

XÁC ĐỊNH ĐIỀU KIỆN ĐỂ BẢO TOÀN THÔNG TIN TRONG LÀM TRƠN PHỔ TÍN HIỆU ĐIỆN TIM NHIỄM NHIỀU

PHẠM TRẦN NHƯ¹, NGUYỄN THỊ QUỲNH LAN², HOÀNG MẠNH HÀ³

¹*Viện Công nghệ thông tin, Viện Khoa học và Công nghệ Việt Nam*

¹*Khoa Công nghệ thông tin, Đại học Thành Đô*

²*Khoa Toán Kinh Tế, Đại học Kinh tế Quốc dân*

³*Đại học Thủ dầu 1*

Tóm tắt. Các thông tin hữu ích của tín hiệu y sinh thường bị biến đổi sai lệch do ảnh hưởng từ nhiễu cộng tính, đơn tần số và thay đổi ngẫu nhiên. Phương pháp khử nhiễu thông dụng là sử dụng bộ lọc triệt tần thích nghi (ANF). Lần đầu tiên được B. Widrow đề xuất năm 1975, ANF có khả năng tự điều chỉnh các tham số lọc để thích ứng với sự thay đổi ngẫu nhiên của nhiễu. ANF có ưu điểm về độ chính xác nhưng tốc độ điều chỉnh chậm. Đề xuất của chúng tôi trong [6] cho phép rút ngắn thời gian thích ứng với sự thay đổi của nhiễu và giảm thiểu gây méo các thông tin hữu ích. Các kết quả đó dựa trên tối thiểu hoá sai số tại bước làm trơn phổ của tín hiệu nhiễm nhiễu. Bài báo này trình bày kết quả đạt được trong việc xác định các yếu tố ảnh hưởng đến chặn trên của xác suất xảy ra sự kiện sai số vượt quá ngưỡng cho phép. Trong các yếu tố đó, độ dài cắt mẫu là tham số duy nhất mà người sử dụng có thể điều chỉnh. Ảnh hưởng của độ dài cắt mẫu được chú ý đặc biệt và được kiểm chứng qua các thực nghiệm.

Abstract. Information of Bio-signals is often corrupted by additive noise. Mono-frequency characterization of such noisy with random translation is recognized as major obstacle for denoising. The usual method of denoising is to pass the composite noisy signal through an adaptive notch filter (ANF), which originated with the pioneering work of B. Widrow. An ANF has only the ability to adjust their parameters automatically with slow adaptation. This was a major motivation for enhancing the convergence rate of processing with the accuracy of the denoising. Our proposed method should have an optimum speed of adaptation and allow the minimization of loss of information of the signal of interest [6]. The results presented in [6] suggest the need to minimize a error of smoothing of a spectrum of the noisy signal and to solve problems of determining the affect on upper bound of probability of event error threshold exceeded. This paper deals with new method for such problem and its basis is stated into a lemma. In addition, extensive simulations are conducted to corroborate the efficiency of the proposed estimator with the theoretical analysis.

1. MỞ ĐẦU

Bài toán khử nhiễu cho tín hiệu y sinh, với nhiễu cộng tính, đơn tần số và thay đổi ngẫu nhiên

Gọi $NoisyECG(n)$, với $n \in Z$, là tín hiệu y sinh bị lẫn nhiễu mà ta thu nhận từ người bệnh. Theo [1], ảnh hưởng của nhiễu lên tín hiệu được mô tả qua công thức sau

$$NoisyECG(n) = S(n) + N(n),$$

trong đó, $S(n)$: Tín hiệu y sinh sạch cần thu nhận; $N(n)$: Nhiễu cộng ảnh hưởng lên tín hiệu y sinh.

Mô hình của nhiễu đã được xác định trong [1,2] như sau:

$$N(n) = A \cos(\omega_0 n + \phi)$$

trong đó, A, ω_0, ϕ lần lượt là biên độ, tần số và góc pha của nhiễu và đều có thể thay đổi một cách ngẫu nhiên.

Định nghĩa 1. Cho $x(n)$ là tín hiệu rời rạc, $X(z)$ được gọi là biến đổi z của $x(n)$ nếu

$$X(z) = \sum_{n=-\infty}^{\infty} x(n)z^{-n},$$

trong đó z là biến phức.

Định nghĩa 2. Gọi $x(n)$ là dãy tín hiệu đầu vào của bộ lọc, $X(z)$ là biến đổi z của $x(n)$ theo Định nghĩa 1. Gọi $y(n)$ là dãy tín hiệu đầu ra của bộ lọc, $Y(z)$ là biến đổi z của $y(n)$ theo Định nghĩa 1. Hàm truyền $H(z)$ của bộ lọc được xác định như sau

$$H(z) = \frac{Y(z)}{X(z)}.$$

Theo [1] và [2], $H_1(z)$ được gọi là hàm truyền của bộ lọc triệt tín hiệu $Noise(n)$, $Noise(n) = A \cos(\omega_1 n + \phi)$ nếu $H_1(z)$ có dạng sau

$$H_1(z) = \frac{z^{-2} - (2 \cos \omega_1)z^{-1} + 1}{\zeta^2 z^{-2} - (2\zeta \cos \omega_1)z^{-1} + 1} \quad (1)$$

trong đó

ω_1 : Tần số triệt của bộ lọc triệt tần. Bộ lọc có hàm truyền là công thức (1) sẽ chặn bất kỳ tín hiệu có tần số ω_1 . Nếu ta đặt $\omega_1 = \omega_0$, trong đó ω_0 là tần số của nhiễu $N(n)$, thì bộ lọc có hàm truyền cho bởi (1) sẽ lọc được nhiễu $N(n)$.

ζ : Tham số quyết định sự suy giảm tới các tín hiệu hữu ích có tần số lân cận tần số trung tâm ω_1 .

