

ĐỀ XUẤT ĐỀ TÀI LUẬN VĂN THẠC SĨ

Họ và tên người hướng dẫn 1: **DƯƠNG ANH ĐỨC**

Cơ quan công tác: **Trường ĐH Công Nghệ Thông Tin – ĐHQG HCM**

Họ và tên người hướng dẫn 2: **LÊ ĐÌNH DUY**

Cơ quan công tác: **Trường ĐH Công Nghệ Thông Tin – ĐHQG HCM**

Đề tài luận văn đối với học viên cao học ngành: **Khoa học Máy tính**

- (1). Tên hướng nghiên cứu: **Thị giác Máy tính**
- (2). Tên đề tài luận văn thạc sĩ: **Tái nhận dạng nhân vật trên mạng các camera**
- (3). Mục tiêu, đối tượng và phạm vi nghiên cứu của đề tài:

Mục tiêu:

- Nghiên cứu các kỹ thuật hiện nay cho bài toán tái nhận dạng nhân vật.
- Đưa ra một số đề xuất giúp cải thiện độ chính xác của quá trình tái nhận dạng. Cụ thể, đề tài sẽ tập trung vào cải tiến các kỹ thuật rút trích đặc trưng thị giác và quá trình hậu xử lý kết quả tái nhận dạng nhằm tăng độ chính xác.
- Xây dựng chương trình thử nghiệm.

Đối tượng và phạm vi nghiên cứu:

- Dữ liệu video được quay từ các camera quan sát.

- (4). Thời gian nghiên cứu và số tín chỉ của đề tài:

Thời gian nghiên cứu: **6 tháng**

Số tín chỉ: **15 tín chỉ**

- (5). Yêu cầu về kiến thức, kỹ năng và có thể gồm tên của HV thực hiện đề tài:

TP. HCM, ngày tháng năm 20

NGƯỜI HƯỚNG DẪN 1

NGƯỜI HƯỚNG DẪN 2

Dương Anh Đức

Lê Đình Duy

ĐỀ CƯƠNG ĐỀ TÀI LUẬN VĂN THẠC SĨ

1. Tên đề tài hoặc hướng NC (gồm cả tiếng Việt và tiếng Anh):

Tên tiếng Việt: *Tái nhận dạng nhân vật trên mạng các camera*

Tên tiếng Anh: *Person Re-Identification on Networks of Cameras*

2. Ngành và mã ngành đào tạo: Ngành Khoa học máy tính, mã ngành: 60.48.01.01

3. Họ tên học viên thực hiện đề tài, khóa-đợt học: Nguyễn Hoàng Vũ

Địa chỉ email, điện thoại liên lạc: hoangvu90@gmail.com

Người hướng dẫn 1: PSG. TS. Dương Anh Đức

Địa chỉ email, điện thoại liên lạc: ducda@uit.edu.vn

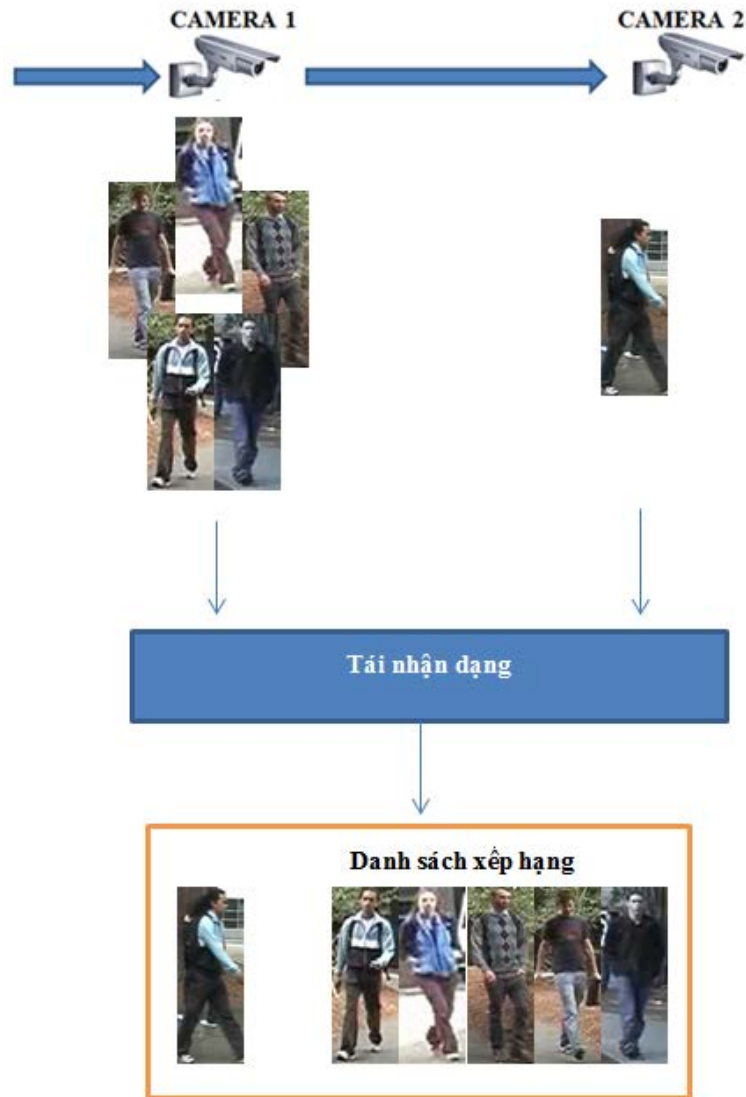
Người hướng dẫn 2: PSG. TS. Lê Đình Duy

Địa chỉ email, điện thoại liên lạc: ldduy@uit.edu.vn

4. Tổng quan tình hình NC trong nước và ngoài nước: mỗi nội dung đề cập cần có trích dẫn tài liệu, công trình; cần phân tích, đánh giá những kết quả NC mới có liên quan đến đề tài hoặc hướng NC; cần chỉ ra những gì còn tồn tại:

4.1 Giới thiệu chung

- Ngày nay, hệ thống camera quan sát đã trở nên rất phổ biến trong các môi trường cần đảm bảo an ninh như siêu thị, ngân hàng, thư viện, trường học,... Việc vận hành, sử dụng hệ thống camera quan sát hiện nay vẫn được thực hiện thủ công chủ yếu bởi các chuyên viên vận hành.. Tuy nhiên, tại những địa điểm như trên, số lượng người xuất hiện thường rất lớn, gây rất nhiều khó khăn cho công việc giám sát của con người.
- Bài toán tái nhận dạng nhân vật (Person Re-Identification) bằng hệ thống máy tính được đưa ra nhằm giải quyết vấn đề này. Tái nhận dạng nhân vật là bài toán so khớp hình ảnh các nhân vật được quay từ hai camera khác nhau có vùng quan sát không chồng lấp. Cụ thể, với mỗi người xuất hiện ở camera thứ hai, hệ thống sẽ output một danh sách những nhân vật đã xuất hiện ở camera thứ nhất được xếp hạng theo mức độ tương tự với người đang xét.



Hình 1: Ví dụ về bài toán tái nhận dạng nhân vật

4.2. Các thách thức

- Đầu tiên phải kể đến là vì các camera quan sát một phạm vi rất lớn, do đó, chất lượng hình ảnh với một số nhân vật rất thấp. Vì vậy, các phương pháp nhận dạng truyền thống như nhận dạng khuôn mặt hay các phương pháp sinh trắc học không thể áp dụng cho bài toán này.
- Các camera được bố trí tại những vị trí khác nhau do đó dẫn đến góc nhìn cũng khác nhau khi thu hình các nhân vật, làm cho các đặc điểm thị giác của các nhân vật thay đổi khi di chuyển từ camera này sang camera khác. Bên cạnh đó điều kiện môi trường tại các camera khác nhau (ví dụ như ánh sáng) là khác nhau, do đó ảnh hưởng rất lớn đến hình ảnh thu được của các nhân vật.
- Để so khớp các hình ảnh thu được trên các camera khác nhau, cần phải tách phần ảnh nền của các ảnh. Các kỹ thuật tách nền hiện nay vẫn chưa cho kết quả đủ tốt, do đó, làm giảm hiệu quả của việc tái nhận dạng.

- Một số thách thức khác như các nhân vật bị che khuất,... cũng ảnh hưởng đến kết quả nhận dạng.

