

THUẬT TOÁN ĐÁNH GIÁ BỀN VỮNG LEVENBERG-MARQUARDT TRONG VẤN ĐỀ KHỬ TIẾNG VỌNG ÂM HỌC

NGUYỄN KỲ TÀI

Khoa điện, Đại học Bách khoa Tp.HCM

Abstract. In this paper, we propose a Levenberg-Marquardt algorithm for combined stereophonic acoustic echo cancellation. The algorithm is developed on the basis of a least mean-square (LMS) error criterion. The adaptive algorithm that has an updating value which is the correlation of the input signal to the weights and the estimation error is derived. Simulation results in Matlab shows that Levenberg-Marquardt algorithm has been successfully applied to acoustic echo cancellation.

Tóm tắt. Trong bài báo này, thuật toán L-M được ứng dụng trong việc khử tiếng vọng trong hệ truyền âm stereo. Thuật toán này được phát triển dựa trên tiêu chuẩn đánh giá bình phương tối thiểu. Thuật toán thích nghi tính toán cập nhật các hệ số của tín hiệu đầu vào và sai số ngõ ra. Mô phỏng trên Matlab cho kết quả tốt khi sử dụng thuật toán L-M để khử tiếng vọng âm học.

1. QUY TRÌNH ĐÁNH GIÁ THÔNG SỐ

Hiện nay, việc khử tiếng vọng trong hệ thống truyền âm đang được ứng dụng nhiều trong điện thoại vô tuyến, điện thoại hình, hội nghị qua điện thoại... đã và đang có nhiều nhà nghiên cứu đưa ra các phương pháp khác nhau để xử lý tiếng vọng trong âm học. Bên cạnh phần tổng quan, trong bài này, tác giả đề xuất một phương pháp đánh giá các thông số của bộ lọc thích nghi dựa vào thuật toán Levenberg-Marquardt để ứng dụng vào việc khử tiếng vọng âm học trong hệ truyền âm stereo.

Một bộ lọc thích nghi có thể tự động thay đổi các thông số để phù hợp với đặc tính của tín hiệu đầu vào. Với khả năng tự điều chỉnh và linh hoạt, bộ lọc thích nghi được thiết kế để sử dụng trong nhiều lĩnh vực khác nhau. Ví dụ như khử tiếng vọng âm thanh, xử lý tín hiệu radar, hệ thống hàng hải, bộ cân bằng của các kênh thông tin, tín hiệu y sinh học ([2–5])

Một bộ lọc thích nghi gồm hai phần: một bộ lọc số với các hệ số thay đổi và một thuật toán thích nghi dùng để điều chỉnh hoặc sửa đổi các hệ số của bộ lọc.

Sử dụng bộ lọc thích trong một hệ thống truyền âm mono được trình bày dưới đây. Xét hệ truyền âm mono, bao gồm các thành phần như sau:

- + Phần phát: tạo tín hiệu $x(n)$ từ nguồn phát.
- + Phần thu: loa, micro, tín hiệu $y(t)$, bộ lọc phần thu với tham số đánh giá $\theta(n)$.
- + Bộ khử nhiễu echo với tham số đánh giá $\hat{\theta}(n)$.
- + $v(t)$ là tín hiệu nhiễu.

Ta có thể viết:

$$y(n) = \theta^T \psi + v(n) \quad (1)$$

Với θ là vectơ thông số:

$$\theta = [\theta_0, \theta_1 \dots \theta_{N-1}]^T$$

Và vectơ quan sát ψ :

$$\psi = [1, x(n), x(n-1) \dots x(n-N+1)]^T$$

Tín hiệu nhiễu $v(n)$ tại phòng thu:

$$v(n) = \sum_{i=N}^{\infty} x(n-i)\theta_i(n) \quad (2)$$

Ứng với hệ thống (1), để tìm các hệ số θ , ta sử dụng mô hình:

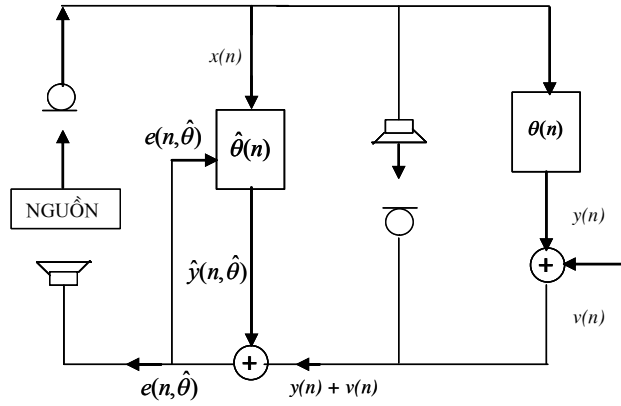
$$\begin{aligned} \hat{Y}(n, \theta) &= \hat{\theta}^T \psi \\ \theta &= [\hat{\theta}_0, \hat{\theta}_1 \dots \hat{\theta}_{N-1}]^T \end{aligned} \quad (3)$$

$N - 1$ là bậc của bộ lọc thích nghi.

Gọi $\varepsilon(n)$ là nhiễu tạo ra giữa các bộ lọc thích nghi, ta có:

$$\varepsilon(n, \hat{\theta}) = x^T(n)[\theta(n) - \hat{\theta}(n)] \quad (4)$$

Hình 1 sau mô tả sơ đồ hệ thống



Hình 1

2. GIẢI BÀI TOÁN

Thuật toán thích nghi được sử dụng để điều chỉnh các hệ số của bộ lọc số nhằm làm giảm tối đa sai số giữa hệ thống thực và mô hình đánh giá.

Có nhiều thuật toán đánh giá được sử dụng, như thuật toán LMS (least mean square), RLS (recursive least squares) và các thuật toán của bộ lọc Kalman. Trong điều kiện tính toán và các thiết bị lưu trữ, thuật toán bình phương tối thiểu LMS đạt hiệu quả nhất, hơn nữa, nó lại mang tính ổn định hơn hai thuật toán trên. Tuy nhiên, thuật toán RLS có tính hội tụ tốt hơn.

Trong bài này, ta sử dụng thuật toán đánh giá bình phương tối thiểu.

Ta xét sai số giữa hệ thống thực và mô hình ở dạng sau:

$$e(n, \hat{\theta}) = y(n) - \hat{y}(n, \hat{\theta}) = y(n) - \hat{\theta}^T \psi(n) \quad (5)$$

Xét tiêu chuẩn bình phương tối thiểu cho sai số, ta có:

$$V(\hat{\theta}) = \frac{1}{2} E\{[e(n, \hat{\theta})]^2\} = \frac{1}{2} E\{[y(n) - \hat{y}(n, \hat{\theta})]^2\} = \frac{1}{2} E\{[y(n) - \hat{\theta}^T \psi(n)]^2\} \quad (6)$$

$E[\cdot]$ là kỳ vọng toán học, $V(\hat{\theta}) \rightarrow \min_{\hat{\theta}}$.

Điều kiện để tối thiểu hóa (6) là gradient theo θ của $V(\theta)$ bằng không.

$$V(\theta) = \frac{1}{2}E\{y^2(n)\} - E\{[y(n)\psi^T(n)]\theta(n)\} + \frac{1}{2}\theta^T(n)E\{\psi(n)\psi^T(n)\}\theta(n) \quad (7)$$

$$\nabla_{\theta}V(\theta) = \frac{\partial V(\theta)}{\partial \theta} = -E\{y(n)\psi(n)\} + E\{\psi(n)\psi^T(n)\}\theta(n) \quad (8)$$

Hoặc ta có thể viết:

$$\hat{\theta} = \operatorname{argmin} \left\{ \frac{1}{N} \sum_{n=1}^N [e(n, \theta)]^2 \right\} \quad (9)$$

Cải biên các kết quả đã có trong tài liệu [4, 5] ta có thể ước lượng tham số tại thời điểm $(n+1)$ dựa trên tham số $\hat{\theta}_n$ tại thời điểm n dựa theo phương pháp Levenberg-Marquardt, ta có thuật toán sau đây:

$$\hat{\theta}_{n+1} = \hat{\theta}_n + \alpha(n) \left[\|\psi(n)\|^2 + \mu(n)I \right]^{-1} e(n, \hat{\theta}_n) \psi(n); \quad (10)$$

I - ma trận đơn vị.

