

**ĐẠI HỌC THÁI NGUYÊN**  
**TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG**

**VŨ VĂN ANH**

**NGHIÊN CỨU**  
**KỸ THUẬT DÁN NHÃN ĐỐI TƯỢNG 2D**

**LUẬN VĂN THẠC SĨ KHOA HỌC MÁY TÍNH**

Thái Nguyên - 2015

## LỜI CAM ĐOAN

Luận văn là sự nghiên cứu các kiến thức mà học viên đã thu thập, tìm hiểu được trong quá trình học tập tại Trường Đại học Công nghệ thông tin và truyền thông- Đại học Thái Nguyên, dưới sự hướng dẫn, giúp đỡ của các thầy cô và bạn bè. Đặc biệt là sự hướng dẫn của thầy giáo hướng dẫn TS. Nguyễn Văn Huân

Học viên xin cam đoan nội dung bản luận văn này không phải là sản phẩm sao chép của bất kỳ tài liệu khoa học nào.

*Thái Nguyên, ngày 12 tháng 5 năm 2015*

**Học viên**

**Vũ Văn Anh**

## LỜI CẢM ƠN

Luận văn sẽ không thể hoàn thành nếu không có sự động viên, hỗ trợ hết mình của rất nhiều người. Trước hết tôi xin gửi lời cảm ơn sâu sắc đến TS Nguyễn Văn Huân người thầy đã chỉ bảo, giúp đỡ tận tình trong cả quá trình học tập, nghiên cứu và hoàn thiện luận văn.

Xin gửi lời cảm ơn đến các thầy cô giáo tại trường Đại học Công nghệ thông tin và truyền thông – Đại học Thái Nguyên, những người đã trang bị các kiến thức cơ sở, nền tảng cho việc nghiên cứu, tiếp thu những tri thức mới, mà từ đó tôi có thể hoàn thành tốt luận văn của mình.

Quá trình thực hiện đề tài không tránh khỏi những thiếu sót. Tôi hi vọng sẽ được sự góp ý chân thành từ phía các thầy, cô giáo, bạn bè, đồng nghiệp để đề tài nghiên cứu được hoàn thiện hơn  
Xin chân trọng cảm ơn!

## MỤC LỤC

LỜI CAM ĐOAN .....	i
LỜI CẢM ƠN .....	iii
MỤC LỤC .....	iv
DANH MỤC HÌNH VẼ TRONG LUẬN VĂN .....	vi
DANH MỤC CÁC KÝ HIỆU, CÁC CHỮ CÁI VIẾT TẮT .....	viii
MỞ ĐẦU .....	1
Chương I. TỔNG QUAN VỀ BÀI TOÁN DÁN NHÃN .....	4
1.1. Giới thiệu bài toán .....	4
1.2. Định nghĩa bài toán.....	5
1.3. Tư tưởng chung giải quyết bài toán dán nhãn .....	7
Chương 2. MỘT SỐ THUẬT TOÁN DÁN NHÃN ĐỐI TƯỢNG .....	9
2.1. Thuật toán dán nhãn điểm GFPL.....	9
2.2. Kỹ Thuật toán ELP .....	15
2.2.1. Thuật toán ELP.....	15
2.2.2. Thuật toán Fast ELP:.....	17
2.3. Giới thiệu kỹ thuật NLP (Node Label Placement) .....	23
2.4. Kỹ thuật MLP (Multiple Label Placement).....	24
2.4.1. Giới thiệu kỹ thuật MLP .....	24
2.4.2. Thuật toán Iterative .....	29
2.4.3. Thuật toán Flow-based.....	32
2.5. Kỹ thuật dán nhãn dựa vào hiệu chỉnh đối tượng.....	36
2.6. Thuật toán dán nhãn đường biên .....	37
2.6.1. Giới thiệu chung .....	37
2.6.2. Thuật toán.....	39
Chương 3. CHƯƠNG TRÌNH CÀI ĐẶT ỨNG DỤNG <b>Error! Bookmark not defined.</b>	6

3.1. Đánh giá thuật toán .....	466
3.2. Yêu cầu bài toán dán nhãn.....	49
3.3. Chương trình:.....	49
3.4. Kết quả thử nghiệm .....	50
KẾT LUẬN .....	51
TÀI LIỆU THAM KHẢO.....	53

