

MỘT THUẬT TOÁN HIỆU QUẢ CHO BÀI TOÁN TÌM CẶP ĐIỂM KHÁC MÀU GẦN NHẤT TRONG TẬP ĐIỂM HAI MÀU TRÊN MẶT PHẪNG

Nguyễn Ngọc Trung*, Trần Thị Diệu Huyền†

1. Mở đầu

Bài toán tìm cặp điểm khác màu gần nhất trong tập điểm hai màu trên mặt phẳng thuộc dạng các bài toán tìm các cặp điểm gần nhất trong mặt phẳng. Bài toán đặt ra là : “Cho n điểm màu xanh và m điểm màu đỏ trên mặt phẳng, làm thế nào tìm được cặp điểm xanh – đỏ gần nhau nhất”. Bài toán trên có thể được giải một cách dễ dàng bằng cách lấy lần lượt từng cặp phần tử là các điểm xanh và và điểm đỏ, xác định khoảng cách giữa chúng và chọn ra cặp có khoảng cách nhỏ nhất. Thuật toán như vậy sẽ có độ phức tạp là $O(n.m)$ trong đó n là số điểm xanh và m là số điểm đỏ trên mặt phẳng. Một câu hỏi được đặt ra là liệu có thể xây dựng một thuật toán tốt hơn cho bài toán này ?

Chúng ta thấy rằng, trong chuyên ngành hình học tính toán, lược đồ Voronoi đóng một vai trò rất quan trọng trong việc giải quyết các bài toán tìm các cặp điểm gần nhất trên mặt phẳng. Điều này cũng dễ hiểu vì lược đồ Voronoi có những tính chất rất đặc trưng về khoảng cách, điều mà rất hay được dùng để rút ngắn thời gian tính toán, cũng như thời gian thực hiện các thuật toán giải những bài toán trên. Chính vì thế trong bài báo này, chúng tôi đã chọn lược đồ Voronoi để làm công cụ xây dựng thuật toán giải bài toán tìm cặp điểm khác màu gần nhất trong số tập điểm hai màu trên mặt phẳng.

Cấu trúc bài báo như sau : mục 2 dành cho việc giới thiệu sơ lược về lược đồ Voronoi và các tính chất của nó. Phần mô tả chi tiết về thuật toán giải quyết bài toán sẽ được trình bày trong mục 3. Phần cuối của bài báo sẽ là một số chứng

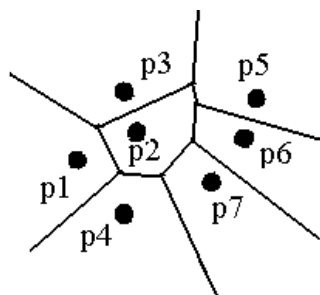
* ThS, Khoa Toán – Tin học, Trường Đại học Sư phạm Tp.HCM

† CN, Sinh viên Khoa Toán – Tin học Trường ĐHSPTp.HCM (Khoá 2001-2005)

minh lý thuyết về tính đúng đắn cũng như những đánh giá về độ phức tạp của thuật toán.

2. Lược đồ Voronoi và các tính chất

Định nghĩa 2.1. Đặt $P = \{p_1, p_2, \dots, p_n\}$ là tập gồm n điểm trong mặt phẳng. Lược đồ Voronoi là sự phân chia mặt phẳng thành n vùng (cell), mỗi vùng chứa một điểm p_i với tính chất một điểm q nằm trong vùng tương ứng với p_i nếu và chỉ nếu $\text{dist}(q, p_i) < \text{dist}(q, p_j), \forall j=1 \dots n$.



Hình 1. Lược đồ Voronoi của $P = \{p_1, p_2, \dots, p_7\}$.

Trong đó, $\text{dist}(q, p)$ là khoảng cách từ q đến p :

$$\text{dist}(q, p) = \sqrt{(q_x - p_x)^2 + (q_y - p_y)^2}$$

Ta kí hiệu : $\text{Vor}(P)$ là lược đồ Voronoi của P và $V(p_i)$ là vùng của $\text{Vor}(P)$ tương ứng với điểm p_i .

Sau đây chúng ta sẽ giới thiệu một số tính chất quan trọng của lược đồ Voronoi. Do khuôn khổ bài báo, chúng tôi sẽ không trình bày phần chứng minh của các tính chất này. Độc giả có thể tham khảo phần chứng minh của các tính chất này trong [1].