Tuy nhiên, như trên đã mô tả, do tần số ω_0 của nhiễu có đặc tính thay đổi ngẫu nhiên, khó xác định nên bộ lọc có hàm truyền cho bởi công thức (1) chưa lọc được hoàn toàn loại nhiễu này.

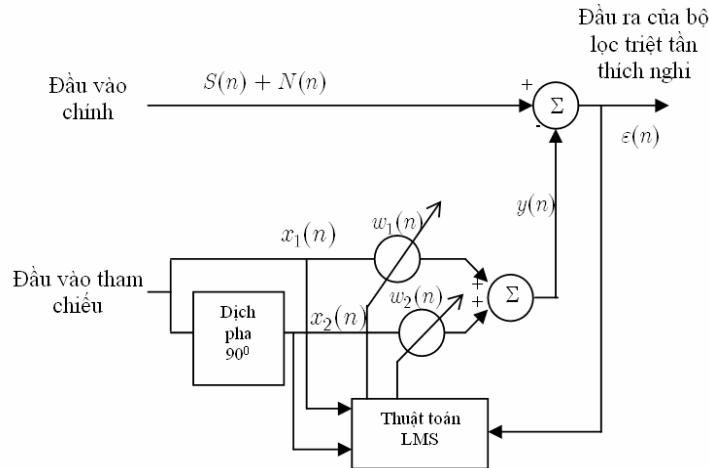
Để giải quyết các vấn đề lọc nhiễu nêu trên, B. Widrow đề xuất mô hình bộ lọc triệt tần thích nghi (ANF) trong Hình 1, trong đó, $NoisyECG(n) = S(n) + N(n)$ là tín hiệu y sinh lẫn nhiễu, được thu nhận từ đầu vào chính.

$x_1(n)$: Tín hiệu nhiễu tham chiếu, với

$$x_1(n) = C \cos(\omega_0 n + \phi). \quad (2)$$

$x_2(n)$: Là $x_1(n)$ được làm trễ pha 90° ,

$$x_2(n) = C \sin(\omega_0 n + \phi). \quad (3)$$



Hình 1. Mô hình bộ lọc triệt tần thích nghi

Mô hình sử dụng thuật toán LMS để điều chỉnh hai trọng số như sau

$$w_1(n + 1) = w_1(n) + 2\mu\epsilon(n)x_1(n), \tag{4}$$

$$w_2(n + 1) = w_2(n) + 2\mu\epsilon(n)x_2(n), \tag{5}$$

trong đó

- $w_1(n + 1)$: Trọng số w_1 tại bước thứ $(n + 1)$;
- $w_2(n + 1)$: Trọng số w_2 tại bước thứ $(n + 1)$;
- μ : Tham số kích thước bước;
- $\epsilon(n)$: Tín hiệu ra tại bước thứ n .

Nếu tham số μ được chọn rất nhỏ, khi thuật toán LMS sẽ hội tụ, mô hình lọc nhiễu triệt tần thích nghi có hàm truyền như sau

$$\hat{H}(z) = \frac{-z^{-2} - (2 \cos w_0)z^{-1} + 1}{(1 - 2\mu C^2)z^{-2} - \{2(1 - \mu C^2) \cos w_0\}z^{-1} + 1}, \tag{6}$$

trong đó

- w_0 : Tần số của nhiễu;
- C : Biên độ của tín hiệu nhiễu tham chiếu.

Có nghĩa là mỗi khi tần số của nhiễu thay đổi, thuật toán LMS điều chỉnh dần các trọng số w_1 và w_2 cho đến khi hàm truyền hội tụ đến công thức (6), khi đó ANF có hàm truyền $\hat{H}(z) \approx H_1(z)$, với tần số triệt tại $w_1 = w_0$ do vậy có khả năng triệt nhiễu tần số w_0 .

Với tham số μ rất nhỏ, thuật toán LMS sẽ hội tụ rất chậm khi được sử dụng trong mô hình lọc nhiễu triệt tần thích nghi [3,4,5]. Do vậy, yêu cầu mới đặt ra là rút ngắn thời gian hội tụ của thuật toán LMS, trong khi vẫn bảo toàn độ chính xác của mô hình. Đã có một số đề xuất giải quyết vấn đề này như sử dụng μ biến đổi như trong [4], sử dụng cấu trúc động (xem [5]). Với cùng mục đích như vậy nhưng với tiếp cận khác, chúng tôi đề xuất mô hình lọc nhiễu và giải thuật dò tìm tần số w_0 , các kết quả đã được công bố tại [6]. Trong đó sử dụng các kỹ thuật của biến đổi Fourier, biến đổi sóng nhỏ và kỹ thuật làm tròn. Độ chính xác

của biến đổi Fourier và biến đổi sóng nhỏ đã được đánh giá trong [1,2,7,8], tuy nhiên các kết quả nghiên cứu độ sai lệch sau khi làm trơn phổ tín hiệu vẫn chưa được công bố. Hơn thế nữa, sai số này có ảnh hưởng trực tiếp đến độ chính xác của phép lọc nhiễu. Do vậy chúng tôi đã tìm cách giải quyết vấn đề này với tiếp cận xác suất, bước đầu đã xác định được điều kiện để sai số nằm trong giới hạn cho phép. Bài báo này nhằm mục đích giới thiệu kết quả đầu tiên trong việc đảm bảo độ chính xác cho bước làm trơn phổ Biên độ - Tần số của tín hiệu y sinh nhiễu nhiễu.

