Hybrid operations for content-based Vietnamese agricultural multimedia information retrieval

- **Pham Minh Nhut**
- **Pham Quang Hieu**
- **Luong Hieu Thi**
- **Vu Hai Quan**

University of Science, VNU-HCM

ABSTRACT

Content-based multimedia information retrieval is never a trivial task even with state-of-the-art approaches. Its mandatory challenge, called "semantic gap," requires much more understanding of the way human perceive things (i.e., visual and auditory information). Computer scientists have spent thousands of hours seeking optimal solutions, only ended up falling in the bound of this gap for both visual and spoken contexts. While an over-the-gap approach is unreachable, we insist on assembling current viable techniques from both contexts, aligned with a domain concept base (i.e., an ontology), to construct an info service for the retrieval of agricultural multimedia information. The development process spans over three packages: (1) building a Vietnamese agricultural thesaurus; (2) crafting a visual-auditory intertwined search engine; and (3) system deployment as an info service. We spring our the thesaurus in 2 sub-boughs: the aquaculture ontology consists of 3455 concepts and 5396 terms, with 28 relationships, covering about 2200 fish species and their related terms; and the plant production ontology comprises of 3437 concepts and 6874 terms, with 5 relationships, covering farming, plant production, pests, etc. These ontologies serve as a global linkage between keywords, visual, and spoken features, as well as providing the reinforcement for the system performances (e.g., through query expansion, knowledge indexing…). On the other hand, constructing a visual-auditory intertwined search engine is a bit trickier. Automatic transcriptions of audio channels are marked as the anchor points for the collection of visual features. These features, in turn, got clustered based on the referenced thesauri, and ultimately tracking out missing info induced by the speech recognizer's word error rates. This compensation technique bought us back 14 % of loss recall and an increase of 9 % accuracy over the baseline system. Finally, wrapping the retrieval system as an info service guarantees its practical deployment, asour target audiences are the majority of farmers in developing countries who are unable to reach modern farming information and knowledge.

Keywords: *semantic information retrieval, content-based video retrieval, agriculture, multimedia, Vietnamese, info service, agricultural ontology.*

INTRODUCTION

In Vietnam, agriculture plays an important part in the country's economic structure. In 2013, agriculture and forestry accounted for 18.4 percent of Vietnam's gross domestic product (GDP) [1]. As a result, information on agriculture comes out in large numbers and in different forms, from textual content to audio or videos. Farmers run into difficulties when searching for this kind of information, because of their lack of subject knowledge and most of the time novice users face insurmountable difficulty in formulating the right keyword queries [2], subsequently induces semantic mismatches between query intension and the fetched documents. Generic search engines such as Google or Bing can give decent results, but a carefully tailored search engine with specific domain knowledge and semantic retrieval techniques [6] can give a better performance. And hence it could bring out the possibilities for these novice seekers to be able to efficiently access to the vast multimedia resources available on the Web.

Multimedia resources, such as videos, are self-contained materials, which carry a large amount of rich information. Researches [3, 4, 5] have been conducted in the field of video retrieval amongst which semantic or content-based (as compared to text- or tag-based) retrieval of video is an emerging research topic [6]. Fig. 1 illustrates a full-fledged content-based video retrieval system, which typically combines text, spoken words, and imagery. Such system would allow the retrieval of relevant clips, scenes, and shots based on queries, which could include textual description, image, audio and/or video samples. Therefore, it involves automatic transcription of speech, multi-modal video and audio indexing, automatic learning of semantic concepts and their representation, advanced query interpretation and matching algorithms,

which in turn impose many new challenges to research. All these topics are entangled in the name "semantic information retrieval" [3].
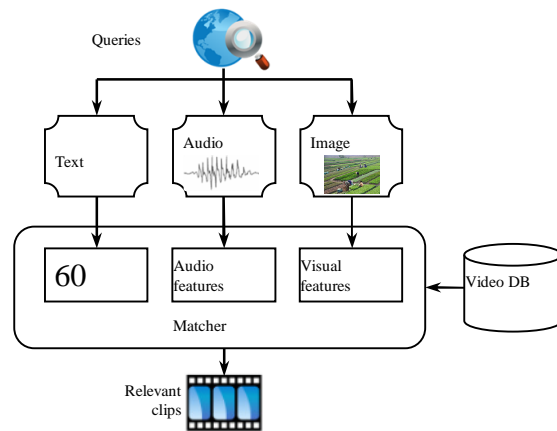

Fig. 1. A full-fledged content-based multimedia retrieval system.

