Tạp chí Khoa học Đại học Huế: Kỹ thuật và Công nghệ; ISSN 2588-1175

Tập 126, Số 2A, 2017, Tr. 153-161; DOI: 10.26459/hueuni-jtt.v126i2A.4444

TÌM HIỂU VÀ ĐÁNH GIÁ VỀ CÁC HƯỚNG TIẾP CẬN NGỮ NGHĨA TRONG NÂNG CAO HIỆU QUẢ TÌM KIẾM ẢNH

Nguyễn Thị Uyên Nhi[1]*, Lê Mạnh Thạnh[1], Hoàng Hữu Hạnh[2]

[1] Trường Đại học Khoa học, Đại học Huế
[2] Phòng nghiên cứu Web ngữ nghĩa, Đại học Huế

Tóm tắt: Dữ liệu đa phương tiện (chẳng hạn như hình ảnh, âm thanh, video) đóng một vai trò quan trọng trong tất cả lĩnh vực. Trong đó, dữ liệu ảnh với sự phát triển mạnh mẽ, trở nên quen thuộc với người sử dụng hơn vì tính biểu cảm trực quan của nó, được ứng dụng rộng rãi vào rất nhiều các hệ thống. Trong thập kỷ qua, rất nhiều nghiên cứu về tìm kiếm dữ liệu ảnh thông qua tiếp cận ngữ nghĩa. Các công nghệ ngữ nghĩa cung cấp cách tiếp cận đầy hứa hẹn cho việc tìm kiếm ảnh, giúp việc tìm kiếm hiệu quả và chính xác hơn. Bài báo này khảo sát, đánh giá và thảo luận về các công trình nghiên cứu tìm kiếm ảnh theo tiếp cận ngữ nghĩa. Bài báo cũng trình bày những ưu, nhược điểm của các công trình nghiên cứu đó để so sánh, đánh giá, từ đó đề xuất định hướng nghiên cứu trong tương lai.

Từ khóa: Ngữ nghĩa, ontology, tìm kiếm ảnh, tiếp cận ngữ nghĩa

1 Mở đầu

Ngày nay, cùng với sự bùng nổ của công nghệ thông tin (CNTT), dữ liệu đa phương tiện (chẳng hạn như hình ảnh, âm thanh, video) đóng một vai trò quan trọng trong tất cả lĩnh vực. Trong đó, dữ liệu ảnh với sự phát triển mạnh mẽ, trở nên quen thuộc với người sử dụng hơn vì tính biểu cảm trực quan của nó, được ứng dụng rộng rãi vào rất nhiều các hệ thống như: thư viện số, hệ thống thông tin địa lý, các nghiên cứu thiên văn học, hệ thống quan sát vệ tinh, hệ thống điều tra hình sự, ứng dụng tìm kiếm ảnh trong y học, giáo dục đào tạo, giải trí, mạng xã hội… Hình ảnh là nguồn nội dung chính trên Internet và đang gia tăng nhanh chóng do sự tiến bộ trong lĩnh vực này. Ảnh số đang được chia sẻ với một lượng dữ liệu khổng lồ và được tạo ra bởi các thiết bị khác nhau như camera, mobile, smartphone, tablet…Theo thống kê của IDC năm 2016 [1], thế giới đã tạo ra 1.138 nghìn tỷ hình ảnh (gấp hơn 700 lần so với năm 2015). Theo báo cáo về chia sẻ ảnh trên toàn cầu, Brandwatch đã tính toán rằng mỗi ngày có 350 triệu hình ảnh được chia sẻ qua Facebook, 95 triệu hình ảnh được chia sẻ qua Instagram, 400 triệu trên Snapchat và 1,6 tỷ hình ảnh trên WhatsApp.

Chính vì thế, tìm kiếm dữ liệu ảnh là một trong những lĩnh vực nghiên cứu thời sự nhất hiện nay. Các nhà khoa học trong thập kỷ qua đã đề xuất rất nhiều phương pháp tiếp cận hiệu quả cho tìm kiếm hình ảnh. Tìm kiếm hình ảnh thông qua ba loại kỹ thuật chính [2] bao gồm tìm kiếm dựa trên văn bản, tìm kiếm dựa trên nội dung và tìm kiếm ngữ nghĩa. Ngày này, tìm kiếm ảnh dựa trên web đang ngày một gia tăng. Tuy nhiên việc tìm kiếm ảnh trong "World

** Liên hệ: ntunhi@sgu.edu.vn*

Nhận bài: 21-8-2017; Hoàn thành phản biện: 4-9-2017; Ngày nhận đăng: 10-9-2017

Wide Web" vẫn chưa đạt được hiệu quả cao. Vì thế cần đặt ra phương pháp giúp người dùng cuối có thể lấy dữ liệu ảnh hiệu quả dựa trên kiến thức chuyên sâu về việc trình bày nội dung ảnh với cấu trúc được xác định rõ ràng.

Trong Semantic Web, dữ liệu ảnh có ý nghĩa rất rõ ràng [3]. Nó cho phép mọi người tạo ra các kho dữ liệu trên web, xây dựng các từ vựng và viết các quy tắc để xử lý dữ liệu. Semantic Web là sự mở rộng của web hiện đại mà trong đó thông tin được định nghĩa rõ rang sao cho con người và máy tính có thể cùng làm việc với nhau một cách hiệu quả hơn. Nó thêm metadata vào các tài liệu hiện có để mở rộng các tài liệu đó thành các thông tin được xác định rõ ràng. Phần mở rộng này cho phép web tự động được xử lý bằng máy móc và được sử dụng bởi con người.

Ontology là một công nghệ chính của web ngữ nghĩa, đại diện cho kiến thức trong web ngữ nghĩa. Ontology được định nghĩa là "một đặc tả rõ ràng chính thức của khái niệm chung". Nó bao gồm các khái niệm, mối quan hệ giữa các khái niệm, các thuộc tính của các khái niệm và các thể hiện. Một số ngôn ngữ ontology là RDF, OWL, SPARQL và SKOS.

Với việc tìm kiếm ảnh theo tiếp cận ngữ nghĩa, người dùng có thể thu thập ảnh theo nội dung hiệu quả cho web ngữ nghĩa. Tuy nhiên, vẫn cần thêm cơ chế để giảm khoảng cách ngữ nghĩa với nội dung của các đối tượng. Đối với các đối tượng trực quan, các tính năng cấp thấp như kích thước, màu sắc, kết cấu,... có thể được sử dụng để tìm các đối tượng thích hợp dựa trên yêu cầu của người dùng. Ngoài ra, các kỹ thuật khác nhau được sử dụng để lấy nội dung dữ liệu ảnh. Bài báo này tìm hiểu, đánh giá và so sánh các công trình nghiên cứu tìm kiếm ảnh theo tiếp cận ngữ nghĩa. Bài báo cũng trình bày những ưu, nhược điểm của các công trình nghiên cứu đó để so sánh, đánh giá, từ đó đề xuất định hướng nghiên cứu trong tương lai.

Phần 2 tìm hiểu, đánh giá các hệ thống tìm kiếm hình ảnh theo tiếp cận ngữ nghĩa trong các công trình nghiên cứu liên quan đã được công bố. Phần 3 so sánh kết quả của các hệ thống tìm kiếm hình ảnh ngữ nghĩa khác nhau để đề xuất hướng nghiên cứu. Cuối cùng, phần 4 kết luận công việc với các cải tiến trong tương lai.

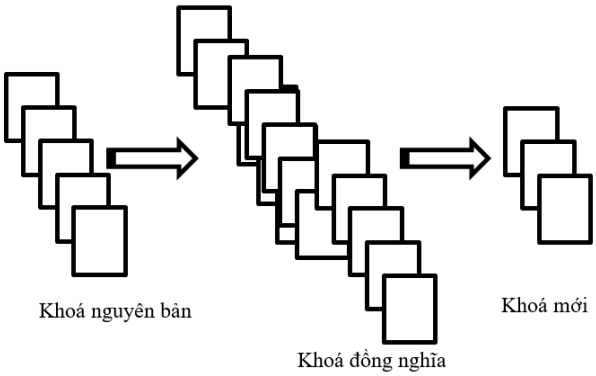
2 Các hệ thống tìm kiếm hình ảnh theo tiếp cận ngữ nghĩa

Tìm kiếm hình ản thông qua ba loại kỹ thuật chính bao gồm tìm kiếm dựa trên văn bản, tìm kiếm dựa trên nội dung và tìm kiếm hình ảnh theo tiếp cận ngữ nghĩa. Hầu hết hệ thống tìm kiếm hình ảnh dựa trên văn bản đều yêu cầu các hình ảnh được chú thích bằng tay [4]. Trong tìm kiếm dựa trên văn bản, các hình ảnh được tìm kiếm thông qua việc kết hợp các truy vấn của người dùng và chú thích từ khoá. Hệ thống tìm kiếm dựa trên từ khoá khớp với tìm kiếm văn bản của người dùng để mô tả văn bản của hình ảnh và trả lại tất cả các hình ảnh có mô tả phù hợp. Tuy nhiên, có thể các kết quả trả lại có chứa hình ảnh không liên quan. Hệ thống truy xuất hình ảnh dựa trên nội dung sử dụng các tính năng trực quan cấp thấp như màu sắc, kết cấu, hình dạng và vị trí… để lấy lại hình ảnh [5,6]. Các tính năng cấp thấp này được trích xuất từ hình ảnh tự động. Các nhà nghiên cứu trong thập kỷ qua đã chứng minh tính hiệu quả và tính chính xác của các kỹ thuật tìm kiếm dựa trên nội dung. Tuy nhiên, kỹ thuật tìm

kiếm dựa trên nội dung vẫn thiếu sự hiểu biết về phân tích ngữ nghĩa của hình ảnh. Kỹ thuật thu hồi hình ảnh dựa trên ngữ nghĩa là sự kết hợp của các tính năng cấp thấp và ngữ nghĩa cấp cao của hình ảnh.

Ban đầu, nhóm nghiên cứu Mezaris, Vasileios, Ioannis Kompatsiaris, và Michael G. Strintzis [7] đề xuất sử dụng một thuật toán phân đoạn hoàn toàn không được giám sát để phân chia các hình ảnh thành các vùng và mang lại hệ thống lập chỉ mục, tìm kiếm với các chức năng dựa trên nội dung. Mô tả mức độ thấp cho màu sắc, vị trí, kích thước và hình dạng của từng khu vực được trích xuất theo thời gian. Các bộ mô tả số học này được tự động kết hợp với các bộ mô tả cấp trung cấp thích hợp, tạo thành một từ vựng đơn giản gọi là đối tượng ontology. Ontology được sử dụng để cho phép định nghĩa định tính các khái niệm cấp cao mà người dùng tìm kiếm (các đối tượng ngữ nghĩa, mỗi từ được đại diện bởi một từ khóa) và các mối quan hệ của chúng. Khi tìm kiếm một đối tượng ngữ nghĩa cụ thể (hoặc các đối tượng), các giá trị mô tả mức trung gian liên kết với cả đối tượng ngữ nghĩa và tất cả các vùng hình ảnh trong bộ sưu tập ban đầu được so sánh, dẫn đến việc từ chối hầu hết các vùng hình ảnh là không liên quan.

Nhóm Yang, Ming Dong và Farshad Fotouhi [8] đề xuất một mô hình xử lý ngôn ngữ tự nhiên để thu thập hình ảnh, kết hợp các thông tin ngữ nghĩa được cung cấp bởi WordNet, một hệ thống tham chiếu trực tuyến và các tính năng trực quan cấp thấp. Hệ thống thứ bậc ngữ nghĩa của từ nghĩa từ WordNet được sử dụng để tăng cường mối liên hệ giữa các hình ảnh và mô tả nguyên bản của một khái niệm. Thuật toán chọn từ khoá thống kê được thực hiện giúp chọn các từ khoá đại diện nhất để chú thích những hình ảnh của khái niệm này.



Hình 1. Quá trình lựa chọn từ khoá

Kết quả thực nghiệm trên bộ ảnh Corel CD với 2500 hình ảnh phong cảnh, cho thấy phương pháp tiếp cận này có thể cải thiện đáng kể độ chính xác của tìm kiếm. Các kết quả cũng cho thấy tiềm năng cao của cách tiếp cận trong việc cung cấp cơ sở ngữ nghĩa để xây dựng ontology cho các dữ liệu hình ảnh.

Nhóm nghiên cứu Wang Huan, Song Liu, and Liang-Tien Chia [9] trình bày sự so sánh giữa tìm kiếm hình ảnh dựa trên từ khóa và dựa trên ontology. Wang Huan chỉ ra rằng cách

tiếp cận dựa trên từ khoá thân thiện với người sử dụng và dễ áp dụng với độ chính xác về tìm kiếm được chấp nhận, trong khi ontology về ngữ nghĩa phong phú, giải quyết nhu cầu mô tả hoàn chỉnh việc tìm kiếm hình ảnh và cải thiện độ chính xác của việc tìm kiếm. Ontology hoạt động tốt hơn với sự kết hợp của các tính năng hình ảnh. Mặc dù có sự cân bằng giữa sự phức tạp và hiệu năng, ontology vẫn là một lựa chọn khả thi khi hiệu suất tốt hơn mong đợi.

Với việc sử dụng ontology đa phương thức, Wang Huan, Liang-Tien Chia và Song Liu đã phát triển một hệ thống truy xuất ngữ nghĩa với việc sử dụng ontology đa phương thức: ontology miền, mô tả văn bản ontology và mô tả trực quan ontology [10]. Nhóm nghiên cứu đã phát triển một cơ chế xếp hạng mới cho chú thích hình ảnh dựa trên ontology đa phương thức và thu hồi bằng cách sử dụng một mối tương quan xếp hạng, tính toán mức độ tương tự về ngữ nghĩa giữa các khái niệm để cải thiện kết quả được tạo ra và các khái niệm được xác định trước. Kết quả trong thử nghiệm đã chứng minh rằng ontology đa phương thức của có thể mở rộng. Tuy nhiên, tiêu chí xếp hạng của nhóm nghiên cứu chỉ liên quan đến bộ ảnh về chó, và một khu vực động vật lớn hơn vẫn đang được xây dựng. Cơ chế xếp hạng vẫn chưa mang lại hiệu quả cao trong tìm kiếm ảnh.

Việc sử dụng ontology để chú thích hình ảnh được nhóm nghiên cứu Koletsis, Pyrros, and Euripides Petrakis [11] đề xuất để chú thích hình ảnh tự động. Ontology được xây dựng mang đặc trưng từ nhiều nguồn thông tin bao gồm mô tả văn bản và các tính năng hình ảnh cấp thấp. Chú thích hình ảnh được thực hiện như một quá trình tìm kiếm bằng cách so sánh hình ảnh đầu vào (tìm kiếm) với hình ảnh đại diện của tất cả các lớp. Xếp hạng truy hồi trung bình (AVR) được áp dụng để tính toán khả năng hình ảnh đầu vào thuộc về mỗi một lớp ontology. Các kết quả của nhóm nghiên cứu chứng minh rằng gần 89% các hình ảnh thử nghiệm được chú thích chính xác. Tuy nhiên, nhóm nghiên cứu chỉ chú thích và tìm kiếm trong bộ ảnh động vật, chưa phân tích hình ảnh phức tạp như xử lý các vị trí khác nhau của đầu thú, thiếu các phân loại chuyên sâu cho hình ảnh.

Nhóm nghiên cứu Magesh, N., and P. Thangaraj [12] đề xuất phương pháp tiếp cận bằng chú thích hình ảnh với ontology để nâng cao khả năng tìm kiếm. Ontology được sử dụng cho chú thích hình ảnh ngữ nghĩa và tìm kiếm trong bộ sưu tập lớn các hình ảnh (2000 ảnh). Tạo ra ontology để tìm kiếm hình ảnh với mối quan hệ hoàn chỉnh, giúp nâng cao khả năng lấy hình ảnh với tốc độ nhanh hơn. Ontology được tạo ra để xác định không gian ngữ nghĩa. Câu ngôn ngữ tự nhiên được chuyển đổi thành câu lệnh SPARQL và các hình ảnh có liên quan được truy cập bằng truy vấn SPARQL. Ontology được trình bày dưới dạng RDF để chú thích các hình ảnh, nhiệm vụ tìm kiếm hình ảnh được đơn giản hóa hơn so với các phương pháp tiếp cận trước đó. Truy xuất được thực hiện bằng cách sử dụng từ khóa mô tả. Cơ sở tri thức được làm giàu bằng các ontology với việc bổ sung siêu dữ liệu metadata của hình ảnh, để xây dựng các câu trả lời có ý nghĩa hơn cho các truy vấn.

Trong một nghiên cứu trong năm 2015, cũng với tìm kiếm hình ảnh theo ngữ nghĩa, nhóm nghiêm cứu Manzoor Umar [13] đề xuất phương pháp tiếp cận dựa trên Ontology. Sử dụng ontology với miền cụ thể để thu thập hình ảnh có liên quan đến tìm kiếm của người dùng. Người dùng có thể đưa khái niệm/từ khoá dưới dạng nhập văn bản hoặc có thể nhập hình ảnh

của chính nó. Tìm kiếm hình ảnh ngữ nghĩa dựa trên phương pháp tiếp cận lai và sử dụng phương pháp tiếp cận dựa trên hình dạng, màu sắc và kết cấu để phân loại mục đích. Hệ thống được đề xuất được đào tạo về bộ dữ liệu "Động vật có vú" và xây dựng các khái niệm ontology cho 900 hình ảnh có chứa hình ảnh của 20 loài động vật có vú khác nhau. Kết quả cho thấy hiệu quả của hệ thống đề xuất.

Nhóm nghiên cứu Filali, Jalila, Hajer Baazaoui Zghal, and Jean Martinet [14] đề xuất một hệ thống tìm kiếm hình ảnh dựa trên từ vựng trực quan và ontology. Xây dựng từ vựng trực quan sử dụng các tính năng cấp thấp và xây dựng bản thể luận dựa trên các khái niệm được trích ra từ các chú thích hình ảnh. Hệ thống tìm kiếm hình ảnh này bao gồm hai giai đoạn chính (giai đoạn ngoại tuyến và giai đoạn trực tuyến). Giai đoạn ngoại tuyến, tương ứng với từ vựng trực quan và giai đoạn xây dựng ontology, bao gồm hai bước: (1) xây dựng vốn từ vựng trực quan và (2) xây dựng ontology. Giai đoạn trực tuyến, tương ứng với giai đoạn thu hồi hình ảnh, bao gồm hai bước: (1) xử lý hình ảnh truy vấn và (2) thu hồi hình ảnh. Các ontology được làm phong phú thêm bởi các khái niệm và mối quan hệ được trích ra từ tài nguyên từ vựng của BabelNet.

Một cách tiếp cận khác của tìm kiếm hình ảnh dựa vào ontology là đề xuất một hệ thống dựa trên đồ thị, kết hợp với tính năng ngữ nghĩa và thị giác của nhóm nghiên cứu Allani Olfa và cộng sự [15]. Nhóm đề xuất cho việc tự động xây dựng một Module ontology dựa trên một bước học tập từ văn bản nguyên bản và tài nguyên thuật ngữ. Sau đó tổ chức các tính năng thị giác trong một mô hình dựa trên đồ thị, trong đó module kết hợp đồ thị biểu diễn một thành phần duy nhất được gọi là "mẫu", từ đó xây dựng một đồ thị mẫu cho tìm kiếm hình ảnh.

Các nghiên cứu trên đều chỉ ra rằng, việc phân tích và tìm kiếm hình ảnh đang trở thành một lĩnh vực nghiên cứu rất nóng hổi và cấp thiết hiện nay. Những nghiên cứu về tìm kiếm ảnh theo ngữ nghĩa, mà cụ thể là tìm kiếm ảnh dựa vào ontology cho ra kết quả vô cùng khả thi khi hiệu suất tìm kiếm nhanh chóng và chính xác hơn mong đợi. Nghiên cứu tìm kiếm ảnh dựa vào ngữ nghĩa còn rất nhiều khía cạnh mà các nhà nghiên cứu chưa khám phá hết, cần đi sâu để tìm hiểu và thực nghiệm, từ đó đưa ra được các đánh giá, so sánh giữa các phương pháp với nhau. Đây là một định hướng nghiên cứu phù hợp với xu thế nghiên cứu chung của thế giới, mang tính cấp thiết cao và có khả năng ứng dụng hiệu quả trong thực tiễn.

3 So sánh các hệ thống tìm kiếm ảnh theo ngữ nghĩa

Tìm kiếm hình ảnh theo cách tiếp cận dựa trên nội dung với các tính năng ngữ nghĩa đang được sử dụng rộng rãi. Cách tiếp cận xử lý ngôn ngữ tự nhiên được tích hợp với các tính năng cấp thấp được sử dụng để tìm kiếm hình ảnh. Nó cung cấp cơ sở ngữ nghĩa cho việc tạo ra ontology của cơ sở dữ liệu hình ảnh và cải thiện độ chính xác của truy xuất. Mô hình lai kết hợp bản thể luận và mạng Bayesian nhằm nâng cao chất lượng phục hồi hình ảnh. Nhưng trong mô hình này, lý lẽ và thứ hạng của các khái niệm cần được cải thiện để có tỷ lệ thu hồi tốt hơn. Phương pháp đa phương thức cung cấp khả năng mở rộng của hệ thống truy xuất và cơ chế xếp hạng đã được cải thiện cho các khái niệm miền lớn. Cách tiếp cận dựa trên khu vực thu hẹp

khoảng cách ngữ nghĩa khi lấy lại hình ảnh và khắc phục những thiếu sót của việc kết hợp hình ảnh và hình ảnh. Chú thích hình ảnh tự động để nâng cao hiệu quả tìm kiếm. Ontology được sử dụng để chú thích các hình ảnh, truy xuất được thực hiện bằng cách sử dụng từ khóa mô tả. Cơ sở tri thức được làm giàu bằng các ontology với việc bổ sung siêu dữ liệu metadata của hình ảnh, để xây dựng các câu trả lời có ý nghĩa hơn cho các truy vấn.

Bảng 1. So sánh các kết quả tìm kiếm hình ảnh theo tiếp cận ngữ nghĩa

STT	Phương pháp tiếp cận	Nhận xét kết quả
1	Tìm kiếm hình ảnh theo khu vực theo Ontology và phản hồi tích cực	- Khắc phục những hạn chế của các phương pháp thông thường như sự không đòi hỏi phải chú thích bằng tay, áp dụng được cho các bộ sưu tập hình ảnh chung, không có khả năng cấu trúc cơ sở tri thức miền cụ thể. - Kết quả thực nghiệm cho thấy hiệu quả của cách tiếp cận về độ chính xác
2	Xử lý ngôn ngữ tự nhiên	- Từ khoá đã chọn được tăng cường ngữ nghĩa về hình ảnh và mang lại trọng lượng cao cho các từ khóa có liên quan ngữ nghĩa, do đó cải thiện tính chính xác của truy vấn. - Cung cấp cơ sở ngữ nghĩa để tạo ra ontology cho các cơ sở dữ liệu hình ảnh.
3	Ontology đa phương thức	- Cơ chế xếp hạng mới cho chú thích hình ảnh dựa trên ontology đa phương thức và thu hồi bằng cách sử dụng một mối tương quan xếp hạng. - Tính toán mức độ tương tự giữa các khái niệm được tạo ra và các khái niệm được xác định trước => giảm khoảng cách giữa nội dung và ngữ nghĩa - Hiệu quả với bộ dữ liệu ảnh nhỏ, cung cấp khả năng mở rộng hơn với ontology đa phương thức.
4	Chú thích ngữ nghĩa ảnh bằng ontology	- Xử lý sự không chắc chắn trong mô tả các lớp và kết hợp thông tin từ nhiều nguồn thông tin để biểu diễn các lớp ontology. - Cung cấp khả năng mở rộng để xử lý nhiều loại hình ảnh hơn và kết hợp phân tích hình ảnh phức tạp hơn
5	Tìm kiếm hình ảnh ngữ nghĩa dựa trên ontology và truy vấn SPARQL	- Tạo ra ontology để thu thập các hình ảnh có mối quan hệ hoàn chỉnh, giúp nâng cao khả năng tìm kiếm hình ảnh với tốc độ nhanh hơn so với việc thu hồi thông thường. - Sử dụng truy vấn SPARQL để lấy hình ảnh và đạt độ chính xác cao hơn so với các truy vấn trước đó. - Cung cấp các metadata để làm giàu ngữ nghĩa cho hình ảnh.
6	Phương pháp dựa trên ontology miền	- Xây dựng ontology miền cụ thể để thu thập hình ảnh có liên quan đến truy vấn của người dùng. - Hiệu quả với bộ sưu tập lớn. Kết quả thực nghiệm cho thấy hiệu quả của kỹ thuật đề xuất.
7	Tìm kiếm hình ảnh với ontology và từ vựng trực quan	- Xây dựng từ vựng trực quan và ontology dựa trên các chú thích hình ảnh để nâng cao độ chính xác tìm kiếm hình ảnh. - Các ontology được làm phong phú thêm bởi các khái niệm và mối quan hệ được trích ra từ tài nguyên từ vựng của BabelNet.
8	Biểu đồ các tính năng ngữ nghĩa của hình ảnh	- Tự động xây dựng, trên các lĩnh vực hình ảnh, một module ontology => tạo thành một biểu đồ thị giác của vùng hình ảnh cho mỗi từ vựng liên quan đến miền. - Cải tiến cơ chế xếp hạng hình ảnh, nhằm tăng hiệu quả tìm kiếm, nâng cao thông tin ngữ nghĩa và giảm khoảng cách ngữ nghĩa với nội dung hình ảnh. - Cung cấp khả năng mở rộng phạm vi tới bộ dữ liệu hình ảnh lớn hơn như ImageNet để đánh giá hệ thống và cải tiến thuật toán phân đoạn để so sánh với các thuật toán phân đoạn trước đó.

Với những cách tìm kiếm hình ảnh theo tiếp cận ngữ nghĩa trên, có thể nhận thấy, mỗi phương pháp được xử lý với những bộ ảnh khác nhau, không thể nói phương pháp nào là tối ưu hoàn toàn cho từng bộ dữ liệu ảnh, nhưng tất cả đều hướng đến mục đích nâng cao hơn nữa hiệu quả của tìm kiếm ảnh. Như ở phương pháp 1 "Tìm kiếm hình ảnh theo khu vực theo Ontology và phản hồi tích cực", nhóm nghiên cứu chú trọng đến việc tìm kiếm ảnh chính xác, thì phương pháp 2 "Xử lý ngôn ngữ tự nhiên" tiếp cận theo hướng tăng ngữ nghĩa cho hình ảnh, nhằm nâng cao trọng lượng của các từ khoá, giúp cho việc tìm kiếm hình ảnh đạt tính chính xác cao hơn. Phương pháp 3 "Ontology đa phương thức" sử dụng cơ chế xếp hạng nhằm giảm khoảng cách giữa nội dung và ngữ nghĩa, thì phương pháp 8 "Biểu đồ các tính năng ngữ nghĩa của hình ảnh" sẽ cải tiến cơ chế xếp hạng hình ảnh này, nhằm tăng hiệu quả tìm kiếm, nâng cao thông tin ngữ nghĩa và giảm khoảng cách hơn nữa giữa ngữ nghĩa với nội dung hình ảnh… Những nghiên cứu ngày nay vẫn đang tiếp tục kế thừa và phát triển từ những nghiên cứu trước đó để nâng cao hiệu quả cho tìm kiếm hình ảnh, cũng như mở rộng ngữ nghĩa cho các tập dữ liệu ảnh mới.

Từ những nghiên cứu được công bố, chúng tôi đề xuất mô hình tìm kiếm dữ liệu ảnh dựa trên nội dung theo tiếp cận ngữ nghĩa. Xây dựng và cải tiến giải thuật tìm kiếm ảnh cho mô hình đã đề xuất nhằm tăng hiệu quả tìm kiếm. Nghiên cứu các phương pháp làm giàu ngữ nghĩa cho tập dữ liệu ảnh:

- Bổ sung metadata cho dữ liệu
- Xây dựng các ontology miền
- Đề xuất ngữ cảnh tìm kiếm, mô tả metadata như con người, địa điểm, thời gian…

4 Kết luận

Bài báo này cung cấp các kỹ thuật khác nhau về tìm kiếm hình ảnh theo tiếp cận ngữ nghĩa, từ quá khứ cho đến hiện tại. Nó cho thấy các hướng tiếp cận ngữ nghĩa ngày càng được quan tâm và nghiên cứu nhiều hơn, nhằm nâng cao hiệu quả tìm kiếm ảnh. Khảo sát và đánh giá các công trình nghiên cứu đã công bố, chúng tôi nhận thấy vai trò của ontology trong việc phát triển tìm kiếm ảnh theo ngữ nghĩa. Ontology được coi là xương sống của web ngữ nghĩa và sẽ là kỹ thuật chính trong tìm kiếm ảnh theo tiếp cận ngữ nghĩa.

Trong nghiên cứu tương lai, chúng tôi đề xuất mô hình tìm kiếm dữ liệu ảnh dựa trên nội dung theo tiếp cận ngữ nghĩa, xây dựng và cải tiến giải thuật tìm kiếm ảnh cho mô hình đã đề xuất nhằm tăng hiệu quả tìm kiếm, nghiên cứu các phương pháp làm giàu ngữ nghĩa cho tập dữ liệu ảnh.

Tài liệu tham khảo

1. Photoindustrie-Verband e.V (2016), "Trend Report 2016", Germany.
2. Y. Alemu, J. Koh, M. Ikram, D.Kim (2009), "Image Retrieval in Multimedia Databases: A Survey", *2009 Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, IEEE.
3. J. Zhai, Y. Cao, Y. Chen (2008), "Semantic Information Retrieval Based on Fuzzy Ontology for Intelligent Transportation Systems," *IEEE*.
4. T. Chang and Y. Huang (2008), "An Ontology Oriented Region-Based Image Retrieval Strategy," *IEEE*.
5. D. C. Pedronette, J. Almeida (2014), R. S. Torres "A scalable re-ranking method for content-based image retrieval", *Information Sciences*, Volume 265, 1 May 2014, Pages 91-104
6. J. Nicolas, K. Todorov, C. Hudelot (2010), "Ontology matching for the semantic annotation of images." *Fuzzy Systems (FUZZ), 2010 IEEE International Conference*.
7. M. Vasileios, I. Kompatsiaris, M. G. Strintzis (2004), "Region-based image retrieval using an object ontology and relevance feedback." *Eurasip Journal on applied signal processing 2004*: 886-901.
8. Y. Changbo, M. Dong, F. Fotouhi (2004), "Learning the semantics in image retrieval-a natural language processing approach." *Computer Vision and Pattern Recognition Workshop, 2004. CVPRW'04. Conference on IEEE*.
9. W. Huan, S. Liu, L. Chia (2006), "Does ontology help in image retrieval?: a comparison between keyword, text ontology and multi-modality ontology approaches." *Proceedings of the 14th ACM international conference on Multimedia*. ACM,
10. W. Huan, S. Liu, L. Chia (2007), "Semantic retrieval with enhanced matchmaking and multi-modality ontology.", *Multimedia and Expo, 2007 IEEE International Conference on. IEEE*
11. K.Pyrros, E. Petrakis (2010), "SIA: Semantic Image Annotation using ontologies and image content analysis." *Image Analysis and Recognition: 374-383.*
12. N. Magesh, P. Thangaraj (2011), "Semantic image retrieval based on ontology and SPARQL query." *International Conference on Advanced Computer Technology (ICACT)*.
13. M. Umar, et al (2015), "Semantic image retrieval: An ontology based approach.", *International Journal of Advanced Research in Artificial Intelligence (IJARAI)* 1.4: 1-8.
14. F. Jalila, H. Baazaoui Zghal, J. Martinet (2016), "Towards Visual Vocabulary and Ontology-based Image Retrieval System." *ICAART (2)*.
15. A. Olfa, et al (2017), "Pattern graph-based image retrieval system combining semantic and visual features." *Multimedia Tools and Applications*: 1-30.

RESEARCH AND EVALUATION ON THE SEMANTIC APPROACHES IN IMPROVING IMAGE RETRIEVAL EFICIENCY

Nguyen Thi Phuong Nhi[1]*, Le Manh Thanh[1], Hoang Huu Hanh[2]

[1] HU – University of Sciences
[2] Hue University

Abstract: Multimedia data (such as images, audio, video) plays a crucial role in all areas. Along with it, the strong development of image data becomes more familiar to the user because of its visual expressiveness, is widely applied in many systems. In the past decade, a lot of research on image data search through semantic approach. Semantic technologies provide a promising approach to image search, making it more efficient and accurate. This article explores, reviews, and discusses image-search studies by semantic approaches. The paper also presents the advantages and disadvantages of these studies for comparison and evaluation, thereby proposing future research orientation.

Keywords: semantics, image search, semantic approach