2. CHẶN TRÊN CỦA XÁC SUẤT XẢY RA SỰ KIỆN SAI SỐ VƯỢT GIỚI HẠN TRONG LÀM TRƠN PHỔ CỦA TÍN HIỆU ĐIỆN TIM NHIỄM NHIỄU

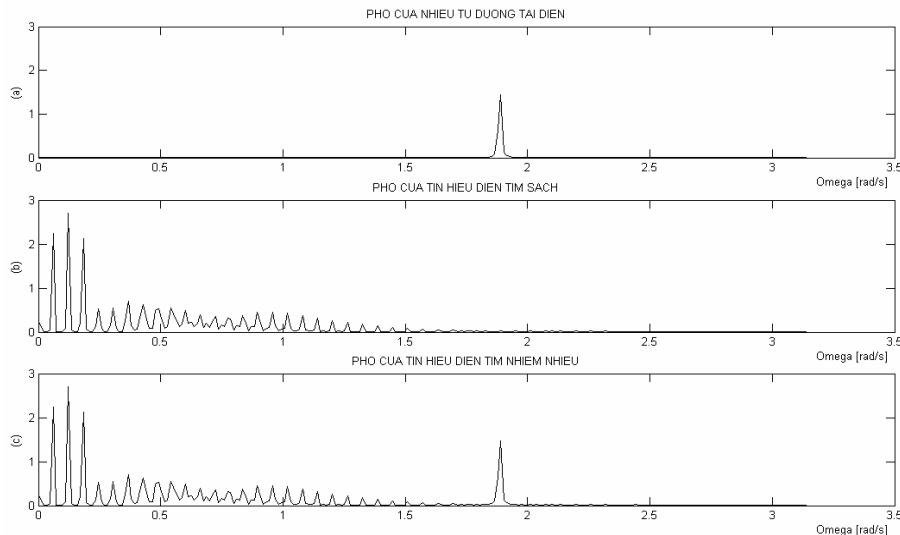
2.1. Phép làm trơn trong bài toán lọc nhiễu

Định nghĩa 3. Nếu $x(n)$ là 1 dãy số thực thoả mãn điều kiện $\sum_{n=-\infty}^{\infty} |x(n)| < \infty$ thì $\mathbf{X}(w) = \sum_{n=-\infty}^{\infty} x(n)e^{-jwn}$ được gọi phép biến đổi Fourier đối với $x(n)$, trong đó, $\mathbf{X}(w)$: Kết quả phép biến đổi Fourier đối với $x(n)$; i : Số ảo, với $i^2 = -1$.

Định nghĩa 4. Nếu $x(n)$ là tín hiệu rời rạc hoặc là đáp ứng xung của bộ lọc số, và $\mathbf{X}(w)$ là kết quả của phép biến đổi Fourier đối với $x(n)$ thì $|\mathbf{X}(w)|$ được gọi là phổ biên độ - tần số của $x(n)$.

Theo [2], $|\mathbf{X}(w)| = |\mathbf{X}(w + 2k\pi)|$ với $k \in Z$. Trong đoạn $[0, 2\pi]$, $|\mathbf{X}(w)|$ có tính chất đối xứng qua π , do vậy phổ của $x(n)$ thường được biểu diễn trong quãng $[0, \pi]$.

Trong Hình 2(a) là $\mathbf{N}(w)$, phổ của riêng nhiễu động, cộng tính và đơn tần số. Hình 2(b), $|\mathbf{S}(w)|$ phổ của tín hiệu điện tim sạch. Hình 2(c), $|\mathbf{NoisyECG}(w)|$ là phổ của tín hiệu điện tim bị nhiễu nhiễu động cộng tính và đơn tần số nêu trên, trong đó phổ của nhiễu được phản ánh qua điểm đột biến tại w_0 .



Hình 2. So sánh phổ của một tín hiệu điện tim nhiễu nhiễu ($|\mathbf{NoisyECG}(w)|$)

Việc phát hiện sự tương đương giữa bài toán tìm vector trọng số cho ANF với bài toán Tìm tọa độ điểm đột biến trên phổ của tín hiệu điện tim nhiễu nhiễu, đã cho phép chúng tôi

sử dụng các kỹ thuật mới để giải quyết bài toán khử nhiễu đơn tần và biến đổi ngẫu nhiên (xem [6]). Vì sau khi làm trơn, $|\mathbf{NoisyECG}(w)|$ chỉ còn lại duy nhất điểm đột biến có tọa độ tại w_0 cần tìm. Do vậy, với mục đích giảm số phép tính và tăng độ chính xác, một cách trực giác, chúng tôi đề xuất làm trơn $|\mathbf{NoisyECG}(w)|$.

Phép làm trơn như sau.

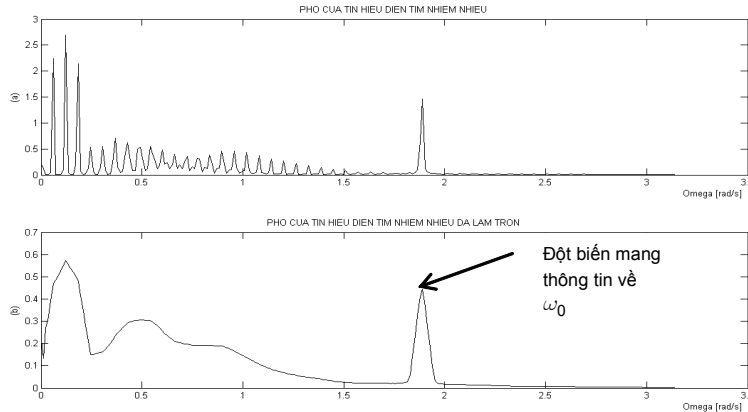
Trong xử lý tín hiệu, hàm cửa sổ chữ nhật đối với biến w được định nghĩa như sau

$$Rect_L(w) = \begin{cases} 1 & 0 \leq w \leq L \\ 0 & w \notin [0, L] \end{cases}.$$

Định nghĩa 5. Cho tín hiệu $y(w)$ xác định trong khoảng $[a, b]$, với $L, \tau \in R^1$, tín hiệu $smoothed_y(w)$ được gọi là kết quả làm trơn đối với $y(w)$ nếu

$$smoothed_y(w) = \frac{1}{L} \int_{-\infty}^{+\infty} y(\tau) Rect_L(w - \tau) d\tau.$$

Kết quả phép làm trơn đối với tín hiệu điện tim nhiễu nhiều được mô tả trong Hình 3.



