Towards Semantic Clustering – A Brief Overview

Phei-Chin Lim pclim@fit.unimas.my
Faculty of Computer Science and Information Technology
Universiti Malaysia Sarawak
94300 Kota Samarahan, Sarawak, Malaysia

Narayanan Kulathuramaiyer nara@fit.unimas.my
Faculty of Computer Science and Information Technology
Universiti Malaysia Sarawak
94300 Kota Samarahan, Sarawak, Malaysia

Dayang NurFatimah Awg. Iskandar dnfaiz@fit.unimas.my
Faculty of Computer Science and Information Technology
Universiti Malaysia Sarawak
94300 Kota Samarahan, Sarawak, Malaysia

Abstract

Image clustering is an important technology which helps users to get hold of the large amount of online visual information, especially after the rapid growth of the Web. This paper focuses on image clustering methods and their application in image collection or online image repository. Current progress of image clustering related to image retrieval and image annotation are summarized and some open problems are discussed. Related works are summarized based on the problems addressed, which are image segmentation, compact representation of image set, search space reduction, and semantic gap. Issues are also identified in current progress and semantic clustering is conjectured to be the potential trend. Our framework of semantic clustering as well as the main abstraction levels involved is briefly discussed.

Keywords: Image Clustering, Semantic Gap, Semantic Clustering, Concept Description, Symbolic Description.

1. INTRODUCTION

Advancement in digital imaging devices, technology and cost-decrease in storage devices contributed to the creation of large-scale digital images in various domains. The question now is how to effectively extract semantically meaningful information (knowledge) from these image collections. One of the fundamental of understanding and learning is to discover the natural groupings of images based on similarity of multiple characteristics or latent aspects of meaning. In this paper, the former referred to image clustering while the latter leads to semantic clustering.

One of the reasons for writing this paper is that we hardly find any work in reviewing image clustering methods by highlighting the challenges and provide association between image clustering and semantic clustering. Semantic clustering has attracted many research efforts after the year 2000 in terms of papers published. A simple experiment is conducted where we searched for publications containing the phrases *image clustering* and *semantic clustering* using the digital libraries of ACM [1], IEEE [2], ScienceDirect [3] and SpringerLink [4] for the 2000 to 2010. Searches included the specific search phrase in publication title only and in publication title or abstract without restriction on research field. As observed in Figure 1 and 2, publication on *semantic clustering* is lower because papers found should be subset of the papers found using *image clustering.* The number of publication for both search terms in Figure 2 are higher as expected, as the search also include abstract field in addition to title field only as recorded in Figure 1. This difference indicates that the research in *semantic clustering* is still in its infancy stage where many publications have not included the term semantic clustering in the title of the paper but have started to appear in the abstract of these publications. Another observation is that the actual growth of publications on *semantic clustering* started in year 2005 and increased gradually throughout the years, which is in accordance with augmented research focus on the issue of semantic gap. Note should be taken that the number of publications shown in this paper is for reference purposes and not to be taken as quantitative proof of the actual publication count.

For the purpose of completeness and better readability for the uninitiated, we have introduced key applications of image clustering and their challenges in Section 2. Discussion on the open issues and summarization of the reviewed works is presented in Section 3. General organization of our solution is illustrated in Section 4 before we conclude in Section 5.

