. Để cải thiện hơn nữa hiệu suất của PEM-AFC2 tôi đề xuất một phương pháp mới dựa trên việc tích hợp nhiễu dò (PN) vào tín hiệu loa của hệ thống, gọi là PEM-AFC2-PN. Phương pháp được đề xuất do đó có khả năng làm giảm hơn nữa sự tương quan giữa tín hiệu loa và tín hiệu đầu vào mic dẫn đến làm giảm đáng kể sai số của việc ước lượng đường phản hồi âm thanh. Kết quả mô phỏng cho thấy phương pháp đề xuất có độ lệch trung bình (MIS) nhỏ hơn và độ khuếch đại ổn định gia tăng trung bình cao hơn (ASG) so với một số phương pháp tiên tiến được dùng để so sánh cho các trường hợp: TH1) đường phản hồi âm thanh không thay đổi (F1, F2), tín hiệu vào là tiếng nói; TH2) đường phản hồi âm thanh thay đổi đột ngột từ F1 sang F1-vật cản, từ F2 sang F2-vật cản sau 40 s, tín hiệu vào là tiếng nói; và TH3) giống TH2 nhưng tín hiệu vào là âm nhạc. Hơn nữa chất lượng của tín hiệu vẫn duy trì ở mức rất tốt (STOI > 0,98).

II. MÔ HÌNH HỆ THỐNG ĐỀ XUẤT



Hình 1. Mô hình hệ thống loại bỏ phần hồi âm thanh đề xuất (PEM-AFC2-PN) cho máy trợ thính

Hình 1 minh họa mô hình hệ thống loại bỏ phần hồi âm thanh đề xuất (PEM-AFC2-PN) cho máy trợ thính. Mô hình này dựa trên mô hình PEM-AFC2 [35], sự khác biệt là trong bài báo này nhiễu dò (PN), kí hiệu $r(k)$, được cộng thêm vào tín hiệu tại loa. Mô hình gồm có 2 mic, trong đó mic chính (mic 1) được đặt trong ống tai (ear canal) còn mic 2 được đặt phía sau tai (Behind The Ear – BTE). Loa được đặt gần với mic 1. Khoảng cách giữa 2 mic phải đủ xa để đảm bảo rằng tín hiệu đường phản hồi thứ 2 suy giảm nhiều hơn tín hiệu đường phản hồi thứ nhất [33], [34]. Mỗi mic nhận tín hiệu tới từ môi trường xung quanh và tín hiệu phản hồi, nghĩa là

$$x_1(k) = u_1(k) + \mathbf{f}_1^T \mathbf{y}(k), \tag{1}$$

$$x_2(k) = u_2(k) + \mathbf{f}_2^T \mathbf{y}(k), \tag{2}$$

trong đó $\mathbf{f}_i = [f_{i,0}, f_{i,1}, \dots, f_{i,L_f-1}]^T$ với $i=1,2$ là bộ lọc

FIR có chiều dài L_f dùng để mô hình hóa đường phản hồi âm thanh thứ i ; $\mathbf{y}(k)$ là vector có chiều dài L_f chi tín

hiệu loa và $u_i(k)$ là tín hiệu đầu vào của mic thứ i . Bộ lọc \mathbf{f}_i được biểu diễn dưới dạng $F_i(q) = \mathbf{f}_i^T \mathbf{q}$ với $\mathbf{q} = [1 \ q^{-1} \ \dots \ q^{-L_f+1}]^T$.

Giả thiết rằng tín hiệu đầu vào mic 2 là $u_2(k)$ có thể được mô hình hóa như sau

$$u_2(k) = G^{-1}(q)w(k), \tag{3}$$

trong đó $w(k)$ là dãy nhiễu Gauss trắng, trung bình không và $G^{-1}(q)$ là bộ lọc toàn cực có thể nghịch đảo được. Mối quan hệ giữa tín hiệu đầu vào mic 1 và mic 2 được giả thiết như sau

$$u_1(k) = \mathbf{h}^T \mathbf{u}_2(k) + \xi(k), \tag{4}$$

trong đó $\mathbf{h} = [h_0, h_1, \dots, h_{L_h-1}]^T$ là đáp ứng xung của hàm truyền đạt tương đối (Relative Transfer Function - RTF) $H(q)$ với chiều dài L_h và $\xi(k)$ là lỗi dư gây ra bởi dưới mô hình hóa (undermodeling) RTF. Giả thiết rằng với vị trí của các mic được sắp xếp như đã cho thì \mathbf{h} là nhân quả.

Tín hiệu mic và loa sau khi đi qua tiền bộ lọc $\hat{G}(q)$ được định nghĩa như sau

$$x_{1p}(k) = \hat{G}(q)x_1(k), \tag{5}$$

$$x_{2p}(k) = \hat{G}(q)x_2(k), \tag{6}$$

$$y_p(k) = \hat{G}(q)y(k), \tag{7}$$

trong đó $\hat{G}(q), x_1(k), x_2(k), y(k)$ lần lượt là ước lượng của bộ lọc $G(q)$, tín hiệu thu được trên mic 1, mic 2, và tín hiệu tại loa.