Tackling on semantic information retrieval requires works on both visual and auditory context of the media. This, however, is not a trivial task even with state-of-the-art approaches. Its mandatory challenge, called "semantic gap," [7] requires much more understanding of the way human perceive things (i.e., visual and auditory information). Computer scientists have spent thousands of hours seeking optimal solutions, only ended up falling in the bound of this gap for both visual and spoken contexts. In the spoken context, content-based retrievals are subjected to text-based retrievals by using an automatic speech recognition system to transcribe speech signal into text. Referenced works from [8] and [9] attained an average performance level around 76 % recall and 71 % precision, reasonable enough in academic but insufficient for field applications. Convictions are blamed on the erroneous generated transcription. On the other hand, pathways of visual information retrieval rely on low-level features for advancement, such as colors [10], textures [11], and sketches [12], etc. Nevertheless, these struggling efforts get us

nowhere near human-level perceptions, but only the mediocre temporary solutions. Recent works [13, 14] also introduce a concept-based approach, which makes use of ontology to expand user queries and knowledge indexing.

While an over-the-gap approach is unreachable, we insist on assembling current viable techniques from both contexts, aligned with a domain concept base (i.e., an ontology), to construct an info service for the retrieval of agricultural multimedia information. The development process spans over three packages: (1) building a Vietnamese agricultural thesaurus; (2) crafting a visual-auditory intertwined search engine; and (3) system deployment as an info service. Automatic transcriptions of audio channels are marked as the anchor points for the collection of visual features. These features, in turn, got clustered based on the referenced thesauri, and ultimately tracking out missing info induced by the speech recognizer's word error rates. Meanwhile, the domain ontologies serve as a global linkage between keywords, visual, and spoken features, as well as providing reinforcement for the system performances (e.g., through query expansion, knowledge indexing…).

The rest of this paper is organized as follows. Section II presents the ontology development process in full details. Section III covers our system's specification. Section IV gives experimental results. And finally, Section V concludes the paper.

METHODS

Ontology development

Taking the same model as in [15], we divide the construction of the Vietnamese agricultural ontology into five stages: (1) Ontology specification, (2) Knowledge acquisition, (3) Conceptualization, (4) Formalization and (5) Implementation.

Ontology specification

In this stage, we define the domain and scope of the ontology. The basic questions are what domain the ontology will cover and for what we are going to use the ontology. In our case, the interested domains are aquaculture and plant production, including their diseases, breeding and harvesting methods, etc. The main purpose of the ontology is to maintain and share the knowledge in the field and increase the retrieval efficiency.

Knowledge acquisition

The first step is to gather and extract as much as possible related knowledge resources from the literature, then categorize them systematically. Common groups of resources are ontology construction guidelines and criteria, related thesauri and dictionaries, and relationship guidelines. For this research, we follow general guidelines and criteria, for example, [16] and [17]. Terms are collected from 5 Vietnamese textbooks. We also extract and translate terms from FishBase [18], a global species database of fish species, and the NAL Thesaurus [19]. Then we organize and summarize all of the related information.
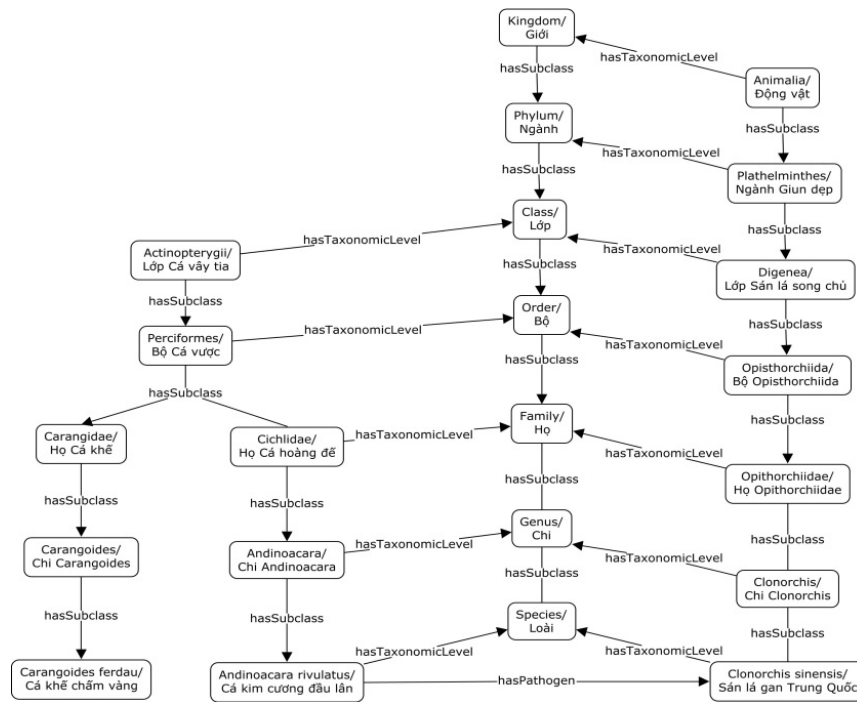
Fig. 2. An example conceptual model of the Vietnamese aquaculture ontology.