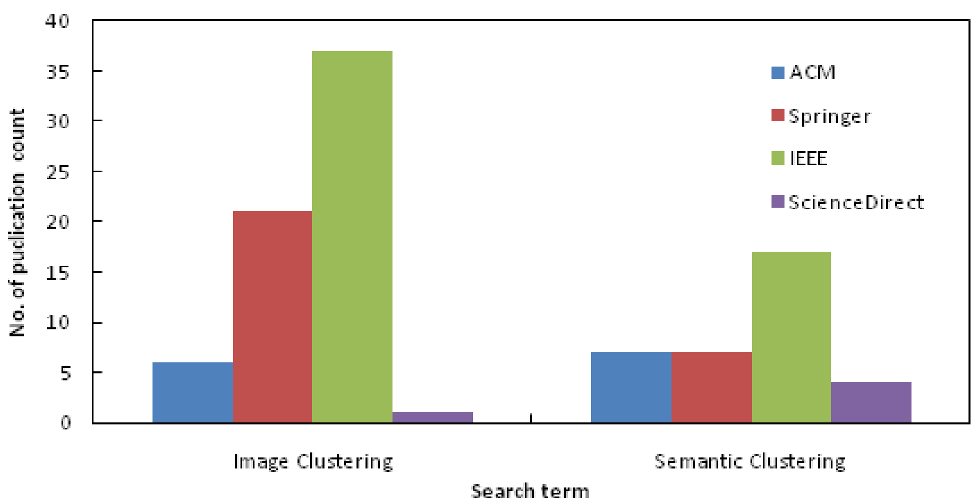


FIGURE 1: Publication count on papers with terms *image clustering* and *semantic clustering* in title only.

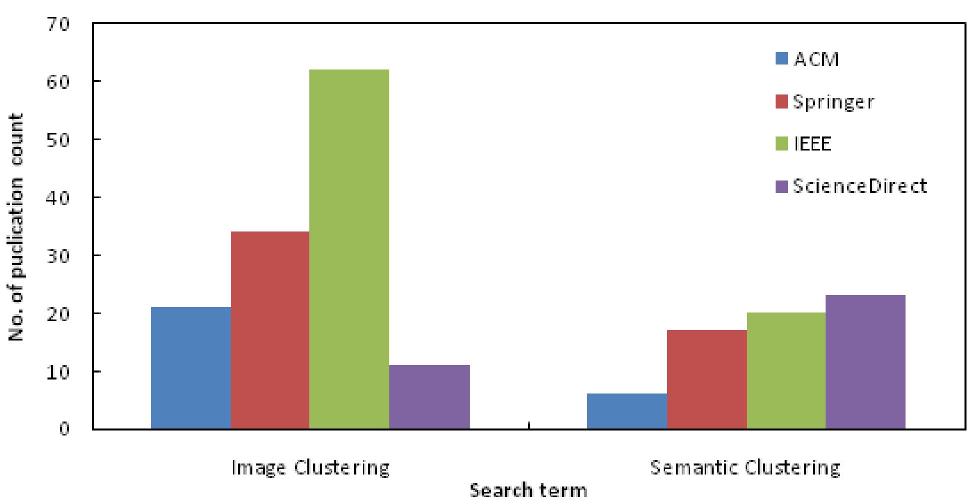


FIGURE 2: Publication count on papers with terms *image clustering* and *semantic clustering* in title or abstract.

2. IMAGE CLUSTERING

There is no universally agreed upon definition for clustering [5], hence, image clustering is generally accepted as the application of cluster analysis in order to support image understanding. Consequently, it inherited the challenges faced by cluster analysis. Interested readers may refer to the vast literature on cluster analysis [5, 6, 7] which will not be cover in this paper. In the following section, reviews on application of image clustering are organized based on the problem addressed and focus is given on employing an unsupervised method to solve these problems.

2.1 Image Segmentation

Initially, unsupervised clustering approach is mostly engaged in Geographic Information System for identifying and segmenting images into desired regions prior to being transformed into medical image analysis. Most work addressing the segmentation-classification problem requires certain level of a priori knowledge. In [8], domain knowledge about real brain CT images is quantified before incorporation into clustering algorithm to cluster region of interest. Images are then classified as normal or abnormal based on the similarity of region of interest. Counterpart of Principal Component Analysis, non-negative matrix factorization is applied on brain CT images to extract visual features and histogram-based semantic features for identifying normal and tumor CT images in [9] while a density function based image clustering analysis is used for the segmentation and identification of abdomen CT images for diagnosis purposes in [10]. In [11], a semantically supervised clustering approach is used to classify multispectral information into geo-images. Again, a priori knowledge is incorporated in the clustering process as selection criteria of the training data. Other work on detecting interesting regions by mean of clustering can be found in [12, 13].