Tín hiệu lỗi sau khi được tiền làm trắng (pre-whitened) được tính như sau

$$e_p(k) = e_{1p}(k) - e_{2p}(k), \tag{8}$$

trong đó $e_{1p}(k) = x_{1p}(k) - \hat{F}_1(q)y_p(k)$ và

$e_{2p}(k) = \hat{H}(q)x_{2p}(k)$. Tín hiệu lỗi $e_p(k)$ được dùng để điều khiển quá trình tính toán thích nghi các tham số của bộ lọc $\hat{F}_1(q)$ và $\hat{H}(q)$.

Bằng cách tối thiểu hóa bình phương trung bình của tín hiệu lỗi $e_p(k)$, ta có thể ước lượng được các hệ số của bộ lọc $\hat{\mathbf{f}}_1$ và $\hat{\mathbf{h}}$ thông qua việc sử dụng thuật toán NLMS (Normalized Least Mean Squares) (xem phân tích chi tiết ở [35]) như sau

$$\hat{\mathbf{f}}_1(k) = \hat{\mathbf{f}}_1(k-1) + \frac{\mu}{\|\mathbf{y}_p(k)\|_2^2 + \delta} \mathbf{y}_p(k)e_p(k) \tag{9}$$

$$\hat{\mathbf{h}}(k) = \hat{\mathbf{h}}(k-1) + \frac{\mu}{\|\mathbf{x}_{2p}(k)\|_2^2 + \delta} \mathbf{x}_{2p}(k)e_p(k) \tag{10}$$

trong đó

$$\mathbf{y}_p(k) = [y_p(k), y_p(k-1), \dots, y_p(k-L_f+1)]^T,$$

$$\mathbf{x}_{2p}(k) = [x_{2p}(k), x_{2p}(k-1), \dots, x_{2p}(k-L_f+1)]^T$$

là các vector có chiều dài L_f ; δ là một giá trị rất nhỏ

được cộng vào mẫu số để tránh chia cho 0.

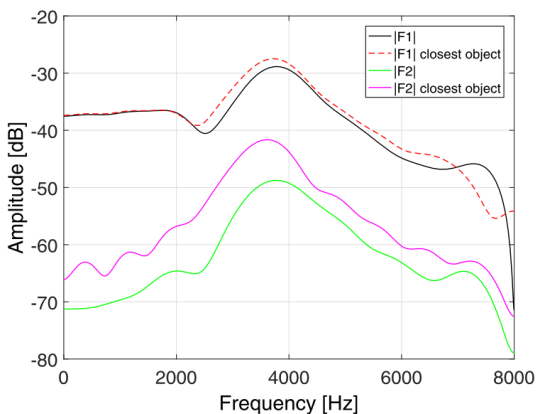
Tín hiệu loa được tạo ra bằng cách làm trễ rồi sau đó khuếch đại tín hiệu lỗi tại mic 1 ($e_1(k)$) trên đường xử lý tín hiệu và cuối cùng là cộng thêm nhiễu $r(k)$ như sau

$$y(k) = K(q)e_1(k) + r(k). \quad (11)$$

Giả thiết rằng đường xử lý tín hiệu gồm một bộ khuếch đại với hệ số khuếch đại bằng $|K|$ và một bộ làm trễ có $d_k \geq 1$ mẫu, thì $K(q) = |K|q^{-d_k}$. Một nhiễu $r(k)$ Gauss trắng có thể được cộng thêm vào tín hiệu loa như trình bày ở công thức (11) trước quá trình thích nghi để tăng cường việc giảm sự tương quan giữa tín hiệu vào mic 1 là $u_1(k)$ và tín hiệu loa $y(k)$. Kết quả là sai số trong việc ước lượng đường phản hồi âm thanh $\hat{\mathbf{f}}_1(k)$ giảm đi đáng kể, góp phần cải thiện các chỉ số độ lệch chuẩn hóa (MIS) và độ khuếch đại ổn định gia tăng (ASG). Tuy nhiên việc cộng thêm nhiễu $r(k)$ có thể gây ảnh hưởng đến chất lượng tín hiệu đầu ra tại loa nếu biên độ của nhiễu đủ lớn so với biên độ của tín hiệu. Thực nghiệm cho thấy rằng nếu ta cộng vào tín hiệu loa một nhiễu có tỷ số SNR = 30 dB thì không những không làm ảnh hưởng đến chất lượng của tín hiệu đầu ra mà còn làm giảm được đáng kể \overline{MIS} đồng thời làm tăng \overline{ASG} .

III. KẾT QUẢ MÔ PHỎNG

Hình 2 biểu diễn phổ biên độ của các đường phản hồi âm thanh được sử dụng trong các mô phỏng. Trong đó các đường phản hồi âm thanh F1 và F2 tương ứng với phản hồi âm thanh từ loa tới mic 1 và mic 2, được đo trong trường hợp không có vật cản ở giữa các mic và loa (bình thường); còn các đường phản hồi âm thanh F1-vật cản (closest object) và F2-vật cản, được đo khi đặt một vật cản ở gần tai để tạo ra vật cản giữa các mic và loa. Ta có thể dễ dàng thấy rằng phổ biên độ của đường phản hồi âm thanh tăng lên khi có vật cản giữa các mic và loa.