Hình 3. Phổ của một tín hiệu điện tim nhiễu nhiều (a) và kết quả làm trơn (b)

Để lấy thông tin về phổ của tín hiệu, chỉ cần biểu diễn $|\mathbf{NoisyECG}(w)|$ trong quãng $w \in 0 \div \pi$. Khi đó, có thể coi $|\mathbf{NoisyECG}(w)|$ là hàm số với biến w , và có đột biến tại w_0 . Với mục đích chuẩn bị cho việc sử dụng các kỹ thuật tính toán với biến rời rạc để xác định tọa độ điểm đột biến ở các phần trình bày sau, chúng tôi đã rời rạc hoá đoạn $0 \div \pi$ của ω . Khi đó, nếu khoảng cách giữa các điểm rời rạc là ω_s , phổ của tín hiệu điện tim nhiễu nhiều là hàm rời rạc $|\mathbf{NoisyECG}(n\omega_s)|$ với $n \in Z^+$, xác định tại những điểm là bội của ω_s . Tương tự như cách biểu diễn tín hiệu trong miền thời gian rời rạc, phổ của tín hiệu điện tim nhiễu nhiều được biểu diễn thành $|\mathbf{NoisyECG}(n)|$, do đó thủ tục tìm ω_0 được chuyển thành tìm n_0 với $\omega_0 = \omega_s n_0$. Hình 4 mô tả $|\mathbf{NoisyECG}(n)|$ trong đoạn $0 \div \pi$ được chia làm 256 điểm, với khoảng cách giữa 2 điểm liên tiếp $\omega_s = 0,0123rad/s$.

Gọi $\mathbf{y}(n) = |\mathbf{NoisyECG}(n)|$ là kết quả rời rạc hoá phổ của tín hiệu điện tim nhiễu nhiều. Theo [2], làm trơn $y(n)$ được định nghĩa như sau

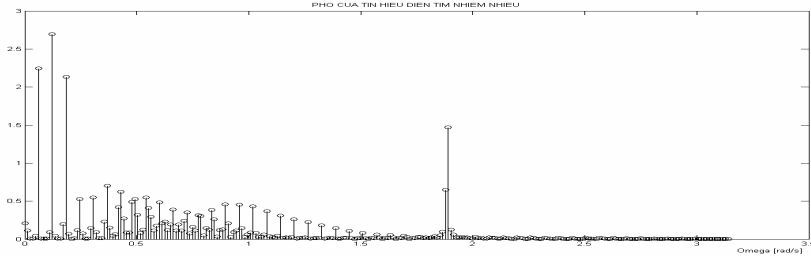
$$\mathbf{f}(n) = \frac{1}{l} \sum_{k=0}^{l-1} y(n+k), \tag{7}$$

trong đó

$\mathbf{y}(n)$: Kết quả rời rạc hoá phổ của tín hiệu điện tim nhiễu nhiều;

$\mathbf{f}(n)$: Kết quả làm trơn trên phổ rời rạc của tín hiệu;

$l \in \mathbb{Z}^+$: Là độ dài cắt mẫu của phép làm trơn.



Hình 4. Phổ rời rạc của một tín hiệu điện tim nhiễu nhiều

2.2. Các yếu tố ảnh hưởng tới xác suất sự kiện sai số nằm trong giới hạn

Khi làm trơn phổ của tín hiệu điện tim nhiễu nhiều theo công thức (7), ta cần tìm các yếu tố ảnh hưởng đến xác suất của sự kiện

$$|\mathbf{y}(n) - \mathbf{f}(n)| < \rho \text{ với } n = 1, 2, \dots, 256,$$

trong đó ρ là giới hạn độ sai lệch cho phép, hoặc có thể viết như sau

$$\text{Prob}[|\mathbf{y}(n) - \mathbf{f}(n)| < \rho] \geq 1 - \nu, \quad (8)$$

trong đó ν là độ không tin cậy.

Xác định các yếu tố ảnh hưởng đến $\text{Prob}[|y(n) - f(n)| < \rho]$ được đề xuất như sau. Coi độ lớn của $\mathbf{y}(n)$ tại $\mathbf{y}(1), \mathbf{y}(2), \dots$ là 1 dãy các sự kiện ngẫu nhiên, biến ngẫu nhiên Y_1, Y_2, \dots tương ứng với sự kiện độ lớn của $y(1), y(2), \dots$

Giả sử Y_1, Y_2, \dots là dãy các biến ngẫu nhiên độc lập và cùng phân phối. Khi đó công thức (8) có thể có thể viết lại như sau :

$$\text{Prob}[|Y_n - M_{n,l}| < \rho] \geq 1 - \nu, \quad (9)$$

trong đó, $M_{n,l} = \frac{1}{l} \sum_{m=1}^l Y_m$ là trung bình số học của l biến ngẫu nhiên, tính từ n của dãy các biến ngẫu nhiên $\{Y_n\}$.

Theo [9], trong thực nghiệm một sự kiện được coi là chắc chắn xảy ra nếu có xác suất xấp xỉ 0,9975. Do vậy ta có định nghĩa sau.

Định nghĩa 6. Sự kiện $|Y_n - M_{n,l}| < \rho$ được coi là chắc chắn xảy ra trong thực nghiệm nếu $\nu \leq 0,0025$.

Các yếu tố ảnh hưởng đến xác suất sự kiện sai số vượt ρ trong bổ đề sau.