2.2 Compact Representation

Image collections are usually heterogeneous, which makes extraction of representation a hard task. Recently, there have been growing interests in employing unsupervised methods to improve the way of representing images sets. Commonly used cluster representation schemes includes representing cluster by their centroid or by a set of distant points, nodes in a classification tree and conjunctive logical expressions [6] or newly emerging graphical representation of heat map [14] which is a multi-feature representation.

Authors in [15] addressed the problem of image representation and clustering based on learning a lower dimensional representation of the image manifold embedded in the ambient tensor space. Listed contributions including a more natural representation by treating images as matrices, computationally efficient compared to traditional dimensionality reduction algorithms which only applicable to vector data. An unsupervised neural network learning and self-organizing tree algorithm that automatically constructs a top-to-bottom hierarchy is proposed in [16]. Objects are firstly cluster according to similarity between objects in term of colour, shape and texture features before clustering images. Comparison with traditional hierarchical agglomerative clustering showed promising accuracy. Other work can be found in [17].

2.3 Search Space Reduction

Assumption and representation of each image as a whole do not really fit what a user is focusing on in an image, which is a part of an image or region in image with semantic meaning. Therefore, representing each image as a set of regions or objects is the desirable setting. Yet, this leads to expanded search space which makes retrieval efficiency a critical issue. In order to tackle the issue, we summarize works into two categories based on the instant clustering algorithm is applied.

First category comprises of works where image region clustering is done offline in the pre-processing stage and will be re-activated only when growth of new images reaches a limit. Reduction of search space is achieved by performing clustering before image retrieval. A massive 74% search space reduction is achieved in [18]. The strength of the algorithm is that the number of clusters is learned from user query log before being refined using outlier detection method.

Authors in [19] proposed an interactive approach where user's feedback on initial search results is required for analyzing user interest. Both works make use of Corel image sets.

Clustering is performed online in the post-processing stage in the second category. In [20], search results from initial query are clustered into semantic groups together with learned labels which provide overview of the content of the results. Users are able to browse through each cluster easily. Similar work can be found in [21, 22]. Humans have higher tendency to use high-level abstract concepts (image semantics) during querying or browsing of image collection rather than using low-level visual features. Consequently, huge amount of irrelevant search results are returned due to the semantic gap described in the previous section. For this reason, cluster search results are converted into symbolically similar clusters in order to filter out the relevant/irrelevant images in [23].

2.4 Semantic Gap

Content-based Image Retrieval (CBIR) is the technology that organizes digital images by their visual content which was haunted by the critical challenge of the semantic gap being defined as "the lack of coincidence between the information that one can extract from the visual data and the interpretation that the same data have for a user in a given situation" [24]. Several taxonomies of methods in addressing the problem of semantic gap were suggested in [25] and [26]. Clustering is adapted in CBIR to solve the semantic gap issue starting mid 2000's onwards.

In a typical CBIR system, query results are a set of images sorted by feature similarities but there is no guarantee that the returned images will be similar in semantic as shown in Figure 3 (a). Figure 3 (b) showed the result of retrieved images that are similar in semantic and content [27], where butterfly wing images were firstly clustered into semantic shape categories before performing similarity search to ensure only images in the same shape category are considered in finding similar images to the query. CLUE [28] approach the same issue differently by grouping semantically similar images with respect to query and return image clusters instead of a set of ordered images. Similar work was done in [20, 21, 29].