Hình 2. Phổ biên độ của các đường phản hồi âm thanh

Các tham số độ lệch chuẩn hóa (MIS) và độ khuếch đại ổn định gia tăng (ASG) được sử dụng để đánh giá hiệu suất của các phương pháp loại bỏ phản hồi âm thanh. Giá trị của MIS càng nhỏ và giá trị của ASG càng cao thì càng tốt. Công thức tính độ lệch chuẩn hóa và độ khuếch đại ổn định gia tăng cho việc ước lượng $\hat{F}_1(q)$ được định nghĩa như sau trong miền tần số chuẩn hóa ω :

$$MIS = 10 \log \left(\frac{\int_0^\pi |F_1(e^{j\omega}) - e^{-j\omega d_b} \hat{F}_1(e^{j\omega})|^2 d\omega}{\int_0^\pi |F_1(e^{j\omega})|^2 d\omega} \right) \quad (12)$$

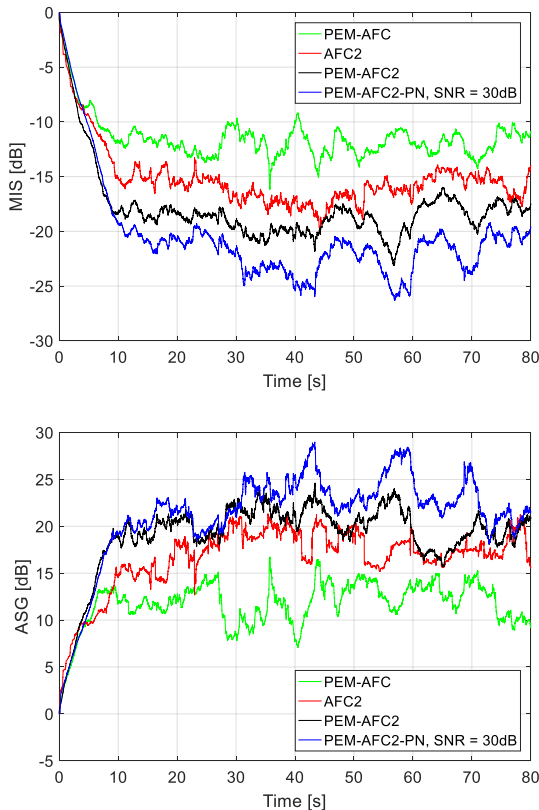
$$ASG = 10 \log \left(\min_{\omega} \frac{1}{|F_1(e^{j\omega}) - e^{-j\omega d_b} \hat{F}_1(e^{j\omega})|^2} \right) - 10 \log \left(\min_{\omega} \frac{1}{|F_1(e^{j\omega})|^2} \right), \quad (13)$$

trong đó d_b là độ trễ trên đường loại bỏ phản hồi âm thanh; $F_1(e^{j\omega})$ và $\hat{F}_1(e^{j\omega})$ lần lượt là các đáp ứng tần số của đường phản hồi âm thanh (F1) thực tế và ước lượng. Các tham số sau được lựa chọn cho tất cả các trường hợp mô phỏng: độ trễ trên đường xử lý tín hiệu $d_k = 32$ mẫu, độ khuếch đại của đường xử lý tín hiệu $|K| = 30$ dB, độ trễ trên đường loại bỏ phản hồi âm thanh $d_b = 16$ mẫu, tần số lấy mẫu $f_s = 16$ kHz, SNR = 30 dB và $\delta = 10^{-6}$. Chiều dài của $F_1(q)$ và $\hat{F}_1(q)$ lần lượt là $L_f = 38, L_f = 22$. Ta chọn $L_h = L_{\hat{h}} = 10$. Đối với các phương pháp AFC sử dụng hai mic các hệ số bước (step-size) để cập nhật các bộ lọc $\hat{F}_1(q)$ và $\hat{H}(q)$ được chọn bằng nhau. Độ lệch chuẩn hóa trung bình (\overline{MIS}) và độ khuếch đại ổn định gia tăng trung bình (\overline{ASG}) được tính trên toàn bộ thời gian mô phỏng (80 s) trong trường hợp 1 (TH1) và được tách thành $\overline{MIS}_i, \overline{ASG}_i$ (thời gian tính là 40 s) với $i=1,2$ tương ứng đường phản hồi bình thường và có vật cản như trong trường hợp 2, 3 (TH2, TH3). Trong trường hợp 1 và 2 (TH1 và TH2) tín hiệu đầu vào các mic là tiếng nói được ghi âm bằng 2 mic với thiết kế như trong [34]. Các nguồn tiếng nói được xây dựng dựa trên 30 câu IEEE được nói bởi 3 nam và 3 nữ lấy từ cơ sở dữ liệu NOIZEUS [36]. Cụ thể là ta ghép 30 câu này với nhau để tạo ra tín hiệu tiếng nói đầu vào mic. Các hệ số bước (step-size) được lựa chọn sao cho các phương pháp AFC đều có tốc độ hội tụ ban đầu tương đương nhau. Cụ thể, hệ số bước $\mu = 0,001$ được chọn cho AFC2, trong khi đó $\mu = 0,0005$ được chọn cho các phương pháp còn lại. Ta sử dụng một mô hình hồi quy tự động (AR model) bậc 20 của tín hiệu vào được tính toán cho tất cả các khung chiều dài 160 mẫu sử dụng thuật toán Levinson-Durbin [37]. Ngoài ra để đánh giá chất lượng của tiếng nói thu được tại loa ta sử dụng hệ số STOI (Short-Time Objective Intelligibility). Hệ số STOI nằm giữa 0 và 1 dùng để xác định độ dễ hiểu của tiếng nói, giá trị của STOI càng cao thì chất lượng tiếng nói càng tốt.