$$\text{Đặt } m(n) = \theta(n) - \hat{\theta}(n). \quad (11)$$

Sau mỗi bước lặp, điều kiện để hệ hội tụ là:

$$E\{\|m(n+1)\|^2\} - E\{\|m(n)\|^2\} < 0 \quad (12)$$

Dựa theo phương pháp Levenberg-Marquardt, ta có thuật toán:

$$m(n+1) = m(n) - \alpha(n) \left[\|\psi(n)\|^2 + \mu(n)I \right]^{-1} e(n, \hat{\theta}_n) \psi(n); \quad (13)$$

Từ (9) và (10), ta có:

$$\begin{aligned} E\{\|m(n+1)\|^2\} - E\{\|m(n)\|^2\} &= \alpha^2(n) E\left\{ e^2(n, \hat{\theta}) \left[\|\psi(n)\|^2 + \mu(n)I \right]^{-1} \right\} \\ &\quad - 2\alpha(n) E\left\{ e(n, \hat{\theta}) \varepsilon \left[\|\psi(n)\|^2 + \mu(n)I \right]^{-1} \right\} < 0 \end{aligned} \quad (14)$$

Như vậy, $\alpha(n)$ cần thỏa điều kiện:

$$0 < \alpha(n) < 2 \frac{E\{e(n, \hat{\theta}) \varepsilon \left[\|\psi(n)\|^2 + \mu(n)I \right]^{-1}\}}{E\{e^2(n, \hat{\theta}) \left[\|\psi(n)\|^2 + \mu(n)I \right]^{-1}\}} \quad (15)$$

$E\{.\}$ là kỳ vọng toán học.

Trong trường hợp $\|\psi(n)\|^2 \approx \text{const}$ và giả sử $v(n)$ và $\varepsilon(n)$ không tương quan nhau, ta có:

$$\alpha(n) = \frac{E\{\varepsilon^2(n)\}}{E\{\varepsilon^2(n)\} + E\{v^2(n)\}} = \frac{E\{\varepsilon^2(n)\}}{E\{e^2(n, \hat{\theta})\}} \quad (16)$$

Từ (7), (8), (10), (12) và (13), ta có :

$$0 < \alpha(n) \leq \min\{1, (12), (13)\} \quad (17)$$

Với cách chọn lựa $\alpha(n)$ trong từng bước lặp, ta có thể viết thuật toán thích nghi để đánh giá cho từng thành phần $\hat{\theta}_i$ của vectơ thông số $\hat{\theta}$ như sau:

$$\hat{\theta}_i(n+) = \hat{\theta}_i(n) + \frac{\alpha(n) e[n, \hat{\theta}_i(n)] \psi_i(n)}{\|\psi_j(n)\|^2} \quad (18)$$

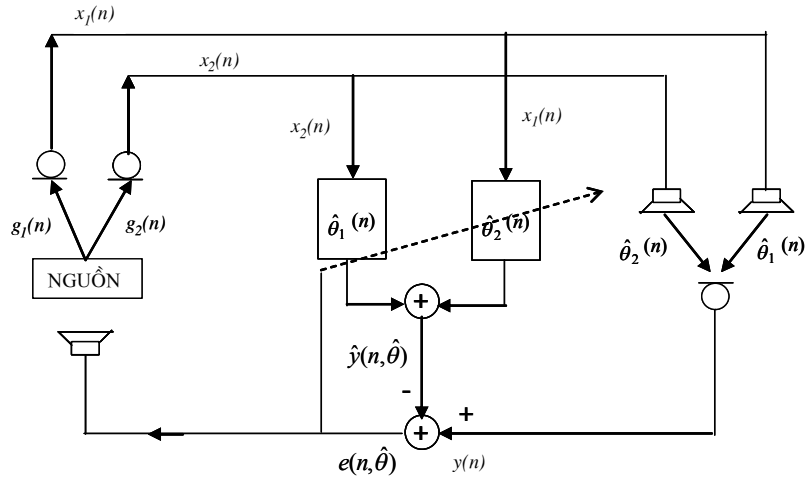
$$\psi_j(n) = \max_{1 \leq i \leq N-1} \{\psi_i(n)\}$$