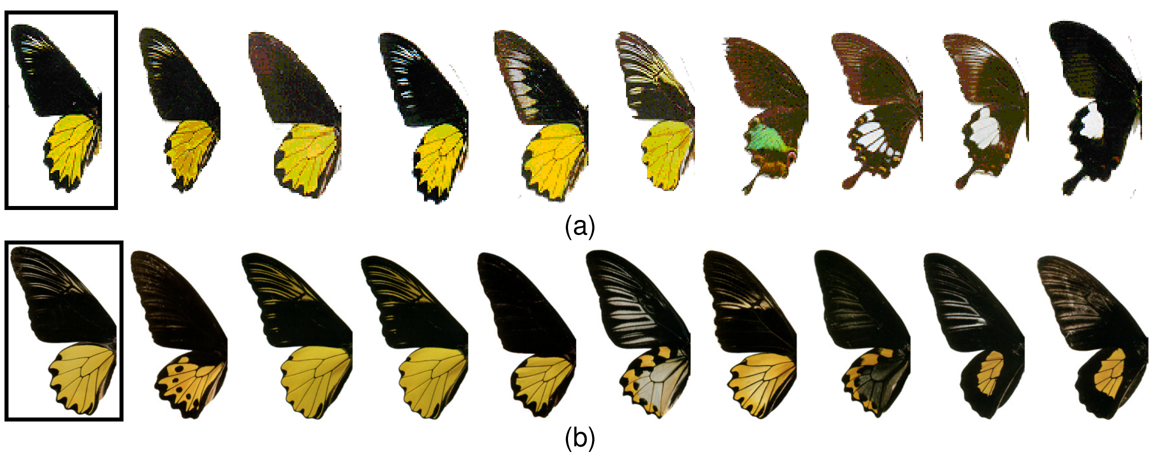


FIGURE 3: Ten most similar images retrieved (query image in highlighted), (a) conventional similarity-based approach, (b) semantic shape category from [27].

3. OPEN ISSUES

In brief, image clustering is the organization of images into different clusters (groups, classes or categories) based on similarity such that images in the same cluster share some common characteristics. A crucial issue of image clustering on large-scale image repository is compact representation for faster image indexing and searching. The derived clusters provide summarization of the image content that can be utilized for tasks such as image search and browsing. Refer to Table 1 for summarization of selected works that employed clustering algorithms.

Problem		Authors	Feature	Approach	Cluster Description
Image Segmentation		Pan et.al [8]	Object's gray level, size, location, elongation, direction	Object clustering followed by image clustering	Normal/Abnormal
		Liu et.al [9]	Visual features; histogram-based semantic features	Non-negative Matrix Factorization	Normal/Tumor
		Torres et.al [11]	Spatial semantics (geometrical & topological properties)	Maximum Likelihood	Categories label (water, urban, crop, landslide, vegetation etc)
		Han et.al [13]	Texture homogeneity	Self-organizing feature map, agglomerative clustering	Nil
Search Space Reduction	Before retrieval	Liu et.al [18]	3 texture features, 2 shape features, 27 colour features	Affinity matrix clustering, network flow - based outlier detection (refine clustering result)	Nil
		Zhang & Chen [19]	13 colour features, 6 texture features	Genetic algorithm	Categories label (horse, eagle, flower, fish, glass etc)
	After retrieval	Wang et.al [22]	6 color moments, 44 auto-correlogram, 14 colour-texture moments; keywords (from image titles & description)	K-means; Image Frequency-Inverse Keyword Frequency ; Random Walk with Restarts	Annotations
		Tahayna et.al [23]	11 colour words, 11 texture words, 8 shape words	Three-layer fuzzy partitioning	Nil
Compact Representation		He et.al [15]	32x32 dimensional matrix	K-means	Nil
		Wang & Khan [16]	Object's shape, colour & texture features	Dynamic Growing Self-Organizing Tree	Nil
Semantic Gap		Lim et.al [27]	Shape features	k-Means	Visual concept & its characterization
		Chen et.al [28]	3 average colours, 3 texture features	K-means; Graph-theoretic clustering	Nil
		Wang et.al [20]	Textual features (phrases)	Salient phrase ranking, Image Search Result Clustering	10-50 final cluster names (dependent on query word) after merge & prune, other cluster
		Gao et.al [29]	Colour histogram, texture histogram	Kernel-based clustering	Nil

TABLE 1: Summarization of selected works on image clustering.

