TẠP CHÍ TIN HỌC VÀ ĐIỀU KHIỂN HỌC TẬP I (1991), Số 3 (1-6)

KHẢO SÁT ĐƯỜNG CONG TÍN HIỆU BẰNG PHƯƠNG PHÁP DELTA VÀ KỸ THUẬT NHÓM ĐIỂM

LÊ TỰ THÀNH
Viện Điện tử Tin học
HOÀNG KIẾM
Viện Tin học

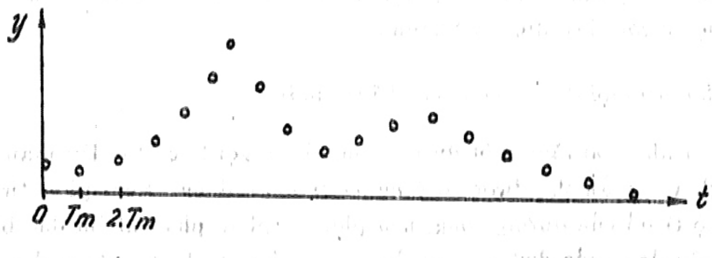
Phương pháp Delta được xây dựng trên cơ sở đánh giá độ biến thiên của đường cong tín hiệu trên một nhóm các điểm lấy mẫu bằng một giá trị xác định (xem [3]). Trên cơ sở những kết quả đã đạt được, chúng ta sẽ tập trung vào giải quyết những vấn đề căn bản của kỹ thuật chọn nhóm là: xác định số điểm trong mỗi nhóm và xác định số điểm tách trong số đó. Từ đó, người ta có thể áp dụng các phương pháp nhận dạng dựa trên tri thức (xem [1]) để xử lý chúng.

I – NGUYÊN TẮC PHÂN NHÓM

Khi khảo sát một đường cong tín hiệu bằng kỹ thuật vi xử lý, có nghĩa là người ta phải thực hiện việc lấy mẫu đường cong tín hiệu đó. Bằng cách này, người ta thu được ảnh của đường cong tín hiệu trên hệ tọa độ biên độ - thời gian. Do tính lặp của chu kỳ lấy mẫu Tm, nên thứ tự của điểm tín hiệu đã cho ta giá trị về thời gian. Từ đó ảnh thu được chỉ được khảo sát theo một chiều là biên độ. Nó được lưu dưới dạng một dãy số :

$$\{y_k | y_k = h(k.Tm).y(k.Tm) \; ; \quad k = 1, 2, ...\} \tag{1}$$

trong đó $y(t)$ là biểu diễn biên độ tín hiệu theo thời gian t, còn hàm $h(t)$ biểu diễn tác động của quá trình lấy mẫu lên giá trị tín hiệu thu nhận được.



Hình 1. Biểu diễn đường cong tín hiệu dưới dạng dãy (1)

Theo [4], sai số gây ra do tác động của quá trình lấy mẫu có thể điều chỉnh về giá trị vô cùng bé. Do vậy, trong quá trình khảo sát ảnh của đường cong tín hiệu, chúng ta sẽ xem (1) là ảnh đúng của tín hiệu.

1. Tư tưởng khảo sát bằng đạo hàm

Bản chất của việc nhận dạng đường cong tín hiệu là nhận biết các biến đổi về giá trị biên độ kết hợp với dáng điệu của đường cong. Do phương pháp Delta được xây dựng để nhận dạng trước hết những đường cong tín hiệu đã được nhận dạng bằng mắt, nên việc nhận ra các đoạn đường cong cơ sở là rất quan trọng (xem [3]).

Trong giải tích toán, khảo sát một đường cong luôn luôn là vấn đề tính đạo hàm của đường cong đó. Như chúng ta đều biết, đó là giá trị được xác định bằng gia số hàm số, trong trường hợp này là biên độ, chia cho gia số của đối số, trong trường hợp này là thời gian, khi gia số của đối số tiến tới 0. Với dãy (1), giới hạn của gia số thời gian chỉ có thể giảm tới Tm. Như vậy có thể nói rằng tư tưởng xây dựng phương pháp Delta hoàn toàn phù hợp với việc khảo sát sự biến thiên của hàm số truyền thống trong giải tích toán trên cơ sở đường cong tín hiệu không liên tục trên máy tính.