Hay ở dạng cải biên nhằm tăng độ bền vững của thuật toán:

$$\hat{\theta}_i(n+) = \hat{\theta}_i(n) + \frac{\alpha(n)e[n, \hat{\theta}_i(n)]\psi_i(n)}{\|\psi_j(n)\|^2} \left(1 + \frac{E\{v^2(n)\}}{E\{e^2(n, \hat{\theta}_i(n))\}}\right) \quad (19)$$

$$\psi_j(n) = \max_{1 \leq i \leq N-1} \{\psi_i(n)\}$$

3. XÉT HỆ HAI KÊNH



Hình 2

Hình 2 mô tả mô hình của hệ gồm hai kênh. Bên trái là phòng truyền âm và bên phải là phòng thu âm.

Từ hai tín hiệu đầu vào $x_1(n)$ và $x_2(n)$ của một nguồn chung, ta gọi M là chiều dài của xung đáp ứng $g_i(n)$ trong phòng truyền âm, L là chiều dài của mô hình bộ lọc thích nghi và N là chiều dài của xung đáp ứng trong phòng thu.

Tại tần số f , từ nguồn phát $s(f)$, ta có:

$$\begin{aligned} x_1(f) &= g_1(f)s(f) \\ x_2(f) &= g_2(f)s(f) \end{aligned} \quad (20)$$

Ta giả sử hệ thống là tuyến tính và thời gian không thay đổi, theo [6] ta có quan hệ như sau:

$$x_{1,M}^T(n)g_{2,M} = x_{2,M}^T(n)g_{1,M} \quad (21)$$

Với:

$$x_{i,M}(n) = [x_i(n), x_i(n-1), \dots, x_i(n-M+1)]^T, \quad i = 1, 2 \quad (22)$$

là véc-tơ của tín hiệu lấy mẫu ngõ ra của micro trong phòng truyền âm. Và:

$$g_{i,M} = [g_{i,0}, g_{i,1}, \dots, g_{i,M-1}]^T, \quad i = 1, 2. \quad (23)$$

$$y(n) = \theta_{1,N}^T x_{1,N}(n) + \theta_{2,N}^T x_{2,N}(n) \quad (24)$$

là các vectơ xung đáp ứng giữa nguồn phát và các micro trong phòng truyền âm (tín hiệu ngõ ra của micro trong phòng thu) trong đó:

$$\theta_{i,N} = [\theta_{i,0}, \theta_{i,1}, \dots, \theta_{i,N-1}]^T, \quad i = 1, 2 \quad (25)$$

là vectơ xung đáp ứng giữa loa và các micro trong phòng thu, tín hiệu lấy mẫu trong phòng thu là:

$$x_{i,N}(n) = [x_i(n), x_i(n-1), \dots, x_i(n-N+1)]^T, \quad i = 1, 2 \quad (26)$$

Để tìm các hệ số θ chưa biết, ta sử dụng mô hình:

$$\hat{y}(n, \theta) = \hat{\theta}_{1,L}^T x_{1,L}(n) + \hat{\theta}_{2,L}^T x_{2,L}(n) \quad (27)$$

$$\hat{\theta}_{i,L} = [\hat{\theta}_{i,0}(n), \hat{\theta}_{i,1}(n), \dots, \hat{\theta}_{i,L-1}(n)]^T, \quad i = 1, 2 \quad (28)$$

Sai số giữa tín hiệu ngõ ra phòng thu $y(n)$ và mô hình đánh giá của bộ lọc thích nghi $\hat{\theta}(n, \theta)$ là:

$$e(n, \hat{\theta}) = y(n) - \hat{y}(n, \hat{\theta}) = y(n) - [\hat{\theta}_{1,L}^T x_{1,L}(n) + \hat{\theta}_{2,L}^T x_{2,L}(n)] \quad (29)$$

Từ (7), ta có:

$$\hat{\theta}_{(i,n+1)} = \hat{\theta}_{(i,n)} + \frac{\alpha(n)e(n)x_i(n)}{1 + \alpha(n) \sum_{\rho=1}^2 \|x_{\rho j(n)}\|^2} \quad (30)$$

$$\|x_{\rho j(n)}\| = \max_{1 \leq i \leq N-1} \|x_i(n)\|$$

với $\alpha(n)$ được chọn theo (17).

