A Knowledge Framework for Histogram-based Image Retrieval

R.C.F. Wong and C.H.C. Leung

Abstract—**With the rapid increase of the volume of digital image collections, effective concept-based image retrieval has been became an important research issue. The effectiveness of image retrieval depends on meaningful indexing. In this paper, we propose an extension of image indexing models which utilizes a knowledge framework with adaptive segmentation of HSV color space. Experiments have been performed to evaluate the performance of this approach, and the results show that a good degree of retrieval effectiveness may be attained.**

Keywords: Image Indexing, Semantic Image Retrieval, HSV color space

1 Introduction and related work

With the rapid increase of the volume of digital image collections, image retrieval has been become one of the most research area. As the number of images available in online repositories is growing dramatically, exploring the frontier between image and language is an interesting and challenging task.

Most important elements in a retrieval system are the features used to express an image and image features can be divided between two main categories: concept-based image retrieval and content-based image retrieval. The former focuses on retrieval by objects and high-level concepts, while the latter focuses on the low-level visual features of the image.

Content-based image retrieval, the problem of searching large image repositories according to image content, has been the subject of a significant amount of research in the last decade. These methods aim at accessing the knowledge embedded in images by extract low-level visual features and capturing image similarity by relying on some specific characteristic of images. Typically, these models are based on color, texture and shape [25, 22, 9, 13, 33, 10, 30, 4, 6]. Some studies [17, 8, 29, 15, 11, 23] of image segmentation using image partitions, sign detection, region segmentation techniques while some studies rely on compute general similarity between images based on statistical image properties [18, 3, 27, 2, 26, 1, 21]. Researches on semantic retrieval of the image database [22, 9, 13, 33, 10, 30, 14] combined with a region based image decomposition is used, which aims to extract semantic properties of images based on spatial distribution of color and texture properties. The advantage of content-based image retrieval methods are they do not incur any indexing cost as they can be extracted by automatic algorithms.

Concept-based image retrieval is to create a set of metadata to describe the images content, namely, concept indexing. Some studies [20, 11] include users in a search loop with a relevance feedback mechanism to adapt the search parameters based on user feedback. Some researches [5] focuses on implicit image indexing which involve an implicit and, in consequence, augments the original indexes with additional concepts that are related to the query. With the advent of Semantic Web technology, knowledge is playing a key role as the core element of knowledge representation architecture. Some effort [12, 24, 28, 7, 19] has been made for image retrieval using Semantic Web techniques.

We propose an integrated framework for image retrieval based on generative modelling approaches. In [31, 32], a semantic indexing technique named Automatic Semantic Annotation (ASA) approach is developed which is based on the use of image parametric dimensions and metadata. Using decision trees and rule induction, a rule-based approach to formulate explicit annotations for images fully automatically is developed, so that, semantic query such as "night scene of Arch of Triumph in Paris in winter" can be answered and indexed purely by machine. In this paper, we propose an extension of such image indexing models by using knowledge-based expansion and contextual feature-based index expansion. Experimental evidence on more than 100,000 web images and over 990,000 tags shows that semantically meaningful retrieval are inferred and it is able to deliver highly competent performance.

2 A Knowledge Framework for Image Retrieval

Our approach provides operations to perform image retrieval with knowledge meaningfully augmented. It aims in introducing concept augmentation into the image retrieval problem and using the sub-objects as surrogate terms for general queries is to improve the precision since, in certain applications, the presence of particular objects in an image often implies the occurrence of other objects.

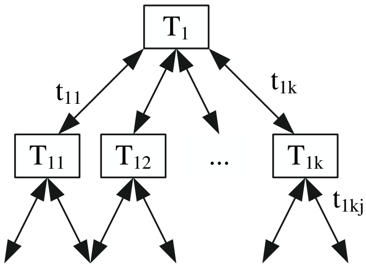

Figure 1: Hierarchical expansion

The application of such inferences will allow the concept of an image to be automatically expanded.

Aggregation hierarchical expansion is a particularly useful technique, which relates to the aggregation hierarchy of sub-objects that constitute an object. These can be classified into two categories, concrete and abstract hierarchical expansion. Concrete hierarchical expansion is the relevant objects are well-defined (Fig. 1). For example, an concept "wedding" expanded to bridge, groom, flower, wedding cake. Abstract hierarchical expansion is the objects are not concretely defined. Although "conflict" is not a definite visual object, it contains certain common characteristics.

In order to perform direct extraction of high-level semantic content automatically, we establish associations between low-level features with high-level concepts, and such associations take the following forms. In the contextual feature-based expansion, the presence of certain low-level features F may suggest a finite number of m object possibilities. Sometimes, a combination of basic features may be used to infer the presence of high-level concepts for inclusion in the semantic index.