Định lý 2.1. Cho P là tập gồm n điểm trong mặt phẳng. Nếu các điểm này thẳng hàng, $\text{Vor}(P)$ sẽ gồm $(n - 1)$ đường thẳng song song. Ngược lại, $\text{Vor}(P)$ liên thông và các cạnh của nó là đoạn thẳng hoặc nửa đường thẳng.

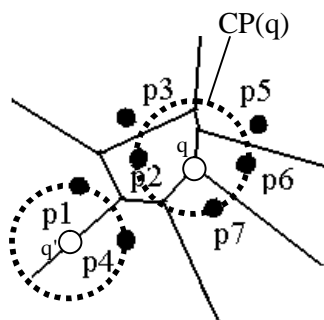
Điều có ý nghĩa lớn nhất trong định lý trên là khi các điểm của P là không thẳng hàng với nhau thì trong lược đồ $\text{Vor}(P)$ chỉ bao gồm tập các đoạn thẳng hoặc nửa đường thẳng chứ không thể có bất kỳ một đường thẳng (mở hai đầu) nào. Tính chất này giúp cho việc xây dựng cấu trúc dữ liệu để biểu diễn lược đồ $\text{Vor}(P)$ được đơn giản đi rất nhiều.

Dễ nhận thấy rằng, lược đồ Voronoi của tập điểm P được xây dựng từ những đường trung trực của các đoạn thẳng nối các cặp đỉnh trong P và hơn nữa,

các đỉnh của $Vor(P)$ cũng sẽ là giao điểm của các đường trung trực này. Tuy nhiên, không phải bất kì đường trung trực nào cũng là cạnh của $Vor(P)$ và cũng không phải bất kì giao điểm nào của các đường trung trực trên cũng là đỉnh của $Vor(P)$. Định lí sau đây sẽ cho ta thấy rõ hơn về nhận xét này.

Định lí 2.2. Cho lược đồ $Vor(P)$ của tập điểm $P = \{p_1, p_2, \dots, p_n\}$. Khi đó :

- a. Một điểm q là đỉnh của $Vor(P)$ nếu và chỉ nếu đường tròn rỗng lớn nhất có tâm là q – được gọi là $CP(q)$ chứa ít nhất ba điểm của P trên biên.
- b. Đường trung trực của đoạn thẳng $p_i p_j$ là một cạnh của $Vor(P)$ nếu và chỉ nếu có một điểm q trên đường trung trực này sao cho $CP(q)$ đi qua p_i, p_j và không chứa bất kì trạm nào khác.



Hình 2. Minh họa tính chất ảnh và cạnh của lược đồ Voronoi

Ví dụ. Theo hình 2, q là một đỉnh của $Vor(P)$ vì $CP(q)$ có chứa p_3, p_6, p_7 . Bên cạnh đó, đường trung trực của đoạn thẳng $p_1 p_4$ là một cạnh của $Vor(P)$ vì có một điểm q' trên đường trung trực này thỏa mãn $CP(q')$ đi qua p_1, p_4 và không đi qua bất kì đỉnh nào khác của P .

Thuật toán xây dựng lược đồ Voronoi của một tập n điểm trên mặt phẳng được xây dựng dựa trên các tính chất quan trọng trên. Thuật toán này sử dụng một dòng quét (sweep line) đi từ trên xuống và dần xác định các đỉnh và các cạnh của lược đồ Voronoi cần tìm. Độc giả quan tâm có thể tìm hiểu chi tiết của thuật toán này trong [1]. Định lí sau đây sẽ đề cập đến độ phức tạp của thuật toán xây dựng lược đồ Voronoi của một tập n điểm trên mặt phẳng.

Định lí 2.3. Thuật toán xây dựng lược đồ Voronoi của tập n điểm trong mặt phẳng sẽ có độ phức tạp về thời gian là $O(n \log n)$ và về không gian lưu trữ là $O(n)$.

Định lí 2.3. Lược đồ Voronoi của n điểm ($n \geq 3$) trong mặt phẳng có tối đa $(2n-5)$ đỉnh và $(3n-6)$ cạnh.