2. Phương pháp phân nhóm đặc biệt

Xuất phát từ tư tưởng nêu trên, cách khảo sát đơn giản nhất là xác định gia số biên độ từ hai điểm liên tiếp nhau. Từ dãy (1) ta có dãy gia số:

$$\{\Delta_k | \Delta_k = y_{k+1} - y_k \ ; \quad k = 1, 2, \dots\} \tag{2}$$

Đây chính là cách phân nhóm đặc biệt, áp dụng cho trường hợp số điểm trong nhóm bằng 2 và số điểm tách bằng 1, mà ta sẽ gọi tắt là nhóm đặc biệt. Đặc biệt trước hết vì đó là nhóm nhỏ nhất có thể nhóm được khi áp dụng phương pháp Delta, và vì đó cũng là hiển nhiên nhất. Phương pháp phân nhóm đặc biệt có các tính chất sau:

a) Đảm bảo độ tỉ mỉ của đường cong được lấy mẫu.

Như chúng ta đã biết (xem [4]), rằng chỉ có thể khảo sát đường cong được lấy từ mẫu tỉ mỉ tới các biến đổi tần số $F_{max} \leq Fm/2$, với Fm là tần số lấy mẫu. Do cách phân nhóm đặc biệt, số nhóm được phân bằng số điểm lấy mẫu trừ 1, với số điểm lớn, việc phân nhóm có thể xem là không làm giảm tần số lấy mẫu. Nhờ vậy, mọi biến đổi trên đường cong có tần số nhỏ hơn F_{max} đều có thể phát hiện được theo định lý Shannon.

b) Tồn tại đối với mọi đường cong có thể lấy mẫu.

Trước hết cần nhắc lại rằng một đường cong được gọi là có thể lấy mẫu với một thiết bị số hóa A/D có nghĩa là mẫu thu được có ý nghĩa để nhận dạng. Xuất phát từ tính chất *a*), do không làm giảm độ tỉ mỉ của đường cong, nên phương pháp phân nhóm đặc biệt sẽ không làm mất ý nghĩa để nhận dạng của đường cong được lấy mẫu. Vì lẽ đó, nhóm đặc biệt luôn tồn tại đối với mọi đường cong có thể lấy mẫu.

c) Số nhóm tạo được đạt cực đại.

Rõ ràng số nhóm tạo ra sẽ đạt cực đại, đó là điều hiển nhiên. Do vậy, trong một số trường hợp sẽ là dư thừa so với các nhóm có số điểm lớn hơn.

d) Chịu hoàn toàn ảnh hưởng của nhiễu có tần số nhỏ hơn 2.Tm.

Tính chất này cũng là một hệ quả tất yếu của tính chất *a*). Khi mọi biến đổi của đường cong với tần số nhỏ hơn hay bằng tần số Shannon đều có thể nhận dạng, thì nhiễu trong vùng tần số đó cũng không thể loại trừ.

3. Phương pháp Delta

Phương pháp Delta (xem [3]) xác lập một ánh xạ tích từ các nhóm N điểm lên một tập các số nguyên. Ta gọi ánh xạ tích đó là hàm Delta f được xác định như sau :

$$f = f1.f2 \quad \text{với : } f2_k = \frac{1}{N}\sum_{i=1}^{N}(y_{k+i} - y_k) \tag{3}$$

Từ giá trị thực của $f2$, chọn các ngưỡng $ng_1, ng_2, \ldots$ với $ng_{i-1} < ng_i$ để xác định giá trị của f1 theo quy tắc :

$$f1_k = \begin{cases} i & ng_{i-1} \le f2_k < ng_i \\ 0 & |f2_k| \le ng_1 \\ -i & -ng_{i-1} \ge f2_k > -ng_i \end{cases} \tag{4}$$

Giá trị của ánh xạ f, hay là hàm Delta, cho thấy mức độ nghiêng của đoạn đường cong tạo bởi nhóm điểm một cách tương đối theo số lượng ngưỡng. Trong phạm vi bài này, chúng ta chỉ quan tâm tới khía cạnh chọn số điểm N sao cho hợp lý.