Ta có thể sử dụng giá trị tại thời điểm $(n-k)$ để xác định $\hat{\theta}_{i,n+1}$ như sau:

$$\hat{\theta}_{i,n+1} = \hat{\theta}_{i,n} + \frac{\alpha(n)e(n-k)x_i(n-k)}{1 + \alpha(n) \sum_{j=1}^2 \|x_j(n-k)\|^2} \quad (31)$$

$$\|x_j(n-k)\| = \max_{1 \leq i \leq N-1} \|x_i(n-k)\|$$

Trong trường hợp phức tạp hơn, ta phải xét đến nhiều quá khứ. Khi đó, ta có:

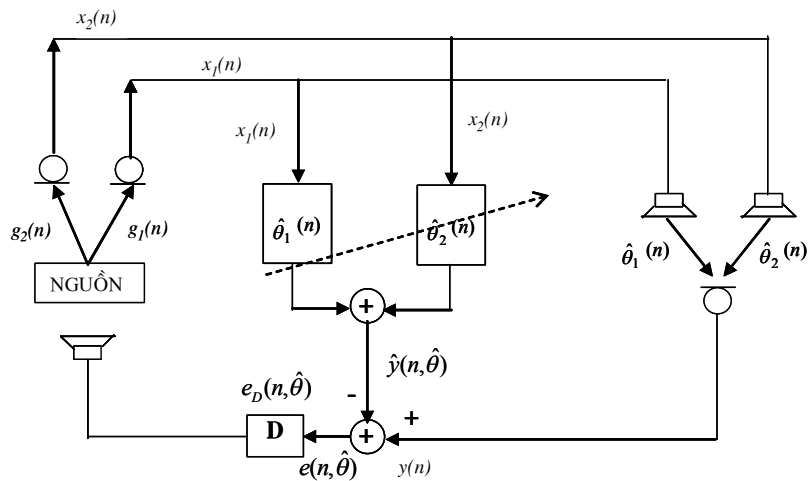
$$e_D(n, \theta) = (1 + d_1 z^{-1} + \dots + d_{n-1} z^{-(n-1)}) e(n, \theta) = D \cdot e(n, \theta) \quad (32)$$

$$D = 1 + d_1 z^{-1} + \dots + d_{n-1} z^{-(n-1)} \quad (33)$$

D - bộ lọc ổn định.

Hình 2 lúc này trở thành hình 3 sau:

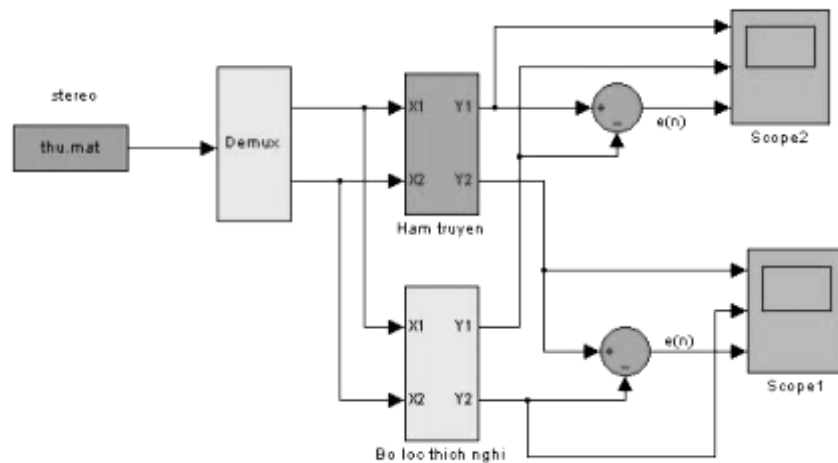
Với thuật toán thích nghi nêu trên, cùng những ước lượng giới hạn các tham số thay đổi, ta có thể tính toán các hệ số thay đổi của bộ lọc số. Việc thay đổi các hệ số của bộ lọc số nhằm đạt đến mục đích làm cho sai số của hệ thống tiến tới một giá trị min có thể chấp nhận được. Điều này cũng có nghĩa là khử bớt tiếng vọng của hệ âm thanh stereo.