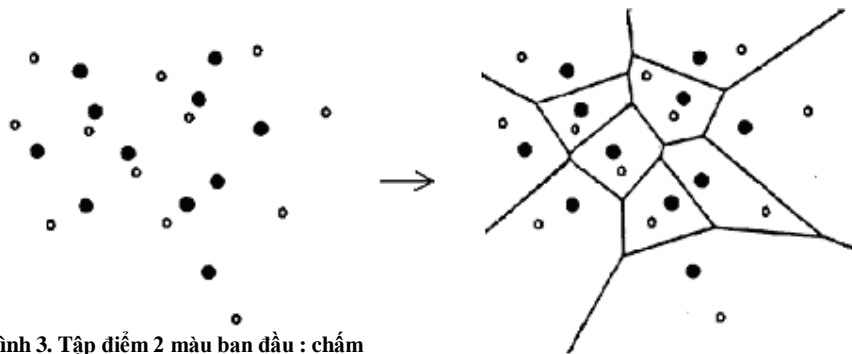
3. Thuật toán giải bài toán tìm cặp điểm khác màu gần nhất trong tập điểm hai màu trên mặt phẳng

Như đã trình bày ở trên, chúng ta có thể giải bài toán này bằng thuật toán vét cạn : kiểm tra hết tất cả các cặp điểm xanh – đỏ trong tập điểm cho trước, từ đó chọn ra cặp điểm có khoảng cách nhỏ nhất. Tuy nhiên, dễ dàng nhận thấy thuật toán này không tốt do trong số tất cả các cặp xanh – đỏ, có rất nhiều cặp không thể trở thành cặp điểm xanh – đỏ gần nhất : chẳng hạn như cặp điểm xanh – đỏ ở rất xa nhau trong khi giữa chúng lại có rất nhiều điểm xanh, đỏ khác.

Để cải thiện nhược điểm này, ta sẽ chỉ quan tâm tới những cặp xanh – đỏ là những “ứng cử viên” cho cặp điểm xanh – đỏ gần nhất. Cụ thể, ứng với mỗi điểm đỏ, ta chỉ quan tâm đến điểm xanh gần nó nhất (so với các điểm xanh còn lại) và đây sẽ là một cặp là “ứng cử viên” cho lời giải của bài toán đặt ra.

Dựa vào tính chất của lược đồ Voronoi “những điểm thuộc cùng một vùng (cell) sẽ gần với trạm tương ứng với vùng đó nhất so với những trạm khác”, chúng ta có thể đưa ra ý tưởng để giải quyết bài toán tìm cặp điểm khác màu gần nhất trong tập điểm hai màu trên mặt phẳng (gọi tắt là **Bài toán tập điểm hai màu trên mặt phẳng**) như sau :

- Đầu tiên, ta xây dựng lược đồ Voronoi cho tập n điểm màu xanh.
- Ứng với mỗi điểm đỏ, tìm điểm xanh gần nó nhất bằng cách xác định điểm đỏ đó đang ở trong vùng tương ứng với điểm xanh nào. Từ đó lập thành một cặp “ứng cử viên” cho lời giải.
- So sánh các cặp ứng cử viên, ta sẽ xác định cặp xanh – đỏ có khoảng cách ngắn nhất.

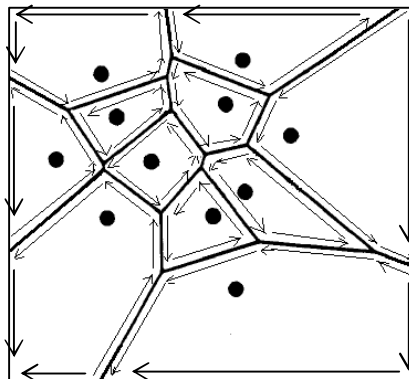


Hình 3. Tập điểm 2 màu ban đầu : chấm trắng là điểm đỏ, chấm đen là điểm xanh

Hình 4. Lược đồ Voronoi của tập điểm xanh đã được xây dựng xong

Sau đây, chúng ta sẽ cụ thể hoá ý tưởng của thuật toán này. Phần quan trọng nhất là làm thế nào để xác định được một điểm đỏ nằm trong vùng nào của lược đồ Voronoi của các điểm xanh. Để việc xác định này được thuận lợi, lược đồ Voronoi của các điểm xanh cũng phải được biểu diễn một cách phù hợp.