Bổ đề 1. Cho $\mathbf{y}(n)$ với $n = 1, 2, \dots, 256$ là các giá trị rời rạc của phổ tín hiệu điện tim nhiễu nhiều. Gọi Y_n với $n = 1, 2, \dots, 256$ là các biến ngẫu nhiên tương ứng với các giá trị của $\mathbf{y}(n)$, $\mathbf{f}(n)$ là kết quả phép làm trơn $\mathbf{y}(n)$ theo công thức (7), l là độ dài cắt mẫu của phép làm trơn, ρ là giới hạn độ sai lệch cho phép. Giả sử Y_n là các biến ngẫu nhiên độc lập và có cùng phân phối chuẩn và có cùng phương sai là σ_Y^2 , nếu

$$\frac{\sigma_Y^2}{\rho^2} \left(1 + \frac{1}{l}\right) \leq 0,0025$$

thì chắc chắn sự kiện $|Y_n - M_{n,l}| < \rho$ sẽ xảy ra trong thực nghiệm.

Chứng minh

Từ bất đẳng thức Chebyshev ta có

$$Prob[|S_{n,l} - ES_{n,l}| \geq k\sigma_S] \geq \frac{1}{k^2},$$

trong đó

$S_{n,l} = \sum_{m=1}^l Y_m$, ta thấy rằng $S_{n,l}$ cũng là biến ngẫu nhiên;

σ_S^2 : Phương sai của biến ngẫu nhiên $S_{n,l}$. Nếu gọi σ_Y^2 là phương sai của biến ngẫu nhiên Y_n thì ta có $\sigma_S = \sqrt{l}\sigma_Y$.

Do vậy ta có

$$Prob[|S_{n,l} - l\mu_n| \geq k\sqrt{l}\sigma] \geq \frac{1}{k^2},$$

trong đó, $\mu_{n,l} = EY_{n,l}$: Kỳ vọng của l biến ngẫu nhiên Y từ vị trí n nên

$$Prob[|M_{n,l} - \mu_n| \geq \frac{k\sigma_Y}{\sqrt{l}}] \leq \frac{1}{k^2}.$$

Đặt $k = \frac{\rho\sqrt{l}}{\sigma_Y}$,

$$Prob[|M_{n,l} - \mu_n| \geq \rho] \leq \frac{\sigma_Y^2}{l\rho^2}. \quad (10)$$

Cũng từ bất đẳng thức Chebyshev,

$$Prob[|Y_n - \mu_n| \geq \rho] \leq \frac{\sigma_Y^2}{\rho^2}. \quad (11)$$

Từ (10) và (11) ta có

$$Prob[|Y_n - M_{n,l}| \geq \rho] \leq \frac{\sigma_Y^2}{\rho^2} \left(1 + \frac{1}{l}\right),$$

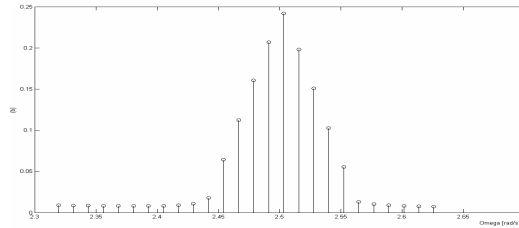
$$Prob[|Y_n - M_{n,l}| < \rho] \geq 1 - \frac{\sigma_Y^2}{\rho^2} \left(1 + \frac{1}{l}\right). \quad (12)$$

So sánh công thức (9) và (12) thấy rằng $\nu = \frac{\sigma_Y^2}{\rho^2} \left(1 + \frac{1}{l}\right)$. Do vậy, nếu $\frac{\sigma_Y^2}{\rho^2} \left(1 + \frac{1}{l}\right) \leq 0,0025$ thì $Prob[|Y_n - M_{n,l}| < \rho] \geq 0,9975$, từ Định nghĩa 6 suy ra sự kiện $|Y_n - M_{n,l}| < \rho$ chắc chắn xảy ra. ■

2.3. Thực nghiệm

2.3.1. Xác định tham số ρ

Theo [7] và [8], tọa độ đỉnh của đột biến nếu đột biến của $Y(n)$ tương ứng với điểm qua không giữa cặp cực đại-cực tiểu cục bộ của phép tính xấp xỉ đạo hàm bậc 1 đối với $Y(n)$. Hình 5 sau mô tả đột biến cần xác định tọa độ trên $\mathbf{y}(n)$.



Hình 5. Độ rộng của đột biến trên $\mathbf{y}(n)$

Qua Hình 5 ta thấy rằng đột biến nhọn trong hình có độ rộng từ $\omega = 2,45\text{rad/s}$ đến $\omega = 2,55\text{rad/s}$, là loại đột biến có độ rộng mức trung bình ($0,1\text{rad/s}$), đột biến có độ rộng lớn nhất là $0,5\text{rad/s}$. Do vậy để xác định tọa độ đỉnh đột biến qua phép tính xấp xỉ đạo hàm bậc 1 đối với $\mathbf{y}(n)$ (xem [7, 8]), độ cao của đỉnh đột biến chỉ cần lớn hơn $0,5$. Trong khi đột biến cần tìm trong phổ của tín hiệu $|\mathbf{N}(n)|$ được xác định là $2 < |\mathbf{N}(n)| < 3$ (xem [1]), do vậy độ sai lệch cho phép sau khi làm tròn $0,5 < \rho < 2,5$.

2.3.2. Xác định σ_Y^2

Việc xác định $\hat{\sigma}_{Y_1}^2, \hat{\sigma}_{Y_2}^2, \dots$ tương ứng với các biến ngẫu nhiên Y_1, Y_2, \dots được mô tả như sau. Để lấy số liệu tính $\hat{\sigma}_{Y_1}^2, \hat{\sigma}_{Y_2}^2, \dots$ ta sử dụng 100 dữ liệu sóng tín hiệu điện tim từ 100 người khác nhau từ cơ sở dữ liệu MIT-BIH, có tại website: <http://ecg.mit.edu/>. Sau đó, từng dữ liệu được biến đổi Fourier theo Định nghĩa 3. Tại bước tiếp theo, phổ của từng dữ liệu sóng điện tim được rời rạc hoá với khoảng cách giữa 2 giá trị rời rạc là $\omega_S = 0,0123\text{rad/s}$, tương ứng với việc chia quãng Nyquist $0 \div \pi$ thành 256 đoạn. Ta ký hiệu $Y_1(n)$ là kết quả rời rạc hoá phổ của dữ liệu tín hiệu điện tim thứ nhất. Tương tự, $Y_{100}(n)$ là kết quả rời rạc hoá phổ của tín hiệu điện tim thứ 100.