The presence of certain basic features alone may not be sufficient to infer the presence of specific objects, but such features if augmented by additional information may lead to meaningful inferences. When a particularly context is known, a concept may be indexed more precisely. Such contextual information will typically be provided through knowledge-based expansion, which may lead to the creation of a new index term, or a revision of the score of an existing index term. An iterative feedback loop will be risen where the determination of new objects will lead to new meaningful feature-object combinations, where further objects may be determined.

The quality assessment of the machine-inferred boundaries between parts of the depicted scenes is based on the precision. Thus, our system is measured quantitatively (given in a later section) in order to compute the effectiveness of our approach. The reliability of a given annotation will given rise to a numerical measure, which signifies how good the annotation is. For annotations with a low value, this would mean that the annotation is not very reliable, or in extreme cases, what is being annotated is absent from an image. A high value indicates that the chance of finding the corresponding object or content in the given image is high. In addition, apart from measuring the likelihood of whether something is present or not, it can be used to indicate the importance of an object in the image. For example, an object which is very prominently present in the foreground of an image would have a much higher value than an object of small size in the remote background. Hence the annotation measure is used to signify two aspects: the likelihood of finding the object in the given image and the prominence of the object in the given image.

A color space is defined as a model for representing color in terms of intensity values.

Many histogram distances have been used to define the similarity of two color histogram representations. The Bhattacharya Distance (BD), Chi-squared Distance (CD) and Euclidean Distance (ED) are also used and, from [25, 16], the definition are listed in Equation 1, 2 and 3 respectively.

$$\Delta_1(x,y) = \sqrt{1 - \sum_{i=1}^{r} \sqrt{x_i y_i(k)}} \qquad (1)$$

$$\Delta_2(x,y) = \sum_{i=1}^{r} \frac{(y_i - x_i(k))^2}{(y_i + x_i(k))} \qquad (2)$$

$$\Delta_3(x,y) = \sqrt{\sum_{i=1}^{r} (x_i - y_i)^2} \qquad (3)$$

where k is the center of the image region, r is the number of bins in the distribution, and x_i and x_i are the weighted histograms of the model and candidate respectively. The application of these distance measures in response to image queries will be evaluated and compared.

3 Evaluation of System Performance

The main purpose in introducing knowledge-based expansion into the image retrieval problem and using the sub-objects as surrogate terms for general queries is to improve the precision in the image sets. In this paper, we mainly focus on the knowledge-based expansion and contextual feature-based index expansion, specially focus on histogram search which characterizes an image by its color distribution. The knowledge concept are organized and used to build the basic content index within a relational database where it is designed for maximum query

ISBN: 978-988-17012-2-0

effectiveness by distributing the semantic elements across different relations. A further concept is built on top of these relations to support rapid discovery.

Our system is evaluated quantitatively in order to compute the effectiveness of our approach. The quality assessment of the machine-inferred boundaries between parts of the depicted scenes is based on the precision. A set of standard evaluation queries are used for experimentation. Comparison is made between base-level indexing and the expanded level indexing, and the widely accepted measures of retrieval performance of precision and recall are used to assess system performance. To numerically assess the accuracy and effectiveness of our annotation approach, we have retrieved 103,521 sets of images with 991,074 associated tags from flickr.com which are a popular photo sharing web site and online community platform offering a fairly comprehensive web-service API that allows developers to create applications that can perform almost any function on images. In our evaluation, we decide that a relevant image must include a representation of the category in such a manner that a human should be able to immediately associate it with the assessed concept.

In [31, 32] by using decision trees and rule induction, a rule-based approach to formulate explicit annotations for images fully automatically has been developed. In relation to image acquisition, many images may be broken down to few basic scenes, such as nature and wildlife, portrait, landscape and sports. In the case of aggregation hierarchical expansion, we decided to test our system using the aggregation hierarchy of basic categories "night scenes" and extend the image hierarchy to find a subscene "night scene of downtown", "downtown" can be expanded to "business district", "commercial district", "city center" and "city district", while "city district" can be expanded to "road", "building", "architecture", "highway" and "hotel".

To extend the original approach, firstly, we annotate night scenes based on the prior rule-based approach to extract 422 out of 103,527 images. We also gather 1108 tags associated with those images and totally 417 unique terms are formed. We list the top 40 out of 417 unique terms list in Fig. 2. We present the results of the evaluation in Table 1.