❖ **Biểu diễn lược đồ Voronoi của các điểm xanh bằng danh sách cạnh kép (double-edge list) :**



Hình 5. Minh hoạ biểu diễn lược đồ Voronoi bằng danh sách cạnh kép

Ta biểu diễn lược đồ Voronoi của các điểm xanh bằng danh sách cạnh kép như sau :

- Thêm vào lược đồ Voronoi một hình chữ nhật bao xung quanh tất cả các đỉnh và cạnh của nó. Như vậy lược đồ Voronoi bây giờ sẽ có thêm các đỉnh mới là 4 đỉnh của hình chữ nhật và giao điểm của các cạnh của nó

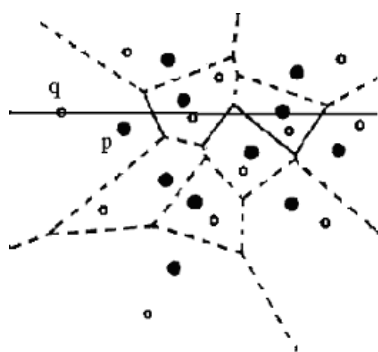
với cạnh của hình chữ nhật. Mặt khác, lược đồ Voronoi bây giờ chỉ bao gồm các đoạn thẳng (không còn nửa đường thẳng nữa).

- Mỗi cạnh của lược đồ Voronoi sẽ được biểu diễn bằng hai vector ngược chiều nhau sao cho trong một vùng, các vector luôn có chiều theo ngược chiều kim đồng hồ (xem hình 5).
- Như vậy mỗi vùng bây giờ đã là một đa giác độc lập và việc kiểm tra một điểm đỏ có nằm trong vùng nào đó hay không sẽ được đưa về một bài toán cơ bản của hình học : kiểm tra một điểm có nằm trong một đa giác hay không.

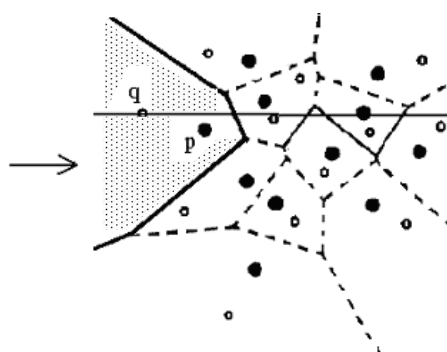
❖ **Thuật toán tìm cặp điểm khác màu gần nhất trong tập điểm hai màu trên mặt phẳng :**

Mặc dù chúng ta đã biết được cách xác định một điểm đỏ có nằm trong một vùng nào đó của lược đồ Voronoi các điểm xanh hay không nhưng điều đó không có nghĩa là mọi việc đều đã được giải quyết. Trong lược đồ Voronoi có rất nhiều vùng, do đó, ứng với một điểm đỏ, ta không thể kiểm tra hết tất cả các vùng để kết luận điểm đỏ này nằm trong vùng nào, thay vào đó, ta chỉ kiểm tra những vùng nào nằm “gần” điểm đỏ đó mà thôi.

Để thực hiện được ý tưởng trên, ta sử dụng một dòng quét (sweep-line) đi từ trên xuống dưới. Dòng quét sẽ lần lượt quét qua các điểm đỏ. Giả sử tại một thời điểm, dòng quét cắt một điểm đỏ, khi đó nó cũng sẽ cắt một số vector của lược đồ Voronoi. Trong số các vector đang bị dòng quét cắt, ta sẽ tìm được một vector nằm gần nhất bên phải điểm đỏ (khoảng cách được xác định trên dòng quét), vùng tương ứng với vector này sẽ chứa điểm xanh gần nhất với điểm đỏ mà dòng quét đang cắt.

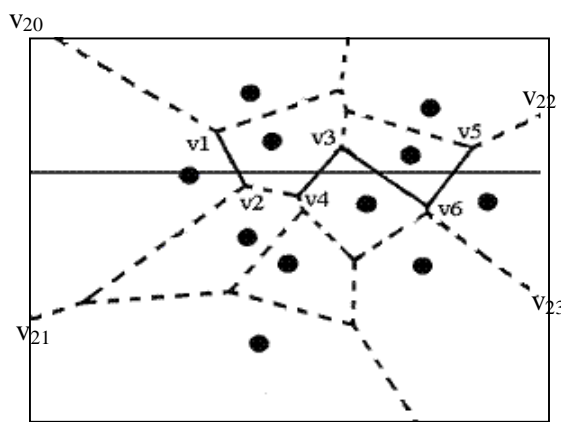


Hình 6. Dòng quét đi qua điểm đỏ q



Hình 7. Xác định được vùng chứa điểm xanh p gần với điểm đỏ q nhất

Nhằm xác định được vector nào nằm bên phải gần nhất so với điểm đỏ, chúng ta cần một cấu trúc lưu trữ các cạnh đang bị dòng quét cắt của lược đồ Voronoi. Cấu trúc này sẽ thường xuyên được cập nhật (thêm, xoá các cạnh, ...) đồng thời phải luôn đảm bảo các cạnh được tổ chức theo một thứ tự nhất định để việc tìm kiếm được dễ dàng, cấu trúc thích hợp nhất là cây nhị phân tìm kiếm cân bằng.