Conceptualization

In this stage, a conceptual model of the ontology will be built, consisting of concepts in the domain and relationships among them. Concepts are organized in hierarchical structures; with each concept has its superclass and subclass concepts. Two main groups of relationships are hierarchical relationships and associative relation-ships. To identify concepts, we use both the top-down and bottom-up approaches [20]. The top-down approach can be used to identify hierarchical structures, while the bottom-up approach completes these structures by identifying bottom-level concepts and defining upper-class concepts until reaching the top. For hierarchical relationships, we use only one relation namely "hasSubclass". Concepts in different hierarchies that are related will be connected by associative relationships. Knowledge modeling tools, i.e. CmapTools [21], can be used for sketching the model. Fig. 2 illustrates an example model in our aquaculture ontology.

Formalization

The conceptual model from the previous stage is transformed into a formal model in this stage. We list all the concepts and relationships in a data sheet. Then for each concept, we define a term representing the concept, which is called "preferred term". Synonym, or "non-preferred term", is a term in a same concept that is not selected to be the preferred term. Then we define the terminology relationships that are concept-to-term relationships, term-to-term relationships, and concept-to-concept relationships. The next step involves filling to formalize the concepts. There are three kinds of data sheet: data sheet for concept lexicalization, data sheet for formalizing concept and hierarchical relationship, and data sheet for formalizing concept and associative relationship.

Implementation

Finally, we can implement the ontology by using the Protégé tool [22]. Protégé is a feature rich ontology-editing environment with full support for the OWL 2 Web Ontology Language.

RESULTS

Ontology development

Following the development process, we have developed two Vietnamese agricultural ontologies in two different sub-domains, namely aquaculture and plant production. Our ontologies come with two languages, Vietnamese and English. We also develop a simple web application for searching terms in the ontologies.

Table 1. Concepts of the aquaculture ontology

Object concept	Functional concept
Plant (weed, moss) / Thựcvật (rong, cỏ dại)	Breeding process / Quá trình sinh sản
Animal (fish, mollusk, and amphibian) / Động vật (cá, giáp xác và lưỡng cư)	Pond preparation process / Quá trình chuẩn bị ao nuôi
Fungi / Nấm	Harvesting process / Phương pháp thu hoạch
Bacteria / Vi khuẩn	Protection and control process / Phương pháp kiểm soát và bảo vệ
Virus / Vi-rút	Cultivation process / Phương pháp nuôi trồng thủy sản
Chemical substance and element / Chất hóa học	
Fish anatomy / Giải phẫu học về cá	
Disease / Bệnh	
Environmental factor / Yếu tố môi trường	

Table 2. Concepts of the plant production ontology

Object concept	Functional concept
Plant (rice, fruit) / Thực vật (cây lúa, trái cây)	Plant genetic and breeding / Gen và nhân giống cây trồng
Animal (pest and natural enemy) / Động vật (sâu bệnh và thiên địch)	Soil preparation process / Quá trình chuẩn bị đất
Fungi / Nấm	Fertilizing process / Phương pháp bón phân
Bacteria / Vi khuẩn	Harvesting process / Phương pháp thu hoạch
Virus / Vi-rút	Protection and control process
Chemical substance and element / Chất hóa học	Cultivation process / Phương pháp nuôi trồng
Plant anatomy / Giải phẫu học về cây trồng	
Disease / Bệnh	
Environmental factor / Yếu tố môi trường Soil / Đất	

Table 3. Number of aquaculture ontology relationships

Relationship	Number
Equivalent relationship	2
Hierarchical relationship	1
Associative relationship	25
Total	28