Nhận xét: Trên cơ sở dữ liệu từ 100 người khác nhau, nếu với cùng n , các giá trị từ $Y_1(n)$ đến $Y_{100}(n)$ tập trung quanh 1 giá trị nhất định. Điều này gợi ý chúng tôi giả thiết các biến ngẫu nhiên Y_1, Y_2, \dots có cùng phân phối Gaussian. Theo [9], trong trường hợp này công thức tính phương sai cho biến ngẫu nhiên Y_1 bằng thực nghiệm như sau

$$\hat{\sigma}_{Y_1}^2 = \frac{1}{100} \sum_{k=1}^{100} (Y_k(1) - \bar{Y}_1)^2,$$

trong đó

$\hat{\sigma}_{Y_1}^2$: Phương sai thực nghiệm của biến ngẫu nhiên Y_1 ;

$Y_k(1)$ là độ lớn của biến ngẫu nhiên Y_1 của bệnh nhân thứ k ;

\bar{Y}_1 là trung bình số học của các độ lớn của biến Y_1 với $\bar{Y}_1 = \frac{1}{100} \sum_{k=1}^{100} Y_k(1)$.

Số liệu để tính $\hat{\sigma}_{Y_1}^2$ được mô tả qua Bảng 1.

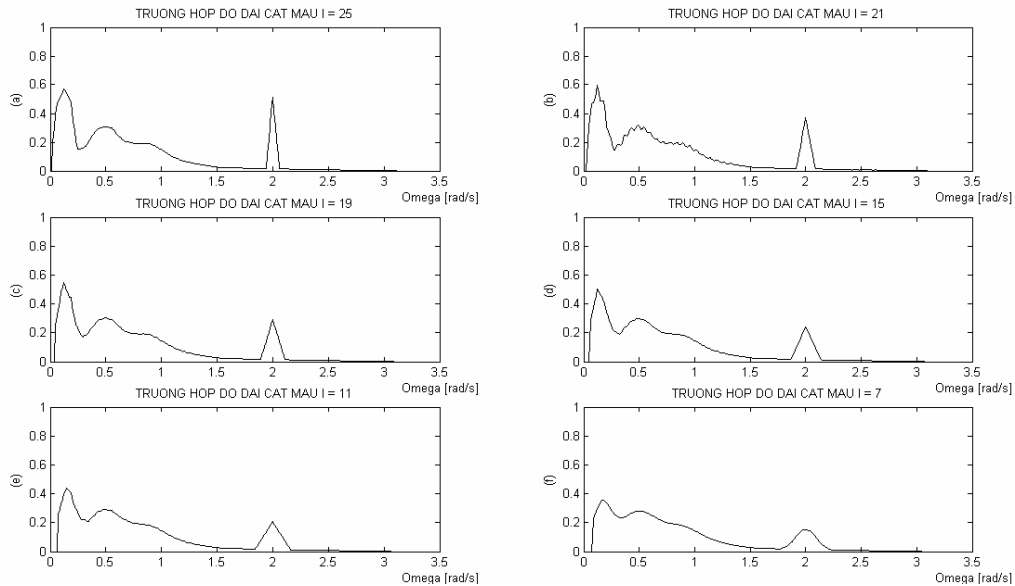
Tương tự, ta tính $\hat{\sigma}_{Y_2}^2$ như sau $\hat{\sigma}_{Y_2}^2 = \frac{1}{100} \sum_{k=1}^{100} (Y_k(2) - \bar{Y}_2)^2$. Qua thực nghiệm, kết quả thu được là $\hat{\sigma}_{Y_1}^2 \approx \hat{\sigma}_{Y_2}^2 \approx \hat{\sigma}_{Y_3}^2 \approx \dots \approx \hat{\sigma}_{Y_{256}}^2 \approx 0,01$. Kết quả này phù hợp với giả thiết các biến ngẫu nhiên Y_1, Y_2, \dots là độc lập cùng phân phối. Như trên ta đã gọi σ_Y^2 là phương sai của biến ngẫu nhiên Y_n với $n = 1 \div 256$.

Bảng 1

k	1	2	3	4	5	6	7	8	9	10
$Y_k(1)$	0,213	0,194	0,241	0,293	0,268	0,219	0,139	0,152	0,435	0,157
k	11	12	13	14	15	16	17	18	19	20
$Y_k(1)$	0,74	0,139	0,180	0,404	0,084	0,145	0,217	0,183	0,243	0,087
k	21	22	23	24	25	26	27	28	29	30
$Y_k(1)$	0,205	0,064	0,174	0,226	0,114	0,211	0,226	0,381	0,189	0,301
k	31	32	33	34	35	36	37	38	39	40
$Y_k(1)$	0,284	0,116	0,247	0,182	0,258	0,082	0,353	0,117	0,257	0,140
k	41	42	43	44	45	46	47	48	49	50
$Y_k(1)$	0,243	0,253	0,095	0,185	0,439	0,129	0,157	0,163	0,105	0,256
k	51	52	53	54	55	56	57	58	59	60
$Y_k(1)$	0,172	0,325	0,025	0,341	0,341	0,155	0,323	0,278	0,246	0,191
k	61	62	63	64	65	66	67	68	69	70
$Y_k(1)$	0,104	0,076	0,279	0,199	0,005	0,127	0,119	0,098	0,205	0,308
k	71	72	73	74	75	76	77	78	79	80
$Y_k(1)$	0,342	0,292	0,254	0,199	0,031	0,166	0,328	0,215	0,319	0,012
k	81	82	83	84	85	86	87	88	89	90
$Y_k(1)$	0,183	0,059	0,360	0,215	0,229	0,315	0,379	0,175	0,073	0,292
k	91	92	93	94	95	96	97	98	99	100
$Y_k(1)$	0,104	0,185	0,082	0,161	0,088	0,035	0,296	0,272	0,125	0,368