In real life, humans tend to query images using high-level abstract concepts rather than visual features. Correlation between image content and the associated text is assumed to be strong but this may not be the case. Hence, most keyword-based image search tends to return a large amount of irrelevant images to a given query. Another spectrum of the issue is due to the computation of image similarities using visual content alone, by minimizing intra-cluster difference or maximizing inter-cluster differences in the feature space. As a result, two semantically similar images may lie far from each other in the feature space while two completely different images may stay close to each other in the same feature space (refer to example in Figure 3).

Other researchers tried to incorporate textual features to solve the mentioned issue [30, 31, 32] with a certain level of success. These textual features are extracted from external sources such as ontologies, Web pages [33] and camera metadata [34]. Different sources face different challenges. Ontologies are scarce and domain dependent where creating one from scratch is time-consuming and laborious. Information provided by Web pages is unstructured in nature where reliability is an open issue. Researchers seem to be gaining interest in utilizing the rich embedded information in camera metadata, yet thorough investigation is required to prove that any value in camera metadata is correlated with image semantic.

Another observation is that the reviewed techniques build clusters solely on the basis of numerical measure. Works that provides explanation through concepts or semantic properties is scarce but is the more desirable future trend. These concepts or semantic properties explain why a set of objects conform into a cluster.

4. TOWARDS SEMANTIC CLUSTERING

Semantic clustering originated from information system field to solve text classification problem. In this paper, we refer semantic clustering as the concept of unsupervised learning to group unstructured images based on latent aspect of meaning.

Even though research for solving the semantic gap is moving towards semantic clustering, most of the work is focusing on direct mapping of visual features to semantic concepts. Little attention is given on employing a symbolic-level abstraction for the mapping, which confronts the human behavior of using high-level concepts in finding images of interest. Figure 4(a) the general direct mapping approaches while Figure 4(b) illustrated our proposed solution respectively.

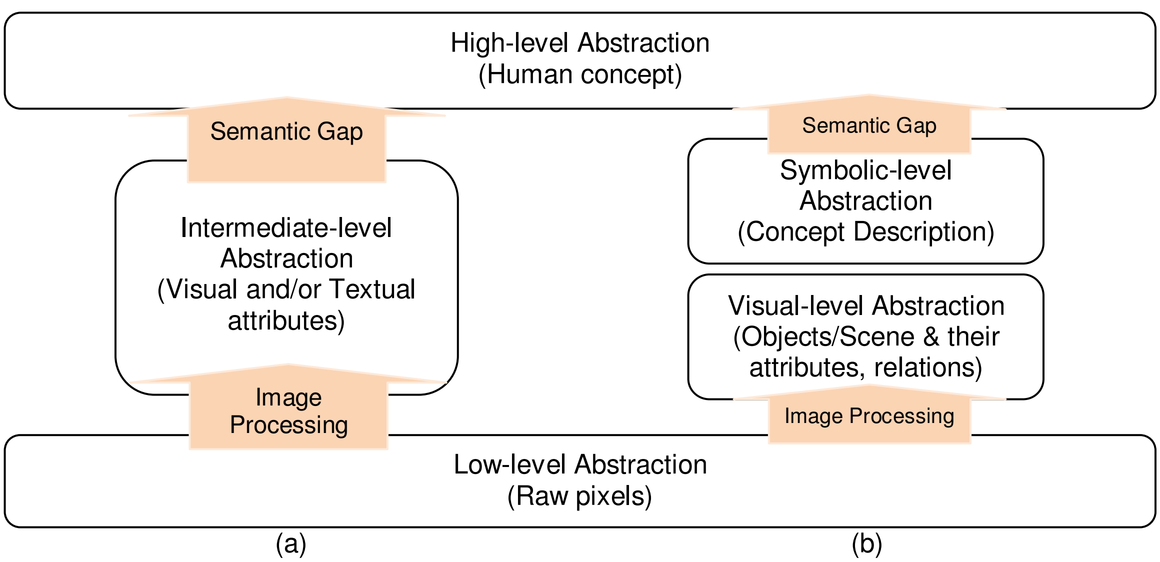


FIGURE 4: Evolution of approaches in bridging the semantic gap: (a) Visual-based; (b) proposed approach.