Đối với trường hợp 3 (TH3) ta chọn 1 đoạn trong bài hát "Imagine" của John Lennon để làm tín hiệu đầu vào mic.

Tiếp theo đây hiệu suất của phương pháp đề xuất sẽ được phân tích cụ thể cho từng trường hợp trong sự so sánh với các phương pháp tiên tiến khác.

Trường hợp 1 (TH1): trong trường hợp này ta so sánh hiệu suất của phương pháp đề xuất (PEM-AFC2-PN) với các phương pháp tiên tiến gồm PEM-AFC, AFC2, PEM-AFC2 với tín hiệu đầu vào các mic là tiếng nói. Các đường phản hồi âm thanh F1, F2 (bình thường, không có vật cản) được sử dụng trong các phương pháp dùng 2 mic, còn F1 được sử dụng cho phương pháp PEM-AFC.



Hình 3. So sánh phương pháp đề xuất (PEM-AFC2-PN) với một số phương pháp AFC tiên tiến; Tín hiệu vào là tiếng nói, đường phản hồi F1 và F2 (bình thường), SNR=30 dB.

Hình 3 cho thấy phương pháp đề xuất có MIS thấp hơn đáng kể và ASG cao hơn so với các phương pháp còn lại. Tốc độ hội tụ của các phương pháp là tương đương.

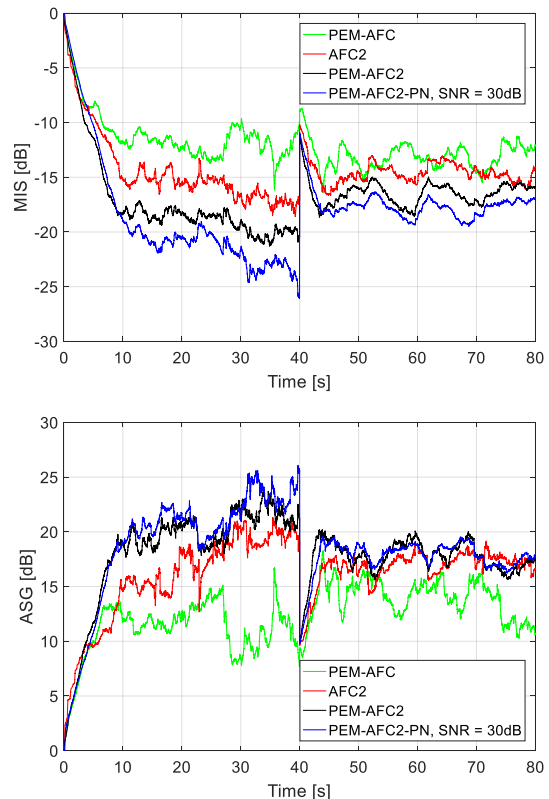
Bảng 1 cũng chỉ ra rằng phương pháp đề xuất có hệ số MIS được cải thiện khoảng 2,64 dB so với PEM-AFC2; 5,31 dB so với AFC2 và lên tới khoảng 9 dB so với PEM-AFC. Hệ số ASG cũng tăng khoảng hơn 2 dB so với PEM-AFC2; 4,59 dB so với AFC2 và lên tới khoảng 9,5 dB so với PEM-AFC. Trong khi đó hệ số chất lượng tín hiệu STOI của phương pháp đề xuất là rất cao (0,985) và tương đương với các phương pháp còn lại.

Bảng 1. So sánh phương pháp đề xuất (PEM-AFC2-PN) với một số phương pháp AFC tiên tiến dựa trên các hệ số MIS , ASG và $STOI$.