Table 4. Number of plant production ontology relationships

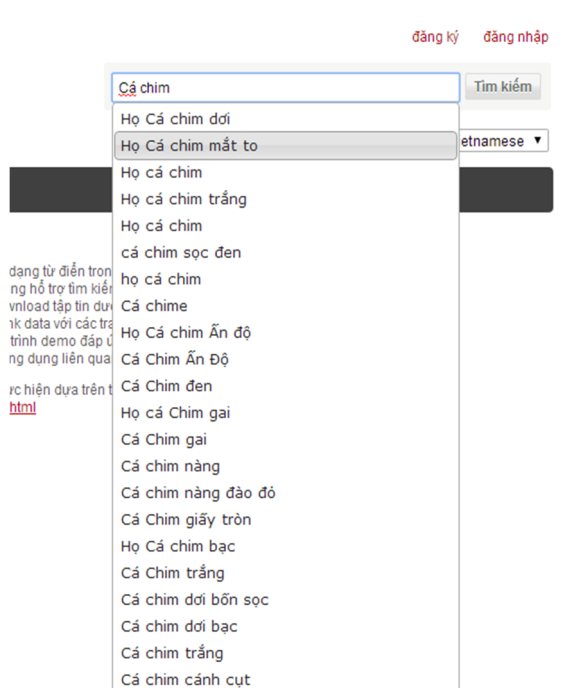| Relationship	Number
Equivalent relationship	3
Hierarchical relationship	1
Associative relationship	1
Total	5

Fig. 3. Ontology searching feature with auto term completion.

The aquaculture ontology consists of 3455 concepts and 5396 terms, with 28 relationships. It covers about 2200 fish species and their related terms. The plant production ontology comprises of 3437 concepts and 6874 terms, with 5 relationships, covering farming, plant production, pests, etc. The ontologies are categorized as classes to provide a comprehensive framework. The categories of the ontologies are summarized in Table I and Table II. The number of relationships is given in Table III and Table IV. While being developed separately, the two ontologies share a fair number of classes, so merging them could be seen in a near future.

There is difference in the number of associative relationships between two ontologies because we use different relationship guidelines. The plant production ontology follows the NAL Thesaurus, which has only one associative relationship, namely "Related to." The aquaculture thesaurus, on the other hand, follows the AGROVOC ontology, where additional relationships are defined, for example, "has Infecting Process," "has Host" or "has Natural Enemy."

A web-based application for searching terms in the ontology was also developed. It provides additional functions to enhance the ontology browsing capability, for instance, bilingual searching (in English and Vietnamese), auto term completion, and external links to other resources. Some of the application's functions are illustrated in Fig. 3.

Content-based agricultural multimedia information retrieval system

The prominent concept of this work basically relies on the composition of visual and auditory (i.e., specifically speech) information, intertwining into each other by their ontology's keyword linkages. Fig. 4 illustrates the construction of this idea – our proposed semantic information retrieval framework.
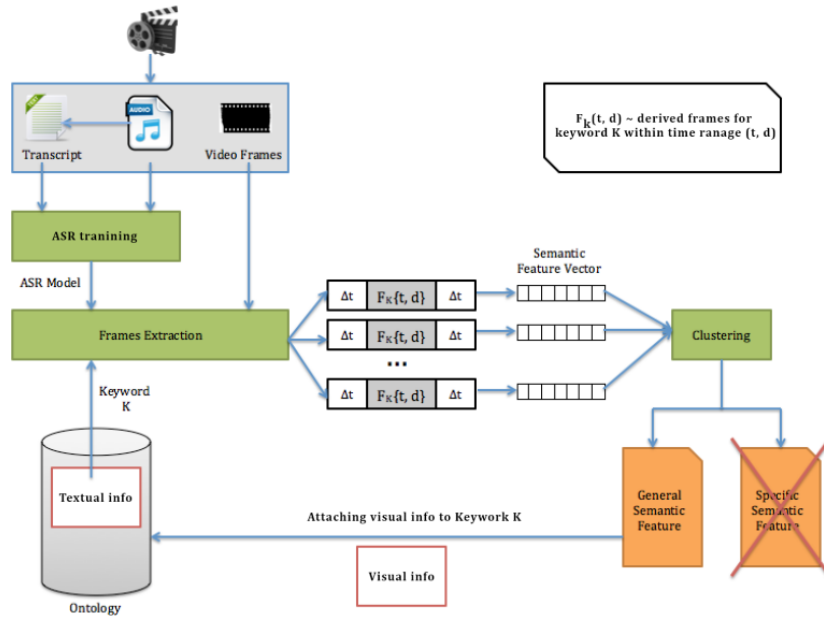
Fig. 4. Intertwined visual-spoken information retrieval framework.