4.3. Tình hình nghiên cứu

Tình hình nghiên cứu ngoài nước:

Một hệ thống tái nhận dạng nhân vật gồm có hai thành phần là trích xuất đặc trưng thị giác và tính khoảng cách giữa các ảnh dựa vào các đặc trưng đã trích xuất. Do đó, các nghiên cứu hiện nay cũng tập trung tiếp cận trên hai vấn đề này.

- **Hướng tiếp cận trích xuất đặc trưng thị giác.**
 - **Đặc trưng cấp thấp.** Trong các loại đặc trưng cấp thấp, màu sắc được xem là một loại đặc trưng đơn giản nhưng rất hiệu quả được sử dụng rất rộng rãi trong các công trình [1, 6, 3]. Bên cạnh đó, đặc trưng hình dáng cũng được giới thiệu trong các công bố [3, 7]. Đặc trưng dựa trên sinh học (Biologically Inspired Feature – BIF), loại đặc trưng khá mới cũng nhận được sự quan tâm trong [6]. Đặc trưng texture cũng được sử dụng trong [8, 9]. Để tăng độ chính xác khi so sánh ảnh hai nhân vật với nhau, người ta thường chia ảnh nhân vật thành nhiều phần để so khớp từng phần. Trong [2], tác giả sử dụng các interest point để so khớp hai ảnh với nhau. Một số phương pháp như [1, 2, 3, 4] chia ảnh nhân vật thành những phần được định nghĩa trước để so khớp.
 - **Đặc trưng ngữ nghĩa:** Các đặc trưng cấp thấp thường rất nhạy cảm, dễ bị ảnh hưởng bởi nhiều yếu tố như ảnh nền, sự thay đổi cường độ sáng qua các camera khác nhau,... nên thời gian gần đây, các đặc trưng ngữ nghĩa được đã được giới thiệu và sử dụng rộng rãi [10, 11]. Đặc trưng ngữ nghĩa là những đặc trưng gắn liền với các khái niệm gần gũi với con người và thường được con người sử dụng để nhận dạng các đối tượng. Các đặc trưng ngữ nghĩa cho nhân vật có thể là *tóc ngắn, tóc dài, mang giày, quần short, áo thun,...* Các đặc trưng ngữ nghĩa được phát hiện dựa trên những đặc trưng cấp thấp và các bộ phát hiện các đặc trưng ngữ nghĩa có độ chính xác cao, giúp nâng cao độ chính xác của bài toán tái nhận dạng nhân vật.



Hình 3: Ví dụ đặc trưng ngữ nghĩa: (1) quần jean; (2) quần short; (3) tóc dài

- **Hướng tiếp cận máy học để tính khoảng cách giữa các ảnh.** Bên cạnh các nghiên cứu về đặc trưng thị giác, các mô hình máy học để tính khoảng cách giữa các ảnh cũng rất được quan tâm. Các công trình tập trung vào phát triển những kỹ thuật máy học nhằm tối ưu hóa các tham số của các mô hình tái nhận dạng hoặc học những đặc trưng hiệu quả từ dữ liệu huấn luyện [7, 8, 9].

Tình hình nghiên cứu trong nước:

- Tại Việt Nam, Nhóm nghiên cứu của TS. Lê Đình Duy và đồng sự tại PTN Truyền thông Đa Phương tiện, ĐH Công nghệ Thông tin có những nghiên cứu về tái nhận dạng nhân vật trong khuôn khổ của đề tài ĐHQG-C về phát hiện trộm dùng mạng cảm biến camera. Trong đề tài này, có một bài toán nhỏ là sử dụng hướng tiếp cận máy học cho việc phát hiện các thành phần cơ thể người (body-parts) đã được nghiên cứu và cài đặt để tăng độ chính xác.
- Bên cạnh đó, nhóm của Tiến sĩ Trương Công Dung Nghi, ĐH Bách Khoa TP Hồ Chí Minh cũng có những nghiên cứu về tái nhận dạng nhân vật trong một đề tài được tài trợ bởi NAFOSTED năm 2013.

Những NC có trước của học viên liên quan đến đề tài luận văn:

- Học viên có tham gia nhóm nghiên cứu về phương pháp xác định các thành phần cơ thể người dựa trên hướng tiếp cận máy học sử dụng mô hình DPM (Deformable Part Models [15]) [R1].
- Học viên đã tham gia nhóm nghiên cứu về đề xuất kỹ thuật hậu xử lý kết quả tái nhận dạng để có được kết quả nhận dạng chính xác hơn [R2].