II – VẤN ĐỀ PHÂN NHÓM ĐỂ TIẾN HÀNH KHẢO SÁT

Trong thực tế nhận dạng đường cong tín hiệu trên cơ sở tri thức chuyên gia, do tần số lấy mẫu Fm lớn hơn tần số của những biến đổi trên đường cong tín hiệu mà các chuyên gia có thể nhận dạng được bằng mắt nhiều lần, nên ít khi người ta áp dụng nhóm đặc biệt. Đối với các phân nhóm bất kỳ, người ta cần phải giải quyết hai vấn đề, đó là số điểm N trong mỗi nhóm và số điểm tách Nt trong số đó.

1. Yêu cầu nhận dạng đúng và số điểm N

Nhìn một cách cảm tính, thì trước hết phân nhóm phải đảm bảo cho ta khả năng nhận dạng được mọi biến đổi đáng quan tâm trên đường cong tín hiệu đã được lấy mẫu. Căn cứ vào cách phân nhóm và định lý Shannon người ta đã đưa ra giới hạn của việc phân nhóm (xem [3]). Tuy vậy do phải xác định đúng tần số F_{max}, nên các giới hạn đó ít có ý nghĩa thực tế. Xuất phát từ thực tế nhận dạng các đường cong tín hiệu trên cơ sở tri thức chuyên gia, chúng ta có mệnh đề sau :

Mệnh đề 1. Chỉ có thể nhận dạng đúng một đoạn đường cong cơ sở bất kỳ nếu và chỉ nếu có thể chia đoạn đó thành M nhóm, với $M \geq 3$.

Chứng minh: Theo định nghĩa các đoạn đường cong cơ sở gồm có ba dạng căn bản là : thẳng (hay đều), lồi và lõm như trên hình 2.1, áp dụng kết quả của mệnh đề nhận dạng (xem [3]), chúng ta nhận thấy cần phải chia được ít nhất hai nhóm trên mỗi đoạn.

a) Thẳng b) Lồi c) Lõm

Hình 2. Các đoạn đường cong cơ sở

Giả sử, giá trị hàm Delta tại các nhóm đó là F_1 và F_2, khi đó đoạn sẽ là :

- thẳng nếu $F_1 = F_2$,
- lồi nếu $F_1 > F_2$,
- lõm nếu $F_1 < F_2$.

Ngược lại, khi chỉ chia được một hay một phần nhóm, khi đó chỉ có thể nhận được nhiều nhất một giá trị f của hàm Delta. Như vậy không thể nhận dạng đúng được đoạn đường cong cơ sở do mối quan hệ giữa hàm Delta và dạng của các đoạn đường cong cơ sở là một - một.

Tuy nhiên, do việc phân nhóm được thực hiện cho toàn bộ cả đường cong, nên có thể đầu nhóm thứ nhất trên đoạn đường cong không trùng với đầu đoạn. Chúng ta dễ dàng chứng minh rằng, nếu có đủ số điểm để phân ra M nhóm trên đoạn đường cong bất kỳ, thì sẽ có ít nhất $M - 1$ nhóm nằm hoàn toàn trong đoạn đó không phụ thuộc vào điểm bắt đầu nhóm. Ta có điều cần phải chứng minh.

2. Khả năng lọc nhiễu và số điểm N

Thực tế, trong quá trình lấy mẫu khó tránh khỏi nhiễu sinh ra do môi trường. Người ta phải xem xét hai loại nhiễu căn bản là nhiễu ngẫu nhiên và nhiễu có chu kỳ (xem [2], [4]).