Amongst the three seemingly independent channels, spoken wordsserves as the mainstream for content inference, while visual features help in salvaging missing contents induced by the speech recognition error rates. Both are pinned to the timeline by textual transcriptions and the concept-based linkages (Ontology). Thus forms the relationships between text, speech, and image in our framework. The following Subsections will describe our system in details.

System construction

For each video crawled from the online sources, we demultiplex it into audio and visual channels, which are later segmented into a sequence of frames. The audio part gets manually transcribed to serve as a training corpus for building the ASR module. This in turn, performs a force-alignment procedure on all video files, making them annotated with timestamps and keywords. Now, we define a concept shot F_k as follow:

$F_k(t, d)$~ derived frames clamped by keyword K begin at timestamp t and last for duration d

With the pre-built agricultural ontologies O, we then proceed to extract the concept shots F_{k-i} defined by all keywords K-i existed in the ontologies, positioned by the timestamps generated from the ASR module. With this way, our video database is now chopped down into segments – a set of concept-shots. We also keep track of their contextual information by padding them with adjacent frames for a short leap Δt. F_k is then refined as:

$$F_{k-i}(t_i - \Delta t, d + 2\Delta t), \quad i \in [1\dots|O|], \quad k_i \in O$$

Despite seeming scattered, concept-shots are closely related to each other, in term of concept relationships and inferring. Consider using a decision tree clustering technique [23], global shots would be divided into local groups where members share the same conceptual representation. HMM-GMM cluster-modeling is then taken place on the group's visual features. With the presence of ontologies, specific semantic visual features are no longer required, and thus low-level features might be sufficient enough (i.e., ontologies take care of rendering the semantic layers). Here, we use a feature bag of

Harris cues, edge, color, blob, and ridge. Fig. 5 shows how concept-shots are shaped and clustered on each other through the linkage of ontologies.

Classification

Any future unseen media collected from the online sources will be auditorily transcribed and visually clustered into one of the available classes of our ontology (i.e., keywords or concept-shots). The classification of concept-shots would definitely compensate for word-error-rates of the transcriptions, and ultimately tracking out missing info potentially available in the media.
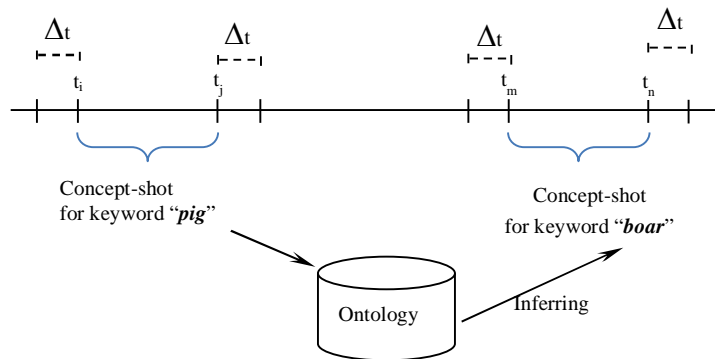

Fig. 5. Illustration of concept-shots and ontology-inferred clustering.

For example, in Fig. 5, if the feature bag of the "boar" shot is classified into the same group as "pig," then we would assume that there would be some kind of pig in that shot (e.g., the wild boar for this case).

Deployment

To make the whole system a viable application, we have wrapped it into an info service, maintained as an AIS structure [25]. Our target audiences are the majority of farmers in developing countries, who are unable to reach the modern farming information and knowledge. The info service is protocol- and platform-independent. It can be accessed by any front-end devices, from traditional mobile phones to PC, or smartphones, etc.

The service is being hosted in its beta stage at:

http://www.ailab.hcmus.edu.vn

This section presents the results captured from our experimental procedure. Comparative analyses between a preset baseline (i.e., the speech-based only system built using the same ASR approach in our previous work [24]) and the proposed system are taken place to measure how well it performs. All of which are conducted in the corpus described below.