Phương pháp AFC	MIS	ASG	$STOI$
PEM-AFC	-11,50	11,75	0,985
AFC2	-15,23	16,67	0,988
PEM-AFC2	-17,92	19,06	0,988
PEM-AFC2-PN	-20,54	21,26	0,985

Trường hợp 2 (TH2): trường hợp này giống TH1, chỉ khác là các đường phản hồi âm thanh được chuyển đột ngột từ F1 sang F1-vật cản, từ F2 sang F2-vật cản sau 40 s. Hình 4 cho thấy PEM-AFC2-PN có tốc độ hội tụ tương đương với các phương pháp dùng để so sánh khác, nhưng phương pháp này có khả năng theo dấu (tracking) sự thay đổi của đường phản hồi âm thanh nhanh hơn hai phương pháp PEM-AFC và AFC2 và tương đương với PEM-AFC2. Ngoài ra phương pháp đề xuất có MIS thấp hơn đáng kể và ASG cao hơn so với các phương pháp còn lại.

Bảng 2 đưa ra so sánh phương pháp đề xuất với một số phương pháp AFC tiên tiến. Trong đó các chỉ số của MIS_1 , ASG_1 và $STOI_1$ với $i=1,2$ tương ứng đường phản hồi bình thường và có vật cản. Ta có thể thấy rõ rằng hệ số MIS của PEM-AFC2-PN có mức cải thiện tương ứng 1,64 dB; 4,48 dB và 7,74 dB so với PEM-AFC2, AFC2, và PEM-AFC trước khi có sự chuyển đột ngột của đường phản hồi âm thanh. Hệ số ASG cũng tăng khoảng 0,78 dB so với PEM-AFC2; 3,37 dB so với AFC2 và lên tới khoảng 8 dB so với PEM-AFC. Sau khi có sự thay đổi đột ngột của đường phản hồi âm thanh, hiệu suất của phương pháp đề xuất được xác định thông qua hệ số MIS_2 , ASG_2 vẫn tốt hơn nhiều so với các phương pháp còn lại. Hơn nữa phương pháp đề xuất có hệ số $STOI$ rất cao ($> 0,98$) và tương đương các phương pháp khác cho cả trước và sau khi đường phản hồi âm thanh chuyển đột ngột. Điều này chứng tỏ rằng chất lượng âm thanh vẫn được duy trì ở mức rất tốt mặc dù nhiều dò được sử dụng.



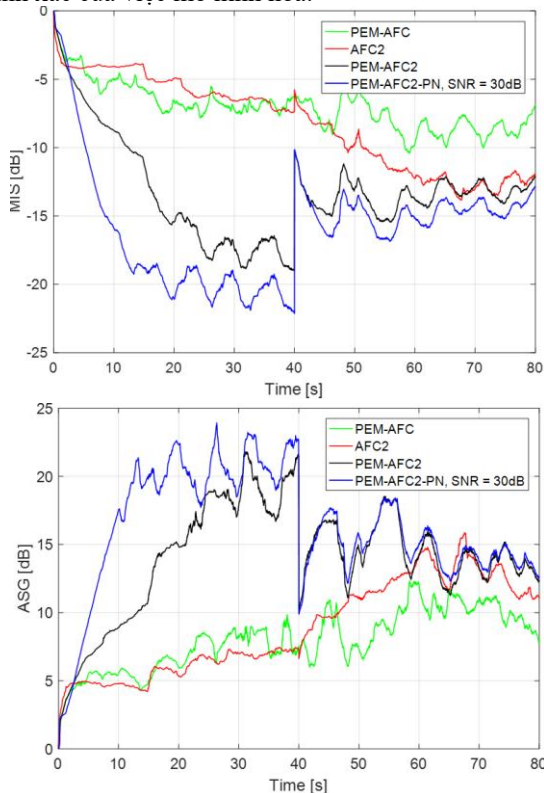
Hình 4. So sánh phương pháp đề xuất (PEM-AFC2-PN) với một số phương pháp AFC tiên tiến; Tín hiệu vào là tiếng nói, các đường phản hồi chuyển đột ngột từ F1 sang F1-vật cản và từ F2 sang F2-vật cản sau 40 s, SNR=30 dB.

Bảng 2. So sánh phương pháp đề xuất (PEM-AFC2-PN) với một số phương pháp AFC tiên tiến, tín hiệu đầu vào là tiếng nói. Các

chỉ số của \overline{MIS}_i , \overline{ASG}_i và \overline{STOI}_i với $i=1,2$ tương ứng các đường phân hồi bình thường và có vật cản.

Phương pháp AFC	\overline{MIS}_1	\overline{ASG}_1	\overline{MIS}_2	\overline{ASG}_2	\overline{STOI}_1	\overline{STOI}_2
PEM-AFC	-11,02	10,97	-13,07	13,90	0,987	0,984
AFC2	-14,28	15,60	-14,56	16,72	0,990	0,986
PEM-AFC2	-17,12	18,19	-16,44	17,67	0,990	0,986
PEM-AFC2-PN	-18,76	18,97	-17,54	17,73	0,986	0,981

Trường hợp 3 (TH3): trong trường hợp này tín hiệu vào mic là âm nhạc; các đường phân hồi âm thanh được chuyển đột ngột từ F1 sang F1-vật cản, từ F2 sang F2-vật cản sau 40 s. Bậc của mô hình hồi quy tự động (AR model) được chọn là 51. Do việc mô hình hóa tín hiệu đầu vào mic là âm nhạc khó hơn so với tín hiệu tiếng nói, trong trường hợp này ta chọn bậc của mô hình AR cao hơn để nâng cao độ chính xác của việc mô hình hóa.