5. Tính khoa học và tính mới của đề tài: *những hạn chế hiện thời, những vấn đề khoa học cần NC giải quyết và tính cấp thiết của nó, hướng giải quyết có gì mới, lợi ích khoa học và thực tiễn khi vấn đề được giải quyết:*

Đề tài sẽ tập trung vào hai bài toán nhỏ của tái nhận dạng nhân vật là trích xuất đặc trưng thị giác và hậu xử lý kết quả tái nhận dạng

Trích xuất đặc trưng thị giác

- Như đã đề cập trong phần tình hình nghiên cứu, để nâng cao độ chính xác, người ta thường chia cơ thể người thành nhiều phần để so khớp. Các phương pháp phát hiện các thành phần cơ thể người hiện nay thường là cố định (chỉ làm việc với một cấu trúc cơ thể người nhất định). Đề tài sẽ nghiên cứu, **đề xuất phương pháp xác định các thành phần cơ thể người tổng quát hơn**, cho phép định nghĩa những cấu trúc cơ thể người khác nhau tùy ý.
- Vì đặc thù của bài toán tái nhận dạng là việc thay đổi góc nhìn, tư thế đối với nhân vật qua nhiều camera, do đó các đặc trưng ngữ nghĩa được chọn phải bất biến với những sự

thay đổi này. Bên cạnh đó, phương pháp chọn các đặc trưng ngữ nghĩa (tự động, bán tự động, thủ công) cũng là một thách thức (các phương pháp hiện nay chọn thủ công theo chủ quan của các tác giả). Đề tài sẽ tập trung vào **đề xuất phương pháp xác định một tập các đặc trưng ngữ nghĩa** phù hợp với bài toán tái nhận dạng nhân vật.

- Để có thể xây dựng các bộ phát hiện những đặc trưng ngữ nghĩa, cần phải có dữ liệu huấn luyện. Đa số các công trình hiện nay sử dụng chính một phần các bộ dữ liệu cho bài toán tái nhận dạng nhân vật. Hướng tiếp cận này có hai hạn chế. Thứ nhất, số lượng mẫu huấn luyện cho các đặc trưng ngữ nghĩa không nhiều. Thứ hai, cần có người đánh nhãn các bộ dữ liệu này một cách thủ công. Trong khi đó, có thể thấy lượng dữ liệu ảnh có sẵn trên Internet cùng với các thông tin đi kèm vô cùng phong phú. Do đó, trong đề tài này, chúng tôi đề xuất **nghiên cứu phương pháp sử dụng đặc trưng ngữ nghĩa cho bài toán tái nhận dạng nhân vật bằng cách khai thác dữ liệu ảnh trên các trang Internet**, cụ thể là các trang web và các kho lưu trữ ảnh và video như Flickr, Youtube,...
- Việc khai thác dữ liệu ảnh trên Internet giúp xây dựng các bộ dữ liệu huấn luyện phong phú, với số lượng mẫu lớn mà không cần đánh nhãn thủ công từ con người. Bên cạnh bài toán tái nhận dạng, kỹ thuật này còn có thể được áp dụng trong nhiều bài toán khác như tìm kiếm ảnh, tìm kiếm video, nhận dạng vật thể,...

Hậu xử lý kết quả nhận dạng

- Các phương pháp tái nhận dạng hiện nay đối xử với các nhân vật xuất hiện ở camera thứ hai một cách độc lập nhau. Tuy nhiên, trong thực tế có thể có nhiều nhân vật xuất hiện cùng một lúc. Đề tài sẽ nghiên cứu và đề xuất **phương pháp tận dụng thông tin xuất hiện đồng thời của nhiều nhân vật** ở camera thứ hai nhằm cải thiện độ chính xác.

6. Mục tiêu, đối tượng và phạm vi NC đề tài *cần hướng tới và khả năng giải quyết: Nêu cụ thể, súc tích theo từng mục; không viết chung chung, lý giải dài dòng,...*:

Mục tiêu:

- Tìm hiểu, phân tích các kỹ thuật tái nhận dạng nhân vật
- Đưa ra các đề xuất nhằm cải thiện độ chính xác của kỹ thuật tái nhận dạng nhân vật. Cụ thể:
 - Phát triển các phương pháp rút trích đặc trưng thị giác cho các ảnh nhân vật.
 - Đề xuất phương pháp hậu xử lý nhằm cải thiện độ chính xác tái nhận dạng.
- Xây dựng chương trình thử nghiệm

Đối tượng và phạm vi nghiên cứu

- Dữ liệu video thu được từ các camera quan sát.