Hình 3

4. KẾT QUẢ MÔ PHỎNG

4.1. Khảo sát sơ đồ truyền âm stereo gồm 1 kênh



Hình 4

Tín hiệu âm thanh stereo qua khối Demux sẽ tách thành hai kênh left và right. Tín hiệu hai kênh cùng được đưa vào hai khối:

- + Khối hàm truyền: với hàm truyền cho trước, tín hiệu qua khối hàm truyền chính là tiếng vọng echo $y(n)$.
- + Khối bộ lọc thích nghi, sử dụng thuật toán thích nghi và bộ lọc FIR để tạo ra tín hiệu đánh giá $\hat{y}(n)$.
- + Sai số $e(n) = y(n) - \hat{y}(n)$ được hiệu chỉnh liên tục sao cho $e(n) \Rightarrow 0$.
- + Tín hiệu echo $y(n)$, tín hiệu nhận dạng $\hat{y}(n)$ và sai số đánh giá $e(n)$ được xem trên hai scop1 và scop2.

Kết quả đánh giá trực quan có thể cho ta biết được sai số tối ưu.

4.2. Chạy mô phỏng với file âm thanh

- Xét hệ 1 kênh, chạy mô phỏng với số bậc bộ lọc là $N = 5; 10; 15; 20$.
- Tìm các giá trị sau:
 - + Sai số giữa tín hiệu và tín hiệu nhận dạng
 - + Sự không đồng bộ giữa đáp ứng xung phòng thu với đáp ứng xung của bộ lọc MIS-ALIGNMENT
 - + Độ tăng của sự suy hao tiếng vọng: ERLE
 - + Đồ thị so sánh giữa đáp ứng hàm truyền và bộ lọc
 - + So sánh các giá trị đã có theo bậc của bộ lọc

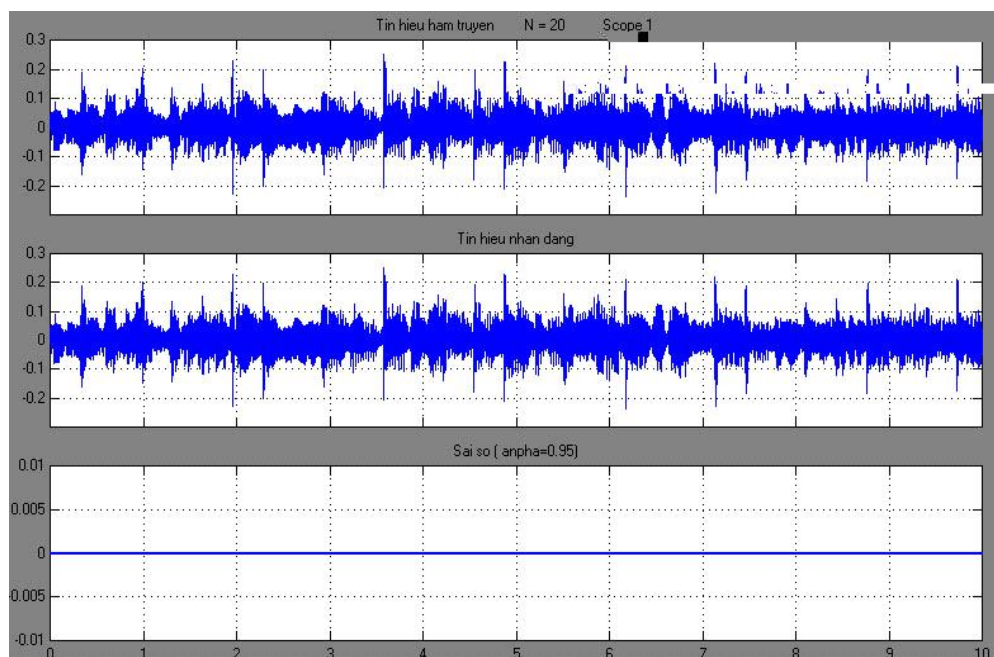
Mô hình 1 kênh

Xét hệ gồm 1 kênh stereo như hình 4. Ta lần lượt khảo sát các tình huống xảy ra khi bậc của bộ lọc thay đổi $N = 5; 10; 15; 20$. Các thông số khảo sát đã được nêu trên.