Hình 5. So sánh phương pháp đề xuất (PEM-AFC2-PN) với một số phương pháp AFC tiên tiến; Tín hiệu vào là âm nhạc, các đường phân hồi chuyển đột ngột từ F1 sang F1-vật cản và từ F2 sang F2-vật cản sau 40 s, SNR=30 dB.

Từ

Hình 5 ta có thể thấy rằng phương pháp đề xuất có các hệ số MIS và ASG vượt trội so với các phương pháp còn lại. Ngoài ra phương pháp đề xuất cũng có tốc độ hội tụ và khả năng theo dấu sự thay đổi của đường phân hồi âm thanh nhanh hơn. Kết quả này hoàn toàn thống nhất với kết quả ở Bảng 3.

Bảng 3 cho thấy rằng trước khi có sự chuyển đột ngột của đường phân hồi âm thanh thì hệ số \overline{MIS}_1 của PEM-AFC2-PN giảm khoảng 4 dB; 11,5 dB và 11 dB tương ứng so với các phương pháp PEM-AFC2, AFC2, và PEM-AFC. Trong khi đó hệ số \overline{ASG}_1 tăng tương ứng khoảng 4 dB; 11,8 dB và 11 dB. Sau khi đường phân hồi âm thanh chuyển sang F1-vật cản và F2-vật cản thì hệ số \overline{MIS}_2 của PEM-AFC2-PN giảm khoảng 1,26 dB; 3,53 dB và 6,59 dB tương

ứng so với các phương pháp PEM-AFC2, AFC2, và PEM-AFC. Hệ số \overline{ASG}_2 cũng tăng tương ứng khoảng 0,44 dB; 2,84 dB và 5,34 dB.

Bảng 3. So sánh phương pháp đề xuất (PEM-AFC2-PN) với một số phương pháp AFC tiên tiến, tín hiệu đầu vào là âm nhạc. Các chỉ số của \overline{MIS}_i , \overline{ASG}_i và \overline{STOI}_i với $i=1,2$ tương ứng đường phân hồi bình thường và có vật cản.

Phương pháp AFC	\overline{MIS}_1	\overline{ASG}_1	\overline{MIS}_2	\overline{ASG}_2
PEM-AFC	-5,99	6,54	-8,10	9,45
AFC2	-5,30	5,82	-11,16	11,95
PEM-AFC2	-13,05	13,73	-13,43	14,35
PEM-AFC2-PN	-17,07	17,64	-14,69	14,79

IV. KẾT LUẬN

Trong bài báo này một phương pháp loại bỏ phản hồi âm thanh mới được đề xuất cho máy trợ thính. Phương pháp này tích hợp nhiều dò lên mô hình PEM-AFC2 nhằm mục đích giảm nhỏ hơn nữa sự tương quan giữa tín hiệu vào mic và tín hiệu loa. Kết quả mô phỏng chỉ ra rằng nếu ta cộng một nhiễu dò có SNR thích hợp (ví dụ SNR=30 dB) thì có thể nâng cao được hiệu suất của hệ thống trong khi đó vẫn duy trì được chất lượng tín hiệu đầu ra tốt. Cụ thể là phương pháp đề xuất cải thiện được đáng kể các hệ số độ lệch chuẩn hóa và độ khuếch đại ổn định gia tăng cũng như khả năng theo dấu sự thay đổi của đường phân hồi âm thanh so với các phương pháp AFC tiên tiến khác. Chất lượng tín hiệu đầu ra của phương pháp đề xuất cũng rất tốt ($\overline{STOI} > 0,98$) và tương đương với các phương pháp dùng để so sánh.

LỜI CẢM ƠN

Nghiên cứu này được tài trợ bởi Học viện Công nghệ bưu chính Viễn thông thông qua đề tài cấp Học viện “Nghiên cứu, đề xuất phương pháp loại bỏ phản hồi âm thanh sử dụng nhiều microphone cho máy trợ thính”.

REFERENCES

- [1] M. G. Siqueira and A. Alwan, “Steady-state analysis of continuous adaptation in acoustic feedback reduction systems for hearing-aids,” *IEEE Trans. Speech Audio Process.*, vol. 8, no. 4, pp. 443–453, Jul. 2000.
- [2] A. Spriet, S. Doclo, M. Moonen, and J. Wouters, “Feedback control in hearing aids,” in *Springer Handbook Speech Process.* New York, NY, USA: Springer, 2008, pp. 979–1000.
- [3] T. van Waterschoot and M. Moonen, “Fifty Years of Acoustic Feedback Control: State of the Art and Future Challenges,” *Proc. IEEE*, vol. 99, no. 2, pp. 288–327, Feb. 2011.
- [4] J. Hellgren and U. Forssell, “Bias of feedback cancellation algorithms in hearing aids based on direct closed loop identification,” *IEEE Trans. Speech Audio Process.*, vol. 9, no. 8, pp. 906–913, Nov. 2001.
- [5] S. Laugesen, K. Hansen, and J. Hellgren, “Acceptable delays in hearing aids and implications for feedback cancellation,” *J. Acoust. Soc. Amer.*, vol. 105, no. 2, pp. 1211–1212, 1999.
- [6] M. Schroeder, “Improvement of acoustic-feedback stability by frequency shifting,” *J. Acoust. Soc. Amer.*, vol. 36, no. 9, pp. 1718–1724, 1964.
- [7] F. Strasser and H. Puder, “Adaptive feedback cancellation for realistic hearing aid applications,” *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 23, no. 12, pp. 2322–2333, Dec. 2015.