2.3.3. Thực nghiệm 1

Để đánh giá độ chính xác, ta sử dụng nhiều mô phỏng theo công thức (3). Các tham số của nhiễu như $C = 1,4$ và $\omega_0 = 2\text{rad/s}$. Hình vẽ 6 dưới đây mô tả kết quả làm trơn phổ bằng công thức (7) với $l = 25, 21, 19, 15, 11, 7$.



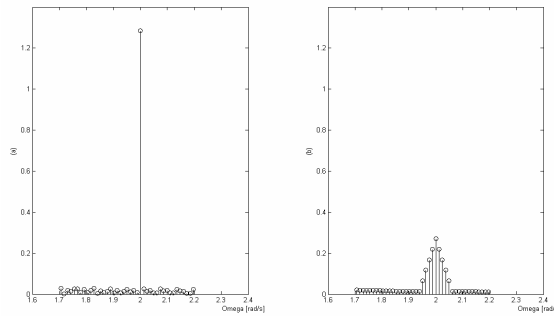
Hình 6. Các kết quả làm trơn phổ của tín hiệu điện tim nhiễu nhiễu

Với $\sigma_Y^2 = 0,01$ và $\rho = 1,5$ được xác định bằng phương pháp thực nghiệm như trên, Bảng 2 giúp kiểm tra điều kiện nêu trong bổ đề.

Bảng 2

l	25	21	19	15	11	7
$\frac{\sigma_Y^2}{\rho^2}(1 + \frac{1}{l})$	0,0046	0,0047	0,00472	0,00475	0,0048	0,0051

Ta thấy rằng, với $\sigma_Y^2 = 0,01$ và $\rho = 1,5$ khi l tăng lên, nhưng $\frac{\sigma_Y^2}{\rho^2}(1 + \frac{1}{l})$ giảm không đáng kể, đều không thoả mãn điều kiện trong bổ đề. Tuy nhiên vì $\frac{\sigma_Y^2}{\rho^2}(1 + \frac{1}{l})$ nhận các giá trị xung quanh 0,005 nên $Prob[|Y_n - M_{n,l}| < \rho] \geq 0,995$, do đó sự kiện sai số nằm trong giới hạn có thể xảy ra với ác suất rất cao. Hình 7 dưới đây mô tả trường hợp độ lệch sau khi làm trơn nằm trong giới hạn cho phép, trong đó Hình 7(a) là điểm đột biến tại $\omega_0 = 2\text{rad/s}$ của $\mathbf{y}(n)$, Hình 7(b) là điểm đột biến tại $\omega_0 = 2\text{rad/s}$ của $f(n)$.



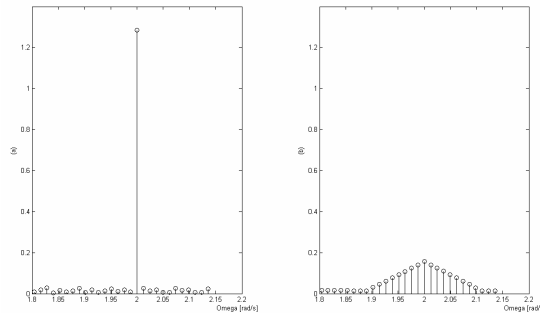
Hình 7. Điểm đột biến trên phổ của tín hiệu điện tim nhiễu nhiều
Hình 7(a). Trước khi làm trơn, Hình 7(b) sau làm trơn

Ta thấy rằng $|\mathbf{y}(\omega_0) - \mathbf{f}(\omega_0)| = 1$, với $\rho = 1,5$ điều này có nghĩa là tại $\omega_0 = 2\text{rad/s}$ sai số giữa trước và sau làm trơn nằm trong giới hạn cho phép.

Với $\sigma_Y^2 = 0,01$ và $\rho = 0,15$, tương tự, Bảng 3 giúp kiểm tra điều kiện nêu trong bổ đề.

Bảng 3

l	25	21	19	15	11	7	3
$\frac{\sigma_Y^2}{\rho^2}(1 + \frac{1}{l})$	0,4622	0,4655	0,4678	0,4741	0,4938	0,5079	0,5926



Hình 8. Điểm đột biến trên phổ của tín hiệu điện tim nhiễu nhiều
Hình 8(a). Trước khi làm trơn, Hình 8(b). Sau làm trơn

Nếu chọn $l = 3$, thì $\frac{\sigma_Y^2}{\rho^2}(1 + \frac{1}{l}) = 0,5926$, tương ứng với

$$Prob[|Y_n - M_{n,l}| < 0,15] \geq 0,4074.$$

Sau đó các thực nghiệm được tiến hành qua 100 dữ liệu sóng điện tim khác nhau, như đã mô tả ở phần tính phương sai bằng thực nghiệm ở trên. Thực hiện thống kê trên kết quả thu được có 58 trường hợp có độ lệch sau làm tròn vượt quá giới hạn cho phép, $|y(\omega_0) - f(\omega_0)| > 0,15$ như trong Hình 8.