Hình 8. Cây nhị phân cân bằng lưu trữ các cạnh đang bị cắt bởi dòng quét



Để cập nhật cây lưu trữ các cạnh này, chúng ta sẽ cho dòng quét dừng lại tại các đỉnh của lược đồ Voronoi, tại các đỉnh này, một số cạnh không có khả năng cắt dòng quét nữa sẽ được xoá khỏi cây, một số cạnh mới sẽ được thêm vào cây.

Như vậy dòng quét không những dừng lại tại các điểm đỏ mà nó còn dừng lại tại các đỉnh của lược đồ Voronoi để cập nhật lại cây lưu trữ các cạnh đang bị

dòng quét cắt. Để dòng quét có thể dừng lại tại các điểm nói trên, ta sẽ xây dựng một hàng đợi Q gồm các điểm đỏ và các đỉnh của lược đồ Voronoi của điểm xanh. Các điểm trong hàng đợi này được sắp theo thứ tự giảm dần về tung độ (điểm có tung độ lớn nhất sẽ được dòng quét đi qua trước).

Tổng kết lại, thuật toán của chúng ta có thể được mô tả như sau :

Thuật toán TÌM_CẶP_ĐỈNH_KHÁC_MÀU_GẦN_NHẤT

Input : Tập P gồm n điểm xanh và m điểm đỏ trên mặt phẳng.

Output : Cặp đỉnh xanh – đỏ gần nhau nhất.

- Bước 1. Xây dựng lược đồ Voronoi cho tập n điểm màu xanh.
- Bước 2. Xây dựng hàng đợi Q gồm các đỉnh của Voronoi và các điểm đỏ - đây là hàng đợi ưu tiên theo tung độ của các điểm.
- Bước 3. Khởi tạo cây tìm kiếm cân bằng T lưu các cạnh đang bị cắt bởi đường quét.
- Bước 4. **WHILE** Q khác rỗng **DO** :
 - Lấy một đỉnh v từ Q.
 - **IF** v là điểm đỏ **THEN**
 - Gọi thủ tục XỬ_LÍ ĐIỂM ĐỎ(v)
 - ELSE**
 - Gọi thủ tục XỬ_LÍ ĐỈNH (v)
- END DO**

Thủ tục XỬ_LÍ ĐỈNH(v)

- Bước 1. Tìm cạnh cần xoá trong T (các cạnh có đỉnh thấp là v) và các cạnh ra khỏi T.
- Bước 2. Thêm các cạnh mới (các cạnh có đỉnh cao là v) vào đúng vị trí vừa xoá.

Thủ tục XỬ_LÍ ĐIỂM ĐỎ(v)

- Bước 1. Thực hiện tìm kiếm trong cây T. Kiểm tra vị trí của điểm đỏ so với nửa cạnh là giá trị của nút gốc, nếu điểm đỏ nằm bên trái thì qua kiểm tra con phải của nó, nếu điểm đỏ nằm bên phải thì qua nút con bên trái và cứ thế tiếp tục. Quá trình sẽ kết thúc khi không thể đi tiếp.

Bước 2. Gọi e là cạnh cuối cùng trước khi dừng lại. Kiểm tra xem điểm đỏ v nằm bên trái của nửa cạnh nào trong hai nửa cạnh tạo bởi cạnh e vừa tìm được, xác định ô chứa nửa cạnh vừa tìm được, từ đó xác định được điểm xanh p gần v nhất.

Bước 3. Tính khoảng cách giữa điểm đỏ v và điểm xanh p vừa tìm được và so sánh nó với cặp nhỏ nhất hiện tại, nếu nhỏ hơn thì gán nó là cặp gần nhất hiện tại.

Trên đây chúng tôi đã trình bày thuật toán tìm cặp điểm khác màu gần nhất trong tập điểm hai màu trên mặt phẳng. Sau đây là đánh giá về độ phức tạp về thuật toán.

Bổ đề 3.1. Thời gian chạy của thuật toán là $O(p \log p)$ với p là tổng số điểm xanh và điểm đỏ.

Chứng minh.

Thời gian thực hiện của thuật toán sẽ bao gồm thời gian xây dựng Voronoi và thời gian xử lý các sự kiện (cập nhật cây, xác định cặp nhỏ nhất hiện tại, ...).