## DANH MỤC HÌNH VẼ TRONG LUẬN VĂN

Hình 1.1. (a) Dán nhãn điểm, (b) Dán nhãn cạnh, (c) Dán nhãn vùng.....	6
Hình 2.1. Bản vẽ hướng nơi nhãn được định vị bằng các kỹ thuật phù hợp cho kỹ thuật GFLP. Các nhãn được đặt song song với trục ngang. Hộp màu xám là nút nhãn và hộp trắng cạnh nhãn.....	12
Hình 2.2. Đồ thị với các vị trí nhãn được dán cho mỗi cạnh .....	13
Hình 2.3. Vị trí nhãn có thể cho một điểm.....	13
Hình 2.4. Xác định các vị trí nhãn cho cạnh .....	18
Hình 2.5. (a) Một bản vẽ đơn giản với các vị trí nhãn cho mỗi cạnh (b) Các đồ thị phù hợp tương ứng .....	19
Hình 2.6. Một kết quả dán nhãn của thuật toán Fast ELP .....	21
Hình 2.7. Kết quả dán nhãn cạnh cho một bản vẽ trực giao có nhiều cạnh nằm ngang, áp dụng Fast ELP. ....	22
Hình 2.8. Bản vẽ hình tròn với các nhãn cạnh, nơi nhãn được phép chồng lên các đối tượng hình khác, được xây dựng bằng kỹ thuật ELP .....	23
Hình 2.9. (a) Phân nhãn thích hợp hơn. (b) Đặt nhãn gây hiểu nhầm. (c) Ràng buộc khoảng cách chặt chẽ. (d) Xác định ràng buộc tự do. ....	27
Hình 2.10. (a) Gán nhãn thích hợp hơn. (b) Việc gán nhãn chấp nhận được. (c) Việc gán nhãn gây hiểu nhầm.....	29
Hình 2.11. Một bản vẽ trực giao với vị trí hai nhãn mỗi cạnh sử dụng thuật toán Iterative.....	31
Hình 2.12. Một bản vẽ phân cấp với vị trí hai nhãn mỗi cạnh sử dụng thuật toán Iterative.....	32
Hình 2.13. Đồ thị Flowbased .....	33
Hình 2.14. Bản vẽ với vị trí hai nhãn mỗi cạnh bởi thuật toán Flow-based ...	34

Hình 2.15. Một bản vẽ vòng tròn với ba nhãn cho mỗi cạnh và nút được định vị bằng thuật toán Flow-based trên. Các ô màu trắng là các nhãn cạnh và các hộp đen là nút nhãn .....	35
Hình 2.16. Bốn bước chính trong kĩ thuật lần, đánh nhãn thành phần .....	40
Hình 2.17. P là điểm đầu của đường biên ngoài, 1 là điểm đen chưa được đánh nhãn .....	41
Hình 2.18. (A) P là điểm khởi đầu của đường biên trong, cũng nằm trên đường biên ngoài. (B) P là điểm khởi đầu của đường biên trong nhưng không nằm trên đường biên ngoài. 1: điểm đen chưa được đánh nhãn, $\Delta$ : điểm đen đã được đánh nhãn. ....	41
Hình 2.19. P là điểm chưa được đánh nhãn, điểm lân cận trái N đã được đánh nhãn	42
Hình 2.20. Các điểm trắng bao quanh đối tượng được đánh dấu bởi số âm...	43
Hình 2.21. Các điểm trắng quanh đối tượng được đánh dấu âm khi đường biên đã được lần .....	43
Hình 2.22. Lần đường biên của một đối tượng .....	44
Hình 2.23. (A) đánh số thứ tự các điểm lân cận của P từ 0 tới 7; (B) Nếu điểm biên trước P nằm ở 3 thì bắt đầu tìm kiếm từ 5 .....	44
Hình 2.24. Lần đường biên trong, ngoài đồng thời đánh dấu các điểm trắng bao quanh .....	45
Hình 2.25. Ví dụ 1 điểm biên P nằm trên 4 đường biên .....	46
Hình 2.26. (A) Nếu P nằm trên đường biên trong $\partial$ , điểm lân cận trái Q là điểm trắng thì $\partial$ phải có điểm U trên Q và P. (B) Ngoài ra, nếu đường kẻ dọc L qua Q không giao với phi trên Q thì Q là điểm nằm ngoài thành phần (đối tượng đang xét) và P là điểm thuộc đường biên ngoài. ....	48

## DANH MỤC CÁC KÝ HIỆU, CÁC CHỮ CÁI VIẾT TẮT

GFLP	Graphical Feature Label Placement
ELP	Edge Label Placement
MLP	Multiple Label Placement
NLP	Node Label Placement