Our solution is inspired from how author in [23] bring together visual semantics and visual features for automatic image classification. A symbolic-level is inserted on top of visual-level abstraction aiming to map visual features to semantic concepts in order to form concept description. Concept descriptions are logical statements describing each cluster's member, which aim at providing human understandable interpretation and user-friendly searching method. Comparison of selected works which complied with the abstraction levels in Figure 4 is listed in Table 2 and described accordingly.

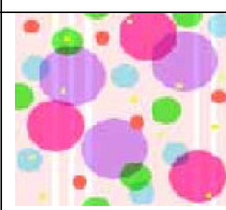
Abstraction level of Images	**Selected work**			
	Awang Iskandar [35] & [36]	Park & Ramamohanarao [37]	Lim [26]	Shin, Kim & Kim [38]
High-level	Object/Character {Bob} Spatial relation {Bob RIGHT OF panel}	Right Apical Edge {Sharply defined} Right Diaphragm {Not sharply defined} … Class {Left Ventricle failure influencing dense ribs OLD inflammatory change at the right apex, bullous change at the left apex Ischemic heart disease and CAL}	Relative Hindwing size {very large} Relative Forewing-Hindwing ratio {Forewing very long} … Genera {*Troides*}	Abstract Emotion {Cute} Object {Bubble}
Symbolic-level	Region tagging→semantic tags/keywords (e.g., Character {Bob}: Region 1 = Body part {head}, Colour {grey}; Region 2 = Body part {eye}, Colour {black})	Symbolic description (e.g. a height of 2cm described as "normal"; height of 2.3cm described as "slightly high")	Visual feature → domain concept (e.g. $\frac{L_f}{L_h}$ → wing ratio) Domain concept → linguistic value (e.g. wing ratio {Forewing very long,…etc})	Visual keywords (e.g. pink, violet, blue, green & circles)
Visual-level	Shape & colour (region area, region mean grey value, compactness & shape boundary), spatial location	Ratio, presence, position, width, size, angle & density	Shape ($\frac{L_f}{L_h}$, $\frac{4*\pi*area}{perimeter^2}$, etc)	130 Kobayashi colour & pattern (Wavelet transform)
Low-level				

TABLE 2: Examples from selected works that complied with the abstraction levels in Figure 4.

Although these selected works are successfully mapped with all the abstraction levels, richness of high-level concept is different. For example, the high-level concept [35] is object label ("comic character") where the usage is to query relevant comic strips, while [38] tried to map colour and pattern to emotional concepts ("Cute", "Classic") in textile domain. Emotional concepts are abstract attributes to describe images and the most difficult indexing level because it is highly subjective and assessments among different users vary greatly. For example, sample textile image in Table 2 may represent "Cute" to one person but "Romantic" to another. Both works do not provide concept description as compared to [26] and [37].

In [37], image features are firstly extracted and analyzed before converted to symbolic descriptions by fuzzy functions (refer to example given in Table 2). Hence, an expert (radiologist) can work with linguistic terms instead of numerical description of object features. Same in [26], visual features are extracted from salient regions before being converted to domain concepts and then further characterization. As final result, each butterfly genera is represented by their corresponding descriptions which is human understandable.

5. CONCLUDING REMARKS

This paper provided an overview of works on image clustering and discussed the challenges faced which led to the shift to semantic clustering. Recent works mostly formed clusters using low-level visual features and directly mapped clusters to high-level human concepts, which is limited in semantic richness. Even though there are attempts to incorporate textual information to enrich image understanding, there is still lack in providing cluster (concept) description. Therefore, a system that is capable to describe clusters symbolically which highlight the interpretability of clusters is highly desired.