- [8] M. Guo, S. H. Jensen, J. Jensen, and S. L. Grant, "On the use of a phase modulation method for decorrelation in acoustic feedback cancellation," in *Proc. 20th Eur. Signal Process. Conf.*, 2012, pp. 2000–2004.
- [9] J. Kates, "Feedback cancellation in hearing aids," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 1990, pp. 1125–1128.
- [10] M. Guo, S. H. Jensen, and J. Jensen, "Novel acoustic feedback cancellation approaches in hearing aid applications using probe noise and probe noise enhancement," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 20, no. 9, pp. 2549–2563, 2012.
- [11] M. Guo, T. B. Elmedyby, S. H. Jensen, and J. Jensen, "On acoustic feedback cancellation using probe noise in multiple-microphone and single-loudspeaker systems," *IEEE Signal Process. Lett.*, vol. 19, no. 5, pp. 283–286, May 2012.
- [12] C. R. C. Nakagawa, S. Nordholm, and W.-Y. Yan, "Feedback cancellation with probe shaping compensation," *IEEE Signal Process. Lett.*, vol. 21, no. 3, pp. 365–369, Mar. 2014.
- [13] J. Hellgren, "Analysis of feedback cancellation in hearing aids with filtered-X LMS and the direct method of closed loop identification," *IEEE Trans. Speech Audio Process.*, vol. 10, no. 2, pp. 119–131, Feb. 2002.
- [14] A. Spriet, I. Proudler, M. Moonen, and J. Wouters, "Adaptive feedback cancellation in hearing aids with linear prediction of the desired signal," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3749–3763, Oct. 2005.
- [15] L. T. T. Tran, H. Schepker, S. Doclo, H. H. Dam, and S. Nordholm, "Proportionate NLMS for adaptive feedback control in hearing aids," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2017, pp. 211–215.
- [16] Lee, K.; Baik, Y.h.; Park, Y.; Kim, D.; Sohn, J. Robust adaptive feedback canceller based on modified pseudo affine projection algorithm. In Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Boston, MA, USA, 30 August–3 September 2011; pp. 3760–3763.
- [17] Pradhan, S.; Patel, V.; Patel, K.; Maheshwari, J.; George, N.V. Acoustic feedback cancellation in digital hearing aids: A sparse adaptive filtering approach. *Appl. Acoust.* 2017, 122, 138–145.
- [18] L. T. T. Tran, H. Schepker, S. Doclo, H. H. Dam, and S. Nordholm, "Improved practical variable step-size algorithm for adaptive feedback control in hearing aids," in *Proc. IEEE 10th Int. Conf. Signal Process. Commun. Syst.*, 2016, pp. 1–8.
- [19] Rotaru, M.; Albu, F.; Coanda, H. A variable step size modified decorrelated NLMS algorithm for adaptive feedback cancellation in hearing aids. In Proceedings of the International Symposium on Electronics and Telecommunications, Timisoara, Romania, 15–16 November 2012; pp. 263–266.
- [20] Albu, F.; Tran, L.T.T.; Nordholm, S. A combined variable step size strategy for two microphones acoustic feedback cancellation using proportionate algorithms. In Proceedings of the Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), Kuala Lumpur, Malaysia, 12–15 December 2017; pp. 1373–1377.
- [21] Bhattacharjee, S.S.; George, N.V. Fast and Efficient Acoustic Feedback Cancellation based on Low Rank Approximation. *Signal Process.* 2021, 182, 107984.
- [22] Nordholm, S.; Schepker, H.; Tran, L.T.T.; Doclo, S. Stability-controlled hybrid adaptive feedback cancellation scheme for hearing aids. *J. Acoust. Soc. Am.* 2018, 143, 150–166.
- [23] Linh T. T. Tran and S. E. Nordholm, "A Switched Algorithm for Adaptive Feedback Cancellation Using Pre-Filters in Hearing Aids," *Journal of Audiology Research*, vol. 11, no. 3, pp. 389–409, 2021.
- [24] J. M. Gil-Cacho, T. van Waterschoot, M. Moonen, and S. H. Jensen, "Transform domain prediction error method for improved acoustic echo and feedback cancellation," in *Proc. IEEE 20th Eur. Signal Process. Conf.*, 2012, pp. 2422–2426.
- [25] L. T. T. Tran, H. H. Dam, and S. Nordholm, "Affine projection algorithm for acoustic feedback cancellation using prediction error method in hearing aids," in *Proc. IEEE Int. Workshop Acoust. Signal Enhancement*, 2016, pp. 1–5.
- [26] G. Bernardi, T. van Waterschoot, J. Wouters, and M. Moonen, "An all-frequency-domain adaptive filter with PEM-based decorrelation for acoustic feedback control," in *Proc. IEEE Workshop Appl. Signal Process. Audio Acoust.*, 2015, pp. 1–5.
- [27] G. Bernardi, T. van Waterschoot, J. Wouters, M. Hillbratt, and M. Moonen, "A PEM-based frequency-domain Kalman filter for adaptive feedback cancellation," in *Proc. IEEE 23rd Eur. Signal Process. Conf.*, 2015, pp. 270–274.
- [28] G. Rombouts, A. Spriet, and M. Moonen, "Generalized sidelobe canceller based combined acoustic feedback-and noise cancellation," *Signal Process.*, vol. 88, no. 3, pp. 571–581, 2008.
- [29] G. Rombouts, T. van Waterschoot, K. Struyve, and M. Moonen, "Acoustic feedback cancellation for long acoustic paths using a nonstationary source model," *IEEE Trans. Signal Process.*, vol. 54, no. 9, pp. 3426–3434, Sep. 2006.
- [30] M. Guo, T. B. Elmedyby, S. H. Jensen, and J. Jensen, "Analysis of acoustic feedback/echo cancellation in multiple-microphone and single-loudspeaker systems using a power transfer function method," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 5774–5788, Dec. 2011.
- [31] Schepker, H.; Nordholm, S.E.; Tran, L.T.T.; Doclo, S. Null-steering beamformer-based feedback cancellation for multi-microphone hearing aids with incoming signal preservation. *IEEE/ACM Trans. Audio Speech Lang. Process.* 2019, 27, 679–691.
- [32] Schepker, H.; Nordholm, S.; Doclo, S. Acoustic Feedback Suppression for Multi-Microphone Hearing Devices Using a Soft-Constrained Null-Steering Beamformer. *IEEE/ACM Trans. Audio Speech Lang. Process.* 2020, 28, 929–940. pp. 1–5.
- [33] C. R. C. Nakagawa, S. Nordholm, and W.-Y. Yan, "Dual microphone solution for acoustic feedback cancellation for assistive listening," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2012, pp. 149–152.
- [34] C. R. C. Nakagawa, S. Nordholm, and W.-Y. Yan, "Analysis of two microphone method for feedback cancellation," *IEEE Signal Process. Lett.*, vol. 22, no. 1, pp. 35–39, Jan. 2015.
- [35] L. T. T. Tran, S. Nordholm, H. Schepker, H. Dam, and S. Doclo, "Two microphone hearing aids using prediction error method for adaptive feedback control," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 26, no. 5, pp. 909–923, May 2018.
- [36] P. C. Loizou, *Speech Enhancement: Theory and Practice*. Boca Raton, FL, USA: CRC Press, 2013.
- [37] J. R. Deller Jr., J. G. Proakis, and J. H. Hansen, *Discrete Time Processing of Speech Signals*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1993.