Trước hết, nhận thấy rằng, thời gian xây dựng lược đồ Voronoi là $O(n \log n)$ như đã trình bày trong phần 2.

Duyệt từ trên xuống qua toàn bộ thuật toán và tính thời gian qua từng bước thực hiện thuật toán :

- Do số đỉnh của lược đồ Voronoi của n đỉnh xanh là không quá $2n-5$ nên thời gian xây dựng hàng đợi Q gồm các đỉnh của lược đồ Voronoi của n điểm xanh và m điểm đỏ sẽ không quá $O(p \log p)$ – cần nhắc lại là $p = m+n$.
- Thời gian lấy một phần tử khỏi Q là $O(1)$, do đó tổng thời gian để lấy p phần tử khỏi Q là $O(p)$.
- Thời gian xử lý sự kiện gặp đỉnh : tổng thời gian này bằng thời gian xây dựng cây, thời gian duyệt cây, thời gian thêm và xoá phần tử khỏi cây.

Do số cạnh của lược đồ Voronoi của n điểm xanh không vượt quá $3n-6$ nên tổng thời gian này không vượt quá $O(n \log n)$.

- Thời gian xử lý sự kiện gặp điểm đỏ : bằng thời gian duyệt cây, cộng thêm thời gian xác định ô cần tìm. Thời gian duyệt cây là $O(n \log n)$, thời gian xác định ô là hằng số, do đó thời gian để xử lý sự kiện gặp điểm đỏ là $O(n \log n)$.
- Vậy tổng thời gian thực hiện của thuật toán là $O(p \log p)$. □

4. Kết luận

Bài toán tìm cặp điểm khác màu gần nhất trong tập điểm hai màu trên mặt phẳng có rất nhiều ứng dụng thực tế. Đơn cử như ta có thể sử dụng thuật toán giải bài toán này để tìm hai trạm gần nhau nhất của hai hệ thống mạng khác nhau, nối chúng lại, tạo thành một mạng liên thông với chi phí ít tốn kém nhất, ...

Những nghiên cứu về những bài toán dạng này cũng được rất nhiều người quan tâm. Dựa trên những kết quả đạt được này, chúng ta có thể tiếp tục nghiên cứu, mở rộng bài toán theo các hướng sau đây :

- Mở rộng với bài toán tìm bộ các điểm khác màu gần nhau nhất trong tập điểm 3, 4 màu.
- Mở rộng bài toán tập điểm hai màu trong mặt phẳng thành bài toán tập điểm hai màu trong không gian : khi đó chúng ta phải sử dụng đến khái niệm lược đồ Voronoi trong không gian.

TÀI LIỆU THAM KHẢO

- [1]. Mark de Berg, Marc van Kreveld, Mark Overmars, Otfried Schwarzkopt (2000), *Computational Geometry Algorithms and Applications*, Springer Verlag, Berlin.
- [2]. Franco P.Preparata, Michael Ian Shamos (1985), *Computational Geometry – An Introduction*, Springer Verlag, Tokyo.
- [3]. O'Rourke, J (1993) : *Computational Geometry in C*, Cambridge University Press.

Tóm tắt :**Một thuật toán hiệu quả cho bài toán tìm cặp điểm khác màu gần nhất trong tập điểm hai màu trên mặt phẳng**

Bài toán tập điểm hai màu trong mặt phẳng được phát biểu như sau : “Cho n điểm màu xanh và m điểm màu đỏ trên mặt phẳng, làm thế nào tìm được cặp điểm xanh – đỏ gần nhau nhất”. Việc xác định lời giải cho bài toán này có thể được thực hiện tương đối dễ dàng bằng thuật toán vét cạn – kiểm tra hết tất cả các cặp điểm khác màu. Tuy nhiên, trong bài báo này chúng tôi sẽ trình bày một thuật toán tốt hơn rất nhiều để giải quyết bài toán này. Công cụ chủ yếu được sử dụng trong thuật toán của chúng tôi là lược đồ Voronoi trên mặt phẳng.

Từ khoá : Voronoi diagram, computational geometry, sweep-line algorithm.

Abstract :**An effective algorithm for the problem to find the nearest different color pair points in the set of two color points on the surface**

The problem of the set of two color points on the surface is suggested “Given n blue points and m red points on the surface, find the nearest blue – red pair point”. Identifying the solution to the problem may be conducted relatively easily with sweep-line algorithm – checking all different color pair points. However, in this article, a much better algorithm to solve this problem is presented. The major tool used in the algorithm is the Voronoi diagram on the surface.