To establish associations between low-level features with high-level concepts, associating basic features with semantic concepts may be applied to arbitrary images for inclusion in the semantic index. Methods using color tone is robust with respect to object movement, rotation and to changes like distortion within an image and may be implemented easily. Histogram is mainly used for color tone information, particularly in the areas of feature detection and feature extraction, to refer to algorithms described in Equation 1, 2 and 3. Here, we adapt color tone histogram algorithms [25, 16] to extract high-level concepts from low-level features.

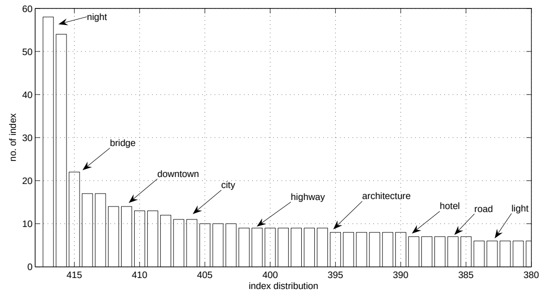

Figure 2: Index distribution associated with night scene images

Table 1: Experimental results on aggregation hierarchical expansion

Semantic content	Precision rate
highway	33.33%
architecture	60.00%
road	66.67%
hotel	75.00%
city	100.0%
downtown	100.00%

We performed experiments to show the good optimality and convergence of our approach. Firstly, we randomly select one "downtown" image from the test set and carried out evaluation by comparing the original Automatic Semantic Annotation (ASA) approach with our approach which combines the original ASA approach using adaptive annotation of HSV color space and distance algorithms (see Fig. 3) and the use of human tags.

In Table 2, experimental results show that tags by human deliver excellent precision rate with 100% precision but this tagging approach relies heavily on human involvement. Significantly better results can be obtained

Table 2: Experimental results on contextual feature-based index expansion

Approach	Precision rate
Original ASA approach	53.20%
ASA approach with BD	67.27%
ASA approach with CD	71.44%
ASA approach with ED	79.90%
ASA approach with edge detection	87.03%
Tags by human	100.00%