Datasets

Roughly 40 hours of agricultural broadcast videos are collected from multiple broadcasting studios in Mekong Delta. We requested the original media instead of the recorded ones for their upper quality. Audio channels are sampled in 16 KHz, 16 bits, mono. And video channels are normalized in standard 480p. The corpus is then manually transcribed and divided into 3 subsets: training, development and test sets. Table V gives a detailed look into these subsets.
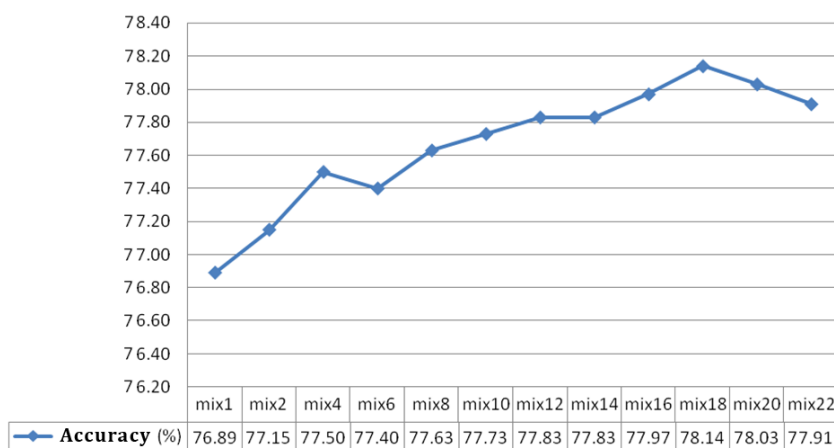
Table 5. Datasets

Corpus	Duration (hours)
Training set	20
Development set	1
Test set	19
Total	40

The training set is used for training ASR and building concept clusters, which are then verified and tuned with the development set. Retrieval performances are finally measured upon the test set.

Parameter tuning

This experiment measures performances of the speech recognizer on the development set to further fine-tune system's parameters. We construct the ASR engine using traditional left-right tied-triphone HMM-GMM pattern. Recognition tasks include 412 utterances segmented from 1-hour speech of agricultural conversation (i.e., development set). Fig. 6 plots the performance function of the recognizer. As the number of mixtures increases, accuracy acceleration slows down and reaches its limit eventually. In the best case, 78.14 % WAR (word accuracy rate) is achieved.

	mix1	mix2	mix4	mix6	mix8	mix10	mix12	mix14	mix16	mix18	mix20	mix22
Accuracy (%)	76.89	77.15	77.50	77.40	77.63	77.73	77.83	77.83	77.97	78.14	78.03	77.91

Fig. 6. Transcription performances.

Independently from unseen data, we choose the best configuration of 18 mixtures. Transcriptions generated by this one alone also are served as the indexed database for the baseline retrieval system.

The same routine applies for choosing a number of mixtures in each cluster-model. Feature bags extracted from 1-hour video are classified into one of 27 concept-classes found in the development set. With each model configuration, we logge down the classification accuracy as in Fig. 7, leading to the selection of 32-mixutre candidate.
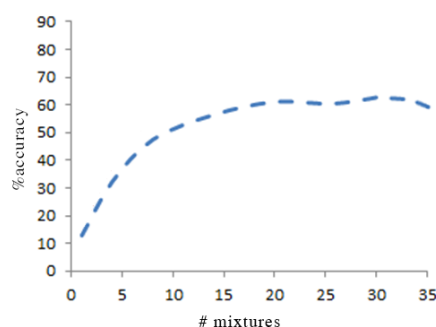
Fig. 7. Clustering performances.

Retrieval evaluations

Having set the ground for the baseline system, ASR engine, and clustering models, we proceed to assess our proposed system upon the remaining 19-hour test set. 500 pseudo test-queries are constructed by randomly choosing queried targets from within 6892 Ontology concepts in mono (e.g., banana) and dual association (e.g., banana cultivation) manners. Pseudo queries without relevant ground-truths are filtered out to ensure the requested documents fall within the corpus's bound, thus making no false claim on missing retrievals.

Table 6 reports average recalls and precisions in a comparative manner for: speech-based system (baseline), vision-based system, and visual-auditory intertwined system. Since the semantic gap is too much for low-level features, vision-based system seems falling back behind, while speech-based system renders recall closely to its transcription accuracy. False alarms did rise, because both system neglects the semantic layer. However, when combining the spoken and visual features together under Ontology's linkages, we found the results shooting upward, attaining absolute increases of 14.3 % recall and 9.1 % precision over the baseline system.