Nhiễu ngẫu nhiên sinh ra, ví dụ do những rung động đột xuất hay biến đổi đột ngột điện thế trong thời gian rất ngắn v.v... Khi biên độ nhiễu lớn hơn so với biên độ cực đại của đường cong tín hiệu B_{max}. Trong phần này chúng ta sẽ không đi vào các phương pháp lọc nhiễu chung đã được công bố ở nhiều tài liệu khác, mà chỉ đi sâu vào khả năng khử nhiễu của phương pháp Delta.

Từ cách xây dựng hàm Delta, tự phương pháp có khả năng làm giảm ảnh hưởng của nhiễu một cách tuyến tính. Mặt khác, nếu xác suất xuất hiện ngẫu nhiên tại một điểm bất kỳ là P, xác suất tồn tại của nhiễu ngẫu nhiên là Pn, giá trị ảnh hưởng của nhiễu là A thì với nhóm N điểm ta có mệnh đề sau.

Mệnh đề 2. Qua biến đổi tạo bởi hàm Delta, giá trị ảnh hưởng A của nhiễu ngẫu nhiên và xác suất tồn tại của nhiễu được đánh giá bởi các công thức:

$$A \leq B_{max}/N \tag{5}$$

$$Pn = 2P/N \tag{6}$$

Chứng minh: Theo (3), giá trị của $f2$ chỉ phụ thuộc vào biên độ điểm đầu và cuối mỗi nhóm. Giả sử nhiễu ngẫu nhiên rơi vào điểm đầu, khi đó (dấu $\prime$ giả sử giá trị có nhiễu):

$$f2'_k = (y_{k+N} - y_k)/N = (Bn_k + y_{k+N} - y_k)/N = f2_k + Bn/N$$

Khi nhiễu ngẫu nhiên rơi vào điểm cuối ta có :

$$f2'_k = (y_{k+N} - y_k)/N = (y_{k+N} - Bn - y_k)/N = f2_k - Bn/N$$

Như vậy, ảnh hưởng của nhiễu về biên độ là:

$$|A| = Bn/N \leq B_{max}/N$$

Mặt khác, do cách xây dựng ánh xạ fi, nếu nhiễu ngẫu nhiên rơi vào điểm bất kỳ trong nhóm trừ điểm đầu và cuối đều bị triệt tiêu. Do không có ràng buộc gì đối với việc chọn điểm đầu tiên cho việc nhóm điểm, nên xác suất để một điểm nào đó rơi vào đầu hoặc cuối nhóm sẽ là $2/N$. Như vậy, xác suất tồn tại của nhiễu ngẫu nhiên qua phép biến đổi Delta sẽ là xác suất có điều kiện sau:

$$Pn = P\,.\,2/N$$

Từ mệnh đề 2 chúng ta nhận thấy rằng, nếu có thể chọn số điểm N trong mỗi nhóm càng lớn, thì ảnh hưởng về biên độ và xác suất gặp nhiễu ngẫu nhiên càng giảm.

Đối với nhiễu có chu kỳ, ví dụ như nhiễu 50 Hz của lưới điện dân dụng, giả sử hàm gây nhiễu là $nh(t)$. Gọi điểm đầu nhóm là ở thời điểm t_1 thì thời điểm cuối nhóm sẽ là $t_1 + (N-1).Tm$, ta có:

$$\begin{aligned} f2'_k = y_{k+N} - y_k &= -nh(t_1).y_k + nh(t_1 + (N-1).Tm).y_{k+N} \\ &= nh(t_1).(y_{k+N} - y_k) - y_{k+N}.[nh(t_1) - nh(t_1 + (N-1).Tm)] \\ &= nh(t_1).f2_k + y_{k+N}.[nh(t_1) - nh(t_1 + (N-1).Tm)] \end{aligned}$$

Từ biểu thức trên chúng ta nhận thấy rằng chỉ trong trường hợp đặc biệt, khi $(N-1).Tm$ là bội của chu kỳ nhiễu, thì mới triệt tiêu được toán hạng thứ hai và

$$f2' = nh(t_1).f2 \quad \text{với } nh(t_1) = \text{const}.$$

Như vậy, ta có thể kết luận rằng phương pháp Delta hầu như không có khả năng lọc nhiễu có chu kỳ và chúng phải được xử lý bằng các phương pháp khác.