6. REFERENCES

1. ACM – Association for Computing Machinery, http://www.acm.org. Accessed November 2010.
2. IEEE – Institute of Electronic and Electrical Engineers, http://ieeexplore.ieee.org. Accessed November 2010.
3. ScienceDirect, http://www.sciencedirect.com. Accessed November 2010.
4. SpringerLink, http://www.springerlink.com. Accessed November 2010.
5. B.S. Everitt. "*Cluster Analysis"* (3rd Edition), Edward Arnold, Ltd., London, UK, 1993.
6. A.K. Jain, M.N. Murty and P.J. Flynn. "*Data Clustering: A Review*". ACM Computing Survey, 31 (3): 264-323, 1999.
7. R. Xu and D. Wunsch. "*Survey of Clustering Algorithms*". IEEE Transactions on Neural Networks, 16 (3): 645-678, 2005.
8. H. Pan, J. Li and W. Zhang. "*Medical Image Clustering with Domain Knowledge Constraint*". W. Fan, Z. Wu, and J. Yang (Eds.): WAIM 2005, LNCS 3739*:* 719 – 724, 2005.
9. W. Liu, F. Peng, S. Feng, J. You, Z. Chen, J. Wu, K. Yuan and D. Ye. "*Semantic Feature Extraction for Brain CT Image Clustering using Nonnegative Matrix Factorization*". D. Zhang (Ed.): ICMB 2008, LNCS 4901*:* 41–48, 2007.
10. Y. Song, C. Xie, Y. Zhu, C. Li and J. Chen. "*Density Function Based Medical Image Clustering Analysis and Research*". K. Elleithy et al. (eds.): Advances in Computer, Information, and Systems Sciences, and Engineering, 149–155, 2006.
11. M. Torres, G. Guzman, R. Quintero, M. Moreno and S. Levachkine. "*Semantic Decomposition of LandSat TM Image*". B. Gabrys, R.J. Howlett, and L.C. Jain (Eds.): KES 2006, Part I, LNAI 4251*:* 550 – 558, 2006.

12. M.S. Hossain, K.A. Rahman, M. Hasanuzzaman and V.V. Phoha. "*A Simple and Efficient Video Image Clustering Algorithm for Person Specific Query and Image Retrieval*". ICIMCS'09. Kunming, Yunnan, China, November 23-25, 2009.

13. E. Han, J. Yang, H. Yang and K. Jung. "*Automatic Mobile Content Conversion using Semantic Image Analysis*". J. Jacko (Ed.): Human-Computer Interaction, Part III, HCII 2007, LNCS 4552: 298–307, 2007.

14. L. Wilkinson and M. Friendly. "*The History of the Cluster Heat Map*". The American Statistician, 63(2):179-184, 2009.

15. X. He, D. Cai, H. Liu and J. Han. "*Image Clustering with Tensor Representation*". MM'05. Singapore, 132-140, November 6-11, 2005.

16. L. Wang and L. Khan. "*A New Hierarchical Approach for Image Clustering*". V. Petrushin and L. Khan (Eds.), Multimedia Data Mining and Knowledge Discovery, 41-57, Springer, London, 2006.

17. S. Xia and E.R. Hancock. "*Clustering using Class Specific Hyper Graphs*". N. da Vitora Lobo et al. (Eds.): SSPR&SPR 2008, LNCS *5342:* 318–328, 2008.

18. Y. Liu, X. Chen, C. Zhang and A. Sprague. "*Semantic Clustering for Region-based Image Retrieval*". Journal of Visual Communication and Image Representation, 20 (2): 157-166, February 2009.

19. C. Zhang and X. Chen. "OCRS*: An Interactive Object-based Image Clustering and Retrieval System*". Multimedia Tools Application, 35: 71-89, 2007.

20. S. Wang, F. Jing, J. He, Q. Du and L. Zhang. "*IGroup: Presenting Web Image Search Results in Semantic Clusters*". CHI'07. San Jose, California, USA, April 28-May 3, 2007.

21. R.H. Leuken, L. Garcia, X. Olivares and R. Zwol. "*Visual Diversification of Image Search Results*". WWW 2009. Madrid, Spain, April 20-24, 2009.

22. C. Wang, F. Jing, L. Zhang and H. Zhang. "*Scalable Search-based Image Annotation*". Multimedia Systems, 14: 205-220, 2008.