Trường hợp $N = 20$

a) *Xét sai số:* Tín hiệu quan sát trên scope1: tín hiệu kênh trái

- + Kênh 1: tín hiệu hàm truyền $y(n)$.
- + Kênh 2: tín hiệu nhận dạng $\hat{y}(n)$ của bộ lọc thích nghi.
- + Kênh 3: sai số $e(n)$ giữa tín hiệu hàm truyền $y(n)$ và tín hiệu nhận dạng $\hat{y}(n)$.



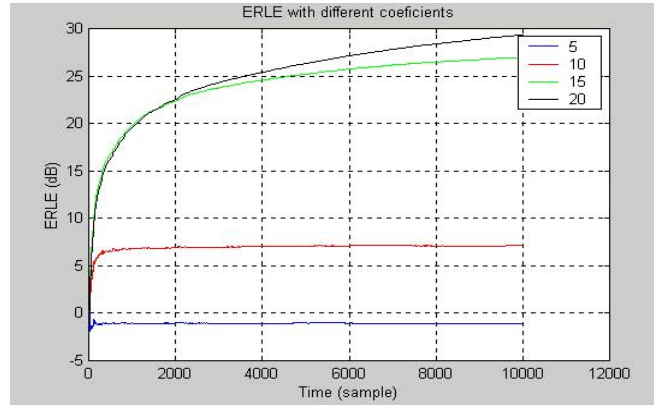
Hình 5

Xem hình 5, ta thấy tín hiệu nhận dạng có biên độ gần bằng tín hiệu hàm truyền. Sai số ngõ ra có biên độ bằng zero. Với kết quả này, bậc của bộ lọc thích nghi $N = 20$ chấp nhận được. Tuy nhiên, ta cần tham khảo thêm các giá trị khác phía sau:

b) *Xét giá trị ERLE (echo return loss enhancement)*

$$ERLE = 10 \log_{10} \left[\frac{\sum_{i=1}^k y^2(i)}{\sum_{i=1}^k e^2(i)} \right]$$

So sánh giá trị ERLE của bốn trường hợp khảo sát, ta có các đồ thị trong hình vẽ 6. Theo hình vẽ, ta thấy với bậc bộ lọc $N = 20$, ta có giá trị ERLE đạt giá trị lớn nhất.



Hình 6

5. KẾT LUẬN

Trong bài này, tác giả đã dùng thuật toán đánh giá bền vững Levenberg-Marquardt ứng dụng cho bộ lọc thích nghi để khử tiếng vọng âm học. Việc đánh giá lại các tham số trong từng bước lặp sẽ giúp hệ thống ổn định hơn. Thuật toán này không chỉ ứng dụng trong hệ 2 kênh truyền âm, mà còn có thể dùng trong hệ thống thông tin di động.

Lời cảm ơn. Tác giả cảm ơn TS Dương Hoài Nghĩa đã hướng dẫn tận tình và có nhiều gợi ý khoa học quý giá, cảm ơn GS.TSKH Nguyễn Thúc Loan đã cung cấp nhiều tài liệu và thông tin nghiên cứu, cảm ơn Ban biên tập đã giúp đỡ trong những bước đầu nghiên cứu.

TÀI LIỆU THAM KHẢO

- [1] Christina Breining, Thomas Shertler, Delay-free low-cost step-gain estimation for adaptive filters in acoustic echo cancellation, *Signal processing* (80) (2000) 1721-1731
- [2] Donald L. Duttweiler, Avoiding Slow band-edge convergence in subband echo cancelers, *IEEE transactions on signal processing* 49 (3) (2001).
- [3] Emmanuel C. Ifeachor - Barrie W. Jervis, *Digital signal processing*, Addison-wesley, 1993.
- [4] Fredric M. Ham, Ph. D and Irica Kostanic, *Principles of Neurocomputing for science and engineering*, Mc. Graw Hill.
- [5] Jianfeng Liu, A novel adaptation scheme in the NLMS algorithm for echo cancellation, *IEEE signal processing letters* vol. S. (1) (2001).
- [6] Tomas Gansler & Jacob Benesty, Stereophonic acoustic echo cancellation and two-channel adaptive filtering: an overview, *Control signal process*, 2000, 14:565-585.

Nhận bài ngày 29 - 11 - 2002