7. Nội dung, phương pháp dự định NC. *Trình bày theo từng bước, mỗi bước cần có: mục tiêu dự kiến kết quả⁺, phương pháp, bố trí NC cụ thể (case-study), phương án phối hợp với các đơn vị, phòng thí nghiệm,... kết quả dự kiến đạt được và sẽ công bố:*

Nội dung 1:

Mục tiêu:

- Khảo sát tổng quan về bài toán tái nhận dạng nhân vật, các phương pháp kỹ thuật hiện nay, các bộ dữ liệu chuẩn được sử dụng cho bài toán.

Kết quả

- Tài liệu tổng hợp, phân tích các phương pháp state-of-the-art hiện nay về tái nhận dạng nhân vật
- Các bộ dữ liệu chuẩn cho bài toán tái nhận dạng.
- Mã nguồn của những phương pháp state-of-the-art (nếu có)

Nội dung 2:

Mục tiêu:

- Đề xuất phương pháp xác định các thành phần cơ thể người một cách tổng quát cho phép định nghĩa những cấu trúc cơ thể người khác nhau tùy ý.

Kết quả

- Phương pháp mới cho phép xác định các thành phần cơ thể người một cách tổng quát với kết quả tái nhận dạng cao hơn các phương pháp hiện nay.
- Một bài báo hội nghị.

Nội dung 3:

Mục tiêu

- Đề xuất phương pháp xác định tập các đặc trưng ngữ nghĩa
- Nghiên cứu phát triển kỹ thuật khai phá dữ liệu ảnh từ internet để xây dựng dữ liệu huấn luyện cho đặc trưng ngữ nghĩa.

Kết quả

- Phương pháp xác định tập các đặc trưng ngữ nghĩa
- Phương pháp khai phá dữ liệu ảnh và video từ internet phục vụ đặc trưng ngữ nghĩa.
- Phương pháp tái nhận dạng nhân vật mới sử dụng các đặc trưng ngữ nghĩa được xây dựng từ những dữ liệu được khai phá từ web, Youtube,... với độ chính xác cao hơn các phương pháp hiện tại.
- Một bài báo hội nghị.

Nội dung 4:

Mục tiêu

- Đề xuất phương pháp hậu xử lý sử dụng thông tin đồng xuất hiện của các nhân vật nhằm cải thiện độ chính xác quá trình tái nhận dạng.

Kết quả

- Phương pháp hậu xử lý sử dụng thông tin đồng xuất hiện của các nhân vật.
- Một bài báo hội nghị

8. Kế hoạch bố trí thời gian NC phải kết hợp nhiệm vụ NC với học tập; làm xêmina, kế hoạch làm việc với người hướng dẫn, kế hoạch viết bản thảo luận văn,... :

Nội dung	Thời gian
Nội dung 1: Khảo sát tổng quan về bài toán tái nhận dạng nhân vật, các phương pháp kỹ thuật hiện nay, các bộ dữ liệu chuẩn được sử dụng cho bài toán.	3/2014
Nội dung 2: Đề xuất các phương pháp cải thiện độ chính xác của bài toán tái nhận dạng nhân vật	4/2014 – 7/2014
Nội dung 3: Viết luận văn	8/2014

9. Tài liệu tham khảo gồm những tài liệu có liên quan nhất với đề tài NC; ngoài những tài liệu kinh điển phải có những tài liệu cận đại (năm năm trở lại) liên quan đến đề tài. Viết theo mẫu⁺, và thuyết minh đề tài phải tham chiếu đến các tài liệu tham khảo này.