ACOUSTIC FEEDBACK CANCELLATION USING TWO MICROPHONES WITH PRE-FILTER AND PROBE NOISE

Abstract: A major problem in open-fit digital hearing aids is acoustic feedback. A popular acoustic feedback cancellation approach uses an adaptive filter to estimate the feedback path. However, the estimation of the feedback path often produces a bias due to a high correlation between the incoming signal and the loudspeaker signal. The prediction error method for a hearing aid with two-microphone and one loudspeaker (PEM-AFC2) is well known for effectively reducing this bias. To further improve the performance of PEM-AFC2, in this paper we propose a new approach which integrates probe noise into the PEM-AFC2. The proposed approach not only improves two important metrics (Normalized Misalignment and Added Stable Gain) but also maintains the good quality of the received signal.

Keywords: Adaptive feedback cancellation, probe noise, prediction error method, PEM-AFC2.



Trần Thị Thục Linh tốt nghiệp đại học ngành Điện tử-Viễn thông của Đại học Bách Khoa Hà Nội, thạc sĩ ngành Truyền thông số của Đại học tổng hợp Kiel, CHLB Đức năm 2005, và tiến sĩ ngành Điện-Công nghệ máy tính của Đại học Curtin, Perth, Úc năm 2018. Từ năm 2019 đến 2020 cô ấy là nghiên cứu sinh sau tiến sĩ tại Đại học Công nghệ Nanyang, Singapore.

Hiện tại cô ấy là giảng viên tại Khoa Công nghệ Điện và Điện tử thuộc Học viện Công nghệ Bưu chính Viễn thông. Hướng nghiên cứu chính bao gồm xử lý tín hiệu âm thanh và tiếng nói, loại bỏ phản hồi âm thanh, đặc biệt là cho các ứng dụng liên quan đến máy trợ thính.