Table 6. Retrieval performances

Metrics	Speech-based system	Vision-based system	Intertwined system
Recall	70.6 %	56.1 %	84.9 %
Precision	79.2 %	64.5 %	88.3 %

CONCLUSION

For long shackled within the semantic gap, we have being pursued a way out and more ideally an optimal solution. But not many achievements had been gained since our first approach of Vietnamese speech-based video retrieval in 2010. As the concept-based retrieval approaches rise in recent years, we made an attempt to plan out a compensation technique that employ the use of visual features and Ontology together. Experimental results did confirm the hypothesis. Despite being a long way from human perceptions, the composite scheme surely shed light on applicable solutions for semantic information retrieval. We also deploy our system as an info service to support agricultural extension in Mekong Delta.

Tiếp cận đa kết hợp cho bài toán truy vấn thông tin nông nghiệp đa phương tiện theo nội dung

- **Phạm Minh Nhựt**
- **Phạm Quang Hiếu**
- **Lương Hiếu Thi**
- **Vũ Hải Quân**

TÓM TẮT

Truy vấn thông tin đa phương tiện theo nội dung là một thách thức lớn, ngay cả đối với hiện trạng công nghệ hiện nay. Vấn đề nan giải của bài toán này, có tên gọi "khoảng cách ngữ nghĩa", đòi hỏi chúng ta phải thấu hiểu cặn kẽ cách mà con người cảm nhận và xử lý thông tin thính-thị giác. Giới nghiên cứu đã triển khai hàng nghìn giờ thực nghiệm, mong muốn tìm ra giải pháp tối ưu, kết cuộc cũng bị hạn chế bởi chính khoảng cách ngữ nghĩa này. Trong khi lời giải toàn cục vẫn chưa có, chúng tôi đề xuất kết hợp kỹ thuật xử lý ngôn ngữ nói và thị giác máy tính, liên kết theo một cơ sở ngữ nghĩa cục bộ (domain Ontology), để đưa ra một giải pháp khả thi trong ứng dụng – xây dựng dịch vụ thông tin nông nghiệp. Quá trình phát triển dàn trải qua 3 phần: (1) xây dựng bộ cơ sở ngữ nghĩa cho nông nghiệp Việt Nam, (2) kiến thiết mô hình lai thính-thị giác cho động cơ tìm kiếm, (3) triển khai hệ thống dưới dạng dịch vụ thông tin. Về cơ sở ngữ nghĩa, chúng tôi phát triển theo 2 hướng: nhánh Ontology thuỷ sản với 3455 khái niệm, 5396 từ, và 28 quan hệ; nhánh Ontology trồng trọt với 3437 khái niệm, 6874 từ, và 5 quan hệ. Hệ Ontology này đóng vai trò liên kết toàn cục giữa từ khoá với các đặc trưng thính-thị giác, đồng thời cũng gia cố thêm cho hiệu năng tìm kiếm thông qua các kỹ thuật như mở rộng truy vấn, đánh chỉ mục tri thức... Ở mặt khác, mô hình lai cho động cơ tìm kiếm được thiết kế theo hướng tận dụng suy diễn ngữ nghĩa trên bộ Ontology để truy vét các thông tin bị mất. Phương pháp thực hiện là gom cụm các đặc trưng thị giác theo từng nhóm ngữ nghĩa được phát sinh từ bộ nhận dạng tiếng nói. Những phân đoạn thông tin nào bị mất mát do sai số trong nhận dạng tiếng nói, chỉ cần cùng cụm thị giác với các phân đoạn đúng, đều có thể truy vấn lại được. Kỹ thuật này đã đem lại cho chúng tôi 14 % gia tăng độ phủ và 9 % gia tăng độ chính xác so với hệ thống nền. Từ kết quả này, chúng tôi cũng đã triển khai một dịch vụ thông tin, hỗ trợ kỹ thuật nông nghiệp cho bà con nông dân vùng sâu vùng xa.

Từ khoá: truy vấn thông tin hướng ngữ nghĩa, truy vấn video theo nội dung, nông nghiệp, đa phương tiện, dịch vụ thông tin, Ontology nông nghiệp

REFERENCES

[1]. General Statistics Office, Thông cáo báo chí Tình hình kinh tế - xã hội năm 2013, retrieved September 18, 2014, from http://www.gso.gov.vn/.

[2]. K. Markey, Twenty-five years of end-user searching, Part 2: Future research directions, *Journal of the American Society for Information Science and Technology*, 58, 8, 1123-1130 (2007).

[3]. A. Amir, et al., A multi-modal system for the retrieval of semantic video events, *Computer Vision and Image Understanding*, 96, 2, 216–236 (2004).

[4]. L. Ballan, M. Bertini, A.D. Bimbo, G. Serra, Semantic annotation of soccer videos by visual instance clustering and spatial/temporal reasoning in ontologies, *Multimedia Tools and Applications*, 48, 2, 313–337 (2010).

[5]. A. Fujii, K. Itou, T. Ishikawa, LODEM: a system for on-demand video lectures, *Speech Communication*, 48, 5, 516–531 (2006).

[6]. A.G. Hauptmann, M.G. Christel, R. Yan, Video retrieval based on semantic concepts, *Proceedings of the IEEE*, 96, 602–622 (2008).

[7]. G. Martens, P. Lambert, R. Walle, Bridging the semantic gap using human vision system inspired features, Self-Organizing Maps, *In Tech Open* (2010).

[8]. M.G. Brown, J.T. Foote, G.J. Jones, K.S. Jones, S.J. Young, Automatic content-based retrieval of broadcast news, In Proceedings of the third ACM international conference on Multimedia, *ACM*, 35-43 (1995).

[9]. B. Adams, G. Iyengar, C. Neti, H.J. Nock, A. Amir, H.H. Permuter, D. Zhang, IBM Research TREC 2002 Video Retrieval System, In TREC (2002).

[10]. T. Gevers, A.W. Smeulders, Pictoseek: Combining color and shape invariant features for image retrieval, *Image Processing, IEEE Transactions*, 9, 1, 102-119 (2000).

[11]. W.Y. Ma, B.S. Manjunath, Netra: A toolbox for navigating large image databases, *Multimedia systems*, 7, 3, 184-198 (1999).

[12]. A.D. Bimbo, P. Pala, Visual image retrieval by elastic matching of user sketches. Pattern Analysis and Machine Intelligence, *IEEE Transactions*, 19, 2, 121-132 (1997).

[13]. A. Jaimes, J.R. Smith, Semi-automatic, data-driven construction of multimedia ontologies, In Multimedia and Expo, 2003. ICME'03, Proceedings. 2003 *International Conference on IEEE*, 1, I-781 (2003).

[14]. L. Hollink, M. Worring, A.T. Schreiber, Building a visual ontology for video retrieval, *In Proceedings of the 13th annual ACM international conference on Multimedia ACM*, 479-482 (2005).

[15]. A. Thunkijjanukij, Ontology development for agricultural research knowledge management: a case study for Thai rice, PhD dissertation, Kasetsart University, Thailand (2009).

[16]. F.N. Natalya, L.M. Deborah, Ontology development 101: A guide to creating your first ontology (2001).

[17]. United States Department of Agriculture, "Agricultural Thesaurus", accessed September 18, 2014, http://agclass.nal. usda.gov.

[18]. M. Uschold, M. Gruninger, Ontologies: Principles, methods and applications, *Knowledge Engineering Review*, 11,02, 93-136 (1996).

[19]. R. Froese, FishBase, *Oceanographic Literature Review*, 43, 3 (1996).

[20]. T.R. Gruber, Toward principles for the design of ontologies used for knowledge sharing?, *International Journal of Human-computer Studies*, 43, 5, 907-928 (1995).

[21]. A.J. Cañas, G. Hill, R. Carff, S. Niranjan, J. Lott, T. Eskridge, G. Gómez, M. Arroyo, R. Carvajal, CmapTools: A knowledge modeling and sharing environment, *Proceedings of the 1st International Conference on Concept Mapping*, 1 (2004).

[22]. H. Knublauch, R.W. Fergerson, N.F. Noy, M.A. Musen, The Protégé-OWL plugin: An open development environment for semantic web applications, *Proceedings of the 3rd International Semantic Web Conference*, Japan (2004).

[23]. Q. Vu et al., A Robust Vietnamese Voice Server for Automated Directory Assistance Application, *VLSP* (2012).

[24]. Q. Vu et al., Soccer Event Retrieval Based on Speech Content: A Vietnamese Case Study, Speech Technologies, Book 2, *Intech Open Access Publisher* (2011).

[25]. A. Hall, Agricultural Innovation Systems: An Introduction, Link-UNU-Merit.