1. M. Farenzena, L. Bazzani, A. Perina, V. Murino, M. Cristani (2010), “*Person Re-Identification by symmetry-driven accumulation of local features*”. Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on , vol., no., pp.2360,2367, 13-18 June 2010
2. N. Gheissari, T. Sebastian, R. Hartley (2006), “*Person reidentification using spatial temporal appearance*”. Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on , vol.2, no., pp.1528,1535, 2006
3. O. Oreifej, R. Mehran, M. Shah (2010), “*Human identity recognition in aerial images*”. 2010 IEEE Conference on , vol., no., pp.709,716, 13-18 June 2010
4. D. S. Cheng, M. Cristani, M. Stoppa, L. Bazzani, V. Murino (2011), “*Custom pictorial structures for re-identification*”. Proceedings of the British Machine Vision Conference, pages 68.1-68.11. BMVA Press, September 2011
5. P. E. Forssen (2007), “*Maximally stable colour regions for recognition and Matching*”. Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on , vol., no., pp.1,8, 17-22 June 2007
6. B. Ma, Y. Su, F. Jurie (2012), “*Bicov: a novel image representation for person re-identification and face verification*”. In Proceedings British Machine Vision Conference 2012. Pages 57.1--57.11
7. W. Schwartz, L. Davis (2009), “*Learning Discriminative Appearance-Based Models Using Partial Least Squares*”. Brazillian Symposium on Computer Graphics and Image Processing 2009
8. D. Gray, H. Tao(2008), “*Viewpoint invariant pedestrian recognition with an ensemble of localized features*”. Proc. European Conference on Computer Vision, 2008
9. B. Prosser, W. S. Zheng, S. Gong, T. Xiang (2010), “*Person re-identification by support vector ranking*”. Proceedings of the British Machine Vision Conference, pages 21.1-21.11. BMVA Press, September 2010
10. R. Layne, T. Hospedales, S, Gong (2012), “*Person re-identification by attributes*”. Proceedings of the British Machine Vision Conference 2012

11. C. Liu, S. Gong, C.C. Loy, X. Lin (2012), “*Person re-identification: What features are important?*”. In: Proceedings of the First International Workshop on Re-Identification (2012)
12. P. F. Felzenszwalb, R. B. Girshick, D. McAllester, D. Ramanan (2010), “*Object detection with discriminatively trained part based models*”. IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI) 2010
- R1. Vu-Hoang Nguyen, Kien Nguyen, Duy-Dinh Le, Duc Anh Duong, Shin'ichi Satoh (2013). “*Person Re-identification Using Deformable Part Models*”. The 20th International Conference on Neural Information Processing, ICONIP 2013
- R2. Vu-Hoang Nguyen, Thanh Duc Ngo, Kien Nguyen, Khang M. T. T. Nguyen, Duy-Dinh Le, Duc Anh Duong (2013), “*Re-Ranking for Person Re-Identification*”. The 5th International Conference of Soft Computing and Pattern Recognition, SoCPaR 2013

NGƯỜI HƯỚNG DẪN 1
(Họ tên và chữ ký)

Dương Anh Đức

TP. HCM, ngày tháng năm 201..
NGƯỜI HƯỚNG DẪN 2
(Họ tên và chữ ký)

Lê Đình Duy

HỌC VIÊN KÝ TÊN
(Họ tên và chữ ký)

Nguyễn Hoàng Vũ

Một số lưu ý: Văn phong nên diễn đạt cô đọng, hình ảnh, sơ đồ, biểu bảng,.. rõ ràng; phân tích ý nghĩa nghiên cứu đề tài và về kết quả dự kiến một cách thuyết phục. *Dự kiến kết quả*, có thể là: Đề xuất quan điểm; Đưa ra báo cáo hiện trạng; Đề xuất các biện pháp về tổ chức, quản lý; Tạo lập được mô hình, sản phẩm, công nghệ mẫu,v.v.. *Tài liệu tham khảo:* viết theo dạng mẫu, là sách hay bài báo, ví dụ:

- + với sách tham khảo:
- [1]. Nguyen P. K. (2005), Flood Simulation in the Vietnam Mekong Delta. *Proceedings of the 3rd International Symposium on Flood Defense*. Taylor & Francis Publ., pp.611-616. Netherlands.
+ với bài báo khoa học:
 - [2]. Hai Bang T., Ngoc Thanh N., Phi Khu N. (2011), “*Fuzzy Ontology Building and Integration for Fuzzy Inference Systems in Weather Forecast Domain*”. LNAI 6592, the 3rd Intl’ Conference, ACIIDS 2011, Korea. Springer, ISSN 0302-9743, ISBN 978-3-642-20042-7.