3. Quan hệ giữa N, Nt và thời gian tính toán

Để giải quyết vấn đề phân nhóm, người ta có thể chọn $Nt = N$. Điều đó có nghĩa là các nhóm hoàn toàn tách nhau. Tuy nhiên, cũng có những trường hợp buộc phải chọn $Nt < N$. Ví dụ ta phải chọn Nt vừa để đáp ứng mệnh đề 1, nhưng với Nt nhỏ sẽ bỏ qua khả năng lọc nhiễu ngẫu nhiên, và do vậy nên chọn $N > Nt$ đủ lớn sao cho vẫn đáp ứng được mệnh đề 1. Trong trường hợp này rõ ràng cần có sự cân nhắc khi chọn các giá trị N và Nt sao cho hợp lý. Như vậy thực tế sẽ tồn tại các cách phân nhóm trùng nhau một số điểm.

Ngoài những ràng buộc được xác định trên cơ sở các mệnh đề 1 và 2 còn có ràng buộc về thời gian tính toán, đặc biệt là khi cần đáp ứng thời gian thực. Nếu gọi thời gian để thực hiện một phép tính cộng hay trừ là T_1, còn thời gian để thực hiện một phép tính nhân hay chia là T_2, thời gian để thực hiện việc tính giá trị của f với nhóm thứ k là T_n, thì ta có:

$$T_n = T_2 + T_1.(1 + [Ng/2]) = \text{const}$$

với Ng là số lượng ngưỡng dùng để xác định giá trị của ánh xạ f. Tổng thời gian để thực hiện việc tính giá trị của f với toàn bộ đường cong với Nd điểm là Ttc:

$$Ttc = T_n.[(Nd - N + Nt)/Nt] = T_n.(1 + [Nd - N)/Nt]) \quad (7)$$

Như vậy, thời gian tính toán tỷ lệ với số lượng nhóm được phân và với số lượng điểm Nd lớn (với Nd nhỏ, thì thường người ta cũng ít cần quan tâm tới thời gian tính toán), có thể coi là tỷ lệ với Nd/Nt. Mặt khác ta cũng phải lưu ý rằng các tính toán dùng để thực hiện công việc nhận dạng ở giai đoạn sau cũng tỷ lệ với số lượng giá trị của f, và do đó là với chính Nd/Nt. Chúng ta có thể nhận xét chung về việc chọn các giá trị N và Nt qua mệnh đề sau.

Mệnh đề 3. Thời gian thực hiện ánh xạ f không phụ thuộc vào số điểm trong mỗi nhóm N, mà chỉ phụ thuộc vào số điểm tách Nt.

III – TÓM LƯỢC

Những vấn đề về việc chọn cách phân nhóm trong phương pháp Delta dùng để nhận dạng các đường cong tín hiệu được lấy mẫu trên máy tính đã được trình bày một cách tương đối trọn vẹn. Đó là việc chọn số điểm trong nhóm sao cho đảm bảo nhận dạng được từng đoạn đường cong cần quan tâm. Đó là ý nghĩa của việc chọn nhóm đối với vấn đề lọc nhiễu ngẫu nhiên.

Trong thực tế, không ít những đường cong tín hiệu cần được nhận dạng với thời gian thực. Trong trường hợp này yếu tố thời gian trở thành một ràng buộc nữa đối với cách chọn nhóm. Và chúng ta đã chứng minh rằng chọn các nhóm trùng nhau không làm tăng thời gian tính toán cần thiết.

Nhận ngày 20 - 9 - 1991

(Xem tiếp trang 13)