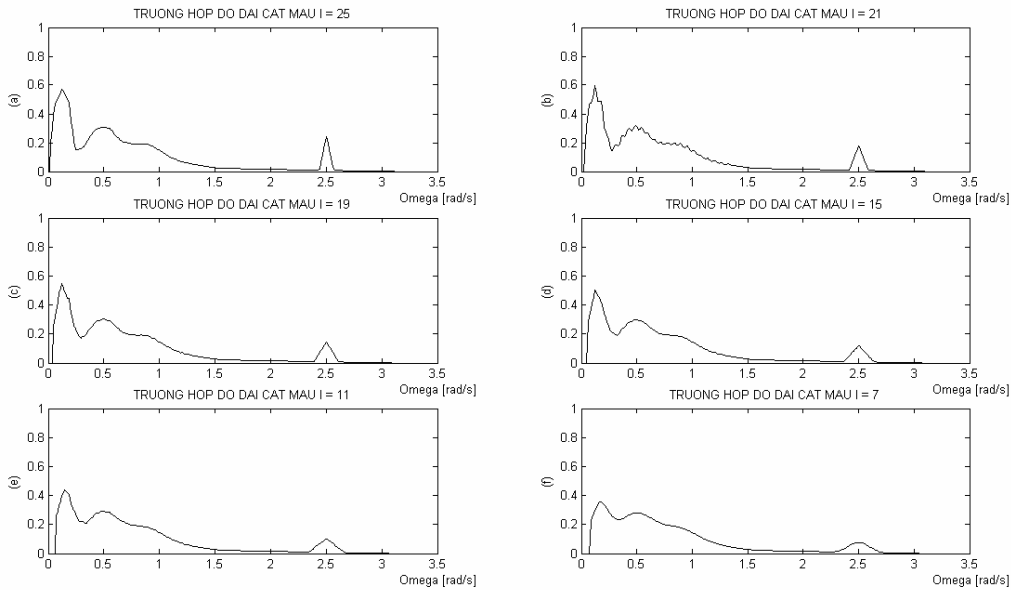
Sự kiện $|Y(\omega_0) - f(\omega_0)| < 0,15$ chiếm 42%, phù hợp với kết quả

$$Prob[|Y_n - M_{n,l}| < 0,15] \geq 0,4074$$

thu được ở trên.

2.3.4. Thực nghiệm 2

Tương tự như trên, để mô phỏng, các tham số của nhiễu $C = 1$ và $\omega_0 = 2,5$ rad/s. Hình 9 mô tả phổ của tín hiệu điện tim nhiễu nhiễu, kết quả làm tròn phổ bằng công thức (7) với $l = 25, 21, 19, 15, 11, 7$.



Hình 9. Các kết quả làm tròn phổ của tín hiệu điện tim nhiễu nhiễu

Trong các thực nghiệm này, với $\sigma_Y^2 = 0,01$ và $\rho = 2,5$ Bảng 4 sau đây mô tả mối quan hệ giữa độ dài cắt mẫu l và xác suất xảy ra sự kiện sai số lớn hơn 1,5 sau khi làm tròn.

Bảng 4

l	25	21	19	15	11	7	3
$\frac{\sigma_Y^2}{\rho^2} (1 + \frac{1}{l})$	0,00161	0,00162	0,00165	0,00168	0,0017	0,0018	0,0021

Bảng 4 cho ta thông tin là từ độ dài cắt mẫu nhỏ nhất, $l = 3$ đến độ dài $l = 25$, điều kiện trong bổ đề vẫn được thoả mãn.

3. KẾT LUẬN

Kết quả đạt được:

- + Xác định được điều kiện để chắc chắn xảy ra sự kiện sai số nằm trong giới hạn cho phép khi làm trơn phổ của tín hiệu điện tim nhiễu nhiễu. Kết quả này góp phần đảm bảo độ chính xác cho giải thuật dò tần số của nhiễu, được đề xuất trong [6].
- + Với đề xuất coi độ lớn của phổ của tín hiệu điện tim nhiễu nhiễu tại mỗi tần số là 1 biến ngẫu nhiên, bằng thực nghiệm đã xác định được các biến ngẫu nhiên này có cùng phương sai. Phát hiện này cho phép sử dụng các công cụ xác suất và Thống kê để đạt được kết quả trong công thức (12).

TÀI LIỆU THAM KHẢO

- [1] M. Akay, *Biomedical Signal Processing*, Academic Press, 1994.
- [2] John G. Proakis, Dimitris G. Manolakis, *Digital signal Processing, Principle, Algorithms and Applications*, Prentice Hall International Inc, 1996.
- [3] B. Widrow, Samuel D. Stearns, *Adaptive Signal Processing, Englewood Cliffs*, Prentice-Hall, 1985.
- [4] Daniel Olgún Olgún, Frantz Bouchereau, Sergio Martínez, Adaptive notch filter for EEG signals based on the LMS algorithm with variable step-size parameter, *Proceedings of the Conference on Information Sciences and Systems*, The John Hopkins University, March 16-18/2005.
- [5] Ju-Won Lee and Gun-Ki Lee, Design of an Adaptive Filter with a Dynamic Structure for ECG Signal Processing, *International Journal of Control, Automation and Systems* **3** (1) (2005) 137–142.
- [6] Pham Tran Nhu, Hoang Manh Ha, Adjustment in central frequency of adaptive notch filter base on wavelet transform in frequency domain, *Proceeding of the Second International Conference on Communications and Electronics (HUT ICCE)*, Hoian, Vietnam, 2008 (590–593).
- [7] S. Mallat, *A Wavelet Tour of signal Processing*, Academic Press, New York, UK, 1998.
- [8] Cuiwei Li et. al., Detection of ECG Characteristic Points Using Wavelet Transforms, *IEEE Transactions on Biomedical Engineering* **42** (1) (1995) 21–28.
- [9] Trần Mạnh Tuấn, *Xác suất & Thống Kê - Lý thuyết và Thực hành tính toán*, NXB Đại học Quốc gia Hà Nội, 2004.

Nhận bài ngày 19 - 7 - 2011
 Nhận lại sau sửa ngày 3 - 11 - 2011