## MỞ ĐẦU

### 1. Lí do chọn đề tài:

Công nghệ thông tin đã ở một bước phát triển cao đó là số hóa tất cả các dữ liệu thông tin, luân chuyển mạnh mẽ và kết nối tất cả chúng ta lại với nhau. Mọi loại thông tin, số liệu âm thanh, hình ảnh có thể được đưa về dạng kỹ thuật số để bất kỳ máy tính nào cũng có thể lưu trữ, xử lý và chuyển tiếp cho nhiều người. Những công cụ và sự kết nối của thời đại kỹ thuật số cho phép chúng ta dễ dàng thu thập, chia sẻ thông tin và hành động trên cơ sở những thông tin này theo phương thức hoàn toàn mới, kéo theo hàng loạt sự thay đổi về các quan niệm, các tập tục, các thói quen truyền thống, và thậm chí cả cách nhìn các giá trị trong cuộc sống. Công nghệ thông tin đã mang lại cho con người những thành tựu to lớn trong nhiều lĩnh vực đời sống như xây dựng kiến trúc, y tế, giáo dục, quảng cáo... Đặc biệt trong lĩnh vực khoa học, nhiều ngành cần có sự hỗ trợ của công nghệ thông tin .

Một trong những ứng dụng quan trọng trong ngành công nghệ thông tin, đặc biệt trong các đối tượng đồ họa 2D như bản đồ địa lí, hình ảnh... là việc dán nhãn cho các đối tượng đó. Tự động đặt vị trí nhãn là một lĩnh vực trong trực quan hóa thông tin. Nhãn là các đoạn văn bản nhằm truyền đạt thông tin, làm rõ ý nghĩa của các cấu trúc phức tạp được biểu diễn ở dạng đồ họa. Bài toán tự động dán nhãn được xác định là một lĩnh vực nghiên cứu quan trọng. Bài toán này có ứng dụng trong nhiều lĩnh vực bao gồm vẽ bản đồ, hệ thống thông tin địa lý và vẽ đồ thị... Hiện trên thế giới đã có nhiều công trình nghiên cứu về bài toán dán nhãn tự động. Tuy nhiên ở Việt Nam, bài toán này còn được đề cập đến một cách hạn chế.

Từ sự định hướng của cán bộ hướng dẫn, căn cứ vào sự phát triển và những ứng dụng của bài toán này, tôi đã quyết định lựa chọn đề tài: “ Nghiên

cứ kỹ thuật dán nhãn cho đối tượng 2D”, bởi đây là một đề tài có mang tính thực tiễn cao và đồng thời mở ra nhiều hướng phát triển nghiên cứu.

## **2. Lịch sử nghiên cứu**

Những tài liệu đầu tiên được công bố liên quan đến bài toán dán nhãn là những bài toán dán nhãn cho bản đồ địa lí, do bài toán khi đặc tả bản đồ có hai yếu tố cần phải đặc tả là yếu tố hình và chữ viết. Một trong những công bố đầu tiên liên quan đến bài toán dán nhãn là của Eduard Imhof[7] vào năm 1962, sau này bài toán dán nhãn tự động được xây dựng bởi một số nhà khoa học như Pinhas Yoeli[5]... Vào những thập niên 80 lượng nghiên cứu đến bài toán dán nhãn tự động có tăng thể hiện qua nhiều bài báo được công bố, nhiều nhà khoa học trong trong lĩnh vực khác nhau cũng quan tâm đến chủ đề này và nó thuộc lớp bài toán tối ưu tổ hợp có độ phức tạp tính toán cao.

## **3. Mục đích đối tượng và phạm vi nghiên cứu**

Bài toán dán nhãn tự động cho một đối tượng cụ thể như bản đồ, đối tượng 2D, đối tượng 3D... đã được nghiên cứu và công bố trên thế giới. Tuy nhiên ở Việt Nam chưa quan tâm đến bài toán này một cách đúng mức. Vì vậy đề tài được xây dựng với các mục đích: Tìm hiểu về bài toán dán nhãn cho đối tượng, thực tiễn ứng dụng của nó trong các lĩnh vực, tìm hiểu một số vấn đề về đồ họa 2D, các thuật toán liên quan từ đó đặt nền tảng nghiên cứu sâu hơn về vấn đề này từ đó thiết kế chương trình thử nghiệm dán nhãn cho một số đối tượng 2D.

Đối tượng và phạm vi nghiên cứu của đề tài là tập chung vào đối tượng 2D và các thuật toán dán nhãn tự động.