ISBN: 978-988-17012-2-0

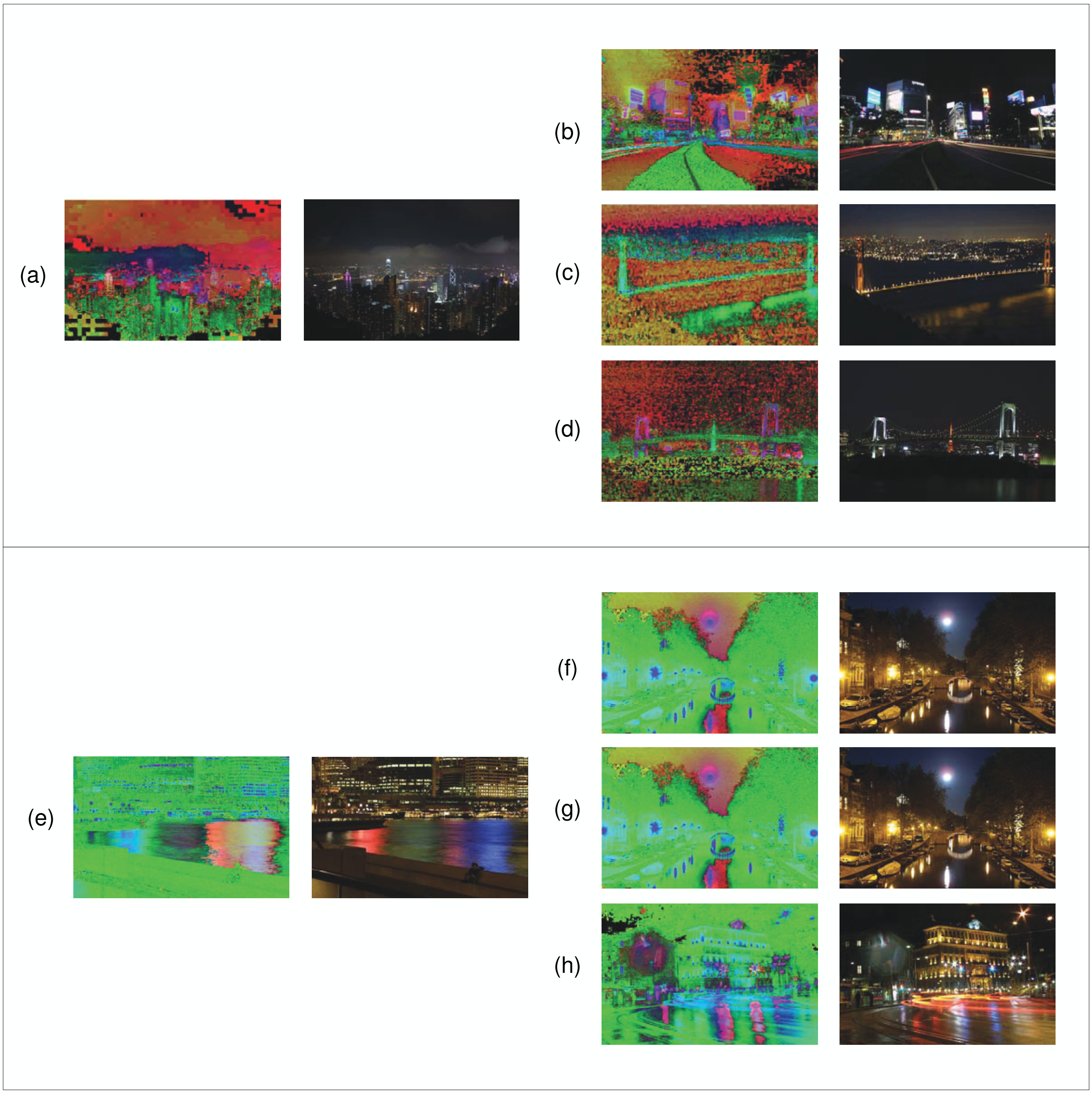

Figure 3: Example of experimental results of nearest histogram distance in RGB and HSV color space: rows (a)(e) query images, rows (b)(f) Bhattacharya Distance(BD), rows (c)(g) Chi-squared Distance(CD), rows (d)(h) Euclidean Distance(ED)

by ASA Approach combining HSV color space histogram distances where the precision rate grows to 67.2%(BD), 71.4%(CD) and 77.9%(ED). Obviously, compared to annotation without the contextual feature-based index expansion enabled, the performance is around 52.8%. From the joint application of these, we can formulate semantic annotations for specific image fully automatically and index images purely by machine without any human involvement.

4 Conclusion

In this paper, a knowledge of image retrieval together with contextual context knowledge augmentation is combined. Our method combines the advantages of original ASA approach and contextual feature-based expansion while preserving the necessary image and knowledge coherence. Our system is evaluated quantitatively, and experimental results indicate that this approach is able to deliver highly competent performance.

[1] G. Amato and C. Meghini. Combining features for image retrieval by concept lattice querying and navigation. *International Conference on Image Analysis and Processing Workshops*, 0:107–112, 2007.

[2] J. Amores, N. Sebe, and P. Radeva. Context-based object-class recognition and retrieval by generalised correlograms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(10):1818–1833, October 2007.

[3] V. Athitsos, J. Alon, S. Sclaroff, and G. Kollios. Boostmap: A method for efficient approximate similarity rankings. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 02:268–275, 2004.

[4] I. Azzam, A. G. Charlapally, C. H. C. Leung, and J. F. Horwood. Content-based image indexing and retrieval with xml representations. *Proceedings of the International Symposium on Intelligent Multimedia, Video and Speech Processing, Hong Kong*, pages 181–185, 2004.

[5] I. A. Azzam, C. H. C. Leung, and J. F. Horwood. Implicit concept-based image indexing and retrieval. In *Proceedings of the IEEE International Conference on Multi-media Modeling*, pages 354–359, Brisbane, Australia, January 2004.

[6] I. A. Azzam, C. H. C. Leung, and J. F. Horwood. A fuzzy expert system for concept-based image indexing and retrieval. *International Conference on Multimedia Modeling Conference*, 0:452–457, 2005.

[7] K. Barnard, P. Duygulu, N. de Freitas, D. Forsyth, D. Blei, and M. Jordan. Matching words and pictures. *Journal of Machine Learning Research*, 3:1107–1135, 2003.

[8] Y. Chen, J. Z. Wang, and R. Krovetz. Content-based image retrieval by clustering. In *MIR'03: Proceedings of the 5th ACM SIGMM international workshop on Multimedia information retrieval*, pages 193–200, New York, NY, USA, 2003. ACM.

[9] J. S. Cho and J. Choi. Contour-based partial object recognition using symmetry in image databases. In *SAC'05: Proceedings of the 2005 ACM symposium on Applied computing*, pages 1190–1194, New York, NY, USA, 2005. ACM Press.

[10] D. Cremers, M. Rousson, and R. Deriche. A review of statistical approaches to level set segmentation: Integrating color, texture, motion and shape. *International Journal of Computer Vision*, 72(2):195–215, 2007.