23. B. Tahayna, M. Belkhatir and Y. Wang. "*Clustering of Retrieved Images by Integrating Perceptual Signal Features within Keyword-based Image Search Engines*". P. Muneesawang et al. (Eds.): PCM 2009, LNCS 5879*:* 956–961, 2009.

24. A.W.M. Smeulders, M. Worring, S. Santini A. Gupta and R. Jain. "*Content-based image retrieval at the end of the early years*". IEEE Transactions on Pattern Analysis and Machine Intelligent, 22(12): 1349–1380, 2000.

25. Y. Liu, D. Zhang, G. Lu and W. Ma. "*Region-based image retrieval with high-level semantic color names*". 11th International Multimedia Modelling Conference. Melbourne, Australia, 180-187, 2005.

26. P.C. Lim, "*A Generalized Framework for Mapping Low-level Visual Features to High-level Semantic Features*", Master's thesis, Universiti Malaysia Sarawak, 2008.

27. P.C. Lim, N. Kulathuramaiyer, F. Abang & Y.C. Wang, "*Classification of Butterfly Wing Images*". International Conference on Intelligent Systems. Kuala Lumpur, Malaysia, December 1-3, 2005.

28. Y. Chen, J.Z. Wang & R. Krovetz, "*CLUE: Cluster-Based Image Retrieval of Images by Unsupervised Learning*". IEEE Transactions on Image Processing, 14(8): 1187-1201, 2005.

29. Y. Gao, J. Fan, H. Luo and S. Satoh, "*A Novel Approach for Filtering Junk Images from Google Search Results*". S. Satoh, F. Nack, and M. Etoh (Eds.): MMM 2008, LNCS 4903*:* 1–12, 2008.

30. P.A. Moellic, J.E. Haugeard and G. Pitel, "*Image Clustering based on a Shared Nearest Neighbours Approach for Tagged Collections*". CIVR'08. Niagara Falls, Ontario, Canada, July 7-9, 2008.
31. L. Cao, J. Luo & T.S. Huang, "*Annotating Photo Collections by Label Propagation according to Multiple Similarity Cues*". MM'08. Vancouver, British Columbia, Canada, October 26-31, 2008.
32. M. Ferecatu, N. Boujemaa & M. Crucianu, "*Semantic Interactive Image Retrieval Combining Visual and Conceptual Content Description*". Multimedia Systems, 13: 309-322, 2007.
33. W. Lu, R. Xue, H. Li & J. Wang, "*A Strategy of Semantic Information Extraction for Web Image*", S. Wang, L. Yu, F. Wen, S. He, Y. Fang & K.K. Lai (Eds.), Business Intelligence: Artificial Intelligence in Business, Industry & Engineering, 2nd International Conference on Business Intelligence and Financial Engineering. Beijing, China, July 24-26, 2009.
34. S. Yang, S. Kim and Y.M. Ro. "*Semantic Home Photo Categorization*", IEEE Transactions on Circuits and Systems for Video Technology, 17(3): 324-335, 2007.
35. D.N.F. Awang Iskandar. "*Image Retrieval using Automatic Region Tagging*", PhD dissemination, School of Computer Science and Information Technology, Royal Melbourne Institute of Technology University, March 2008.
36. D.N.F. Awang Iskandar, J.A. Thom and S.M.M. Tahaghoghi, "*Content-based Image Retrieval Using Image Regions as Query Examples*". A. Fekete and X. Lin (Eds.): 19th Australasian Database Conference (ADC2008). CRPIT, 75: 39-75. Wollongong, NSW, Australia, 2008.
37. M. Park and K. Ramamohanarao. "*Automatic extraction of semantic concepts in medical images*". IEEE International Conference on Image Processing. Singapore, October 24-27, 2004.
38. Y. Shin, Y. Kim and E.Y. Kim. "*Automatic textile image annotation by predicting emotional concepts from visual features*". Image and Vision Computing, 28(3): 526-537, 2010.