

Content-based Image Retrieval Using Image Regions as Query Examples

D. N. F. Awang Iskandar James A. Thom S. M. M. Tahaghoghi

School of Computer Science and Information Technology, RMIT University
Melbourne, Australia

Email: {dayang.awgiskandar, james.thom, seyed.tahaghoghi}@rmit.edu.au

Abstract

A common approach to content-based image retrieval is to use example images as queries; images in the collection that have low-level features similar to the query examples are returned in response to the query. In this paper, we explore the use of image regions as query examples. We compare the retrieval effectiveness of using whole images, single regions, and multiple regions as examples. We also compare two approaches for combining shape features: an equal-weight linear combination, and classification using machine learning algorithms. We show that using image regions as query examples leads to higher effectiveness than using whole images, and that an equal-weight linear combination of shape features is simpler and at least as effective as using a machine learning algorithm.

Keywords: CBIR, Query-by-example, machine learning, linear combination.

1 Introduction

Search for images has been explored using text and image content. In content-based image retrieval (CBIR), image content is frequently represented through image features. Commonly used features include colour, texture, or shape descriptors for objects found within an image (Vasconcelos & Kunt 2001, Lew et al. 2006, Gevers & Smeulders 2000, Jain & Vailaya 1996). Searching using a combination of more than one image feature — for example region and colour — improves retrieval effectiveness (Jain & Vailaya 1996). In this paper, we demonstrate work done in retrieving similar regions using colour and shape features.

To help bridge the semantic gap, we need to capture a user's information need by allowing the user to express their query to the CBIR system. This paradigm is known as query-by-example (QBE), where the query is expressed as one or more example images (Smeulders et al. 2000). Presenting an image as the query example limits the user to express the exact information need. Therefore, a study by Enser (1993) reveals the need for expressing region or regions of interest as the query into CBIR systems. In this work, we distinguish between query-by-image-example (QBIE) as using an image or images as the query, and query-by-region-example (QBRE) as using region or regions of interest as the query.

Copyright ©2008, Australian Computer Society, Inc. This paper appeared at the Nineteenth Australasian Database Conference (ADC2008), Wollongong, Australia, January 2008. Conferences in Research and Practice in Information Technology (CRPIT), Vol. 75, Alan Fekete