[11] R. Datta, J. Li, and J. Z. Wang. Content-based image retrieval: approaches and trends of the new age. In *MIR'05: Proceedings of the 7th ACM SIGMM international workshop on Multimedia information retrieval*, pages 253–262, New York, NY, USA, 2005. ACM.

[12] L. Fan and B. Li. A hybrid model of image retrieval based on ontology technology and probabilistic ranking. *Web Intelligence*, 0:477–480, 2006.

[13] J. Gausemeier, J. Fruend, C. Matysczok, B. Bruederlin, and D. Beier. Development of a real time image based object recognition method for mobile ar-devices. In *International conference on Computer graphics, virtual Reality, visualisation and interaction in Africa*, pages 133–139, New York, NY, USA, 2003. ACM Press.

[14] K. Hornsby. Retrieving event-based semantics from images. *International Symposium on Multimedia Software Engineering*, 00:529–536, 2004.

[15] M. Jian, J. Dong, and R. Tang. Combining color, texture and region with objects of user's interest for content-based image retrieval. *International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing*, 01:764–769, 2007.

[16] A. Kale, K. Kwan, and C. Jaynes. Epipolar contrained user pushbutton selection in projected interfaces. *Computer Vision and Pattern Recognition Workshop*, 10:156–164, 2004.

[17] R. Krishnapuram, S. Medasani, S. H. Jung, Y. S. Choi, and R. Balasubramaniam. Content-based image retrieval based on a fuzzy approach. *IEEE Transactions on Knowledge and Data Engineering*, 16(10):1185–1199, 2004.

[18] J. Li and J. Z. Wang. Automatic linguistic indexing of pictures by a statistical modeling approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(9):1075–1088, 2003.

[19] H. Lieberman and H. Liu. Adaptive linking between text and photos using common sense reasoning. *Adaptive Hypermedia and Adaptive Web-Based Systems, Second International Conference, AH 2002, Malaga, Spain*, pages 2–11, May 2002.

[20] D. Liu and T. Chen. Content-free image retrieval using bayesian product rule. *IEEE International Conference on Multimedia and Expo*, 0:89–92, 2006.

[21] A. P. Natsev, A. Haubold, J. Tešić, L. Xie, and R. Yan. Semantic concept-based query expansion and re-ranking for multimedia retrieval. In *International conference on Multimedia*, pages 991–1000, New York, NY, USA, 2007. ACM.

[22] R. Pawlicki, I. Kókai, J. Finger, R. Smith, and T. Vetter. Navigating in a shape space of registered models. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1552–1559, 2007.

[23] A. Perina, M. Cristani, and V. Murino. Natural scenes categorization by hierarchical extraction of typicality patterns. *International Conference on Image Analysis and Processing*, pages 801–806, 2007.

[24] A. Popescu, G. Grefenstette, and P. A. Moellic. Using semantic commonsense resources in image retrieval. In *International Workshop on Semantic Media Adaptation and Personalization*, pages 31–36, Washington, DC, USA, 2006. IEEE Computer Society.

[25] M. Riaz, G. Kang, Y. Kim, S. Pan, and J. Park. Efficient image retrieval using adaptive segmentation of hsv color space. *Computational Science and its Applications, International Conference*, 0:491–496, 2008.

[26] K. Stevenson and C. H. C. Leung. Comparative evaluation of web image search engines for multimedia applications. *IEEE International Conference on Multimedia and Expo*, 0:4 pp., 2005.

[27] Y. Sun, S. Shimada, and M. Morimoto. Visual pattern discovery using web images. *MIR'06: Proceedings of the 8th ACM international workshop on Multimedia information retrieval*, pages 127–136, 2006.

[28] A. M. Tam and C. H. C. Leung. Semantic content retrieval and structured annotation: Beyond keywords. *ISO/IEC JTC1/SC29/WG11 MPEG00/M5738, Noordwijkerhout, Netherlands*, March 2000.

[29] N. Vasconcelos. From pixels to semantic spaces: Advances in content-based image retrieval. *IEEE Computer*, 40(7):20–26, 2007.

[30] J. Vogel, A. Schwaninger, C. Wallraven, and H. H. Bülthoff. Categorization of natural scenes: Local versus global information and the role of color. *ACM Transactions on Applied Perception*, 4(3):19, 2007.

[31] R. C. F. Wong and C. H. C. Leung. Automatic semantic annotation of real-world web images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(11):1933–1944, November 2008.

[32] R. C. F. Wong and C. H. C. Leung. Knowledge-based expansion for image indexing. In *International Computer Symposium*, volume 1, pages 161–165, November 2008.

[33] T. Zöller and J. M. Buhmann. Robust image segmentation using resampling and shape constraints. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(7):1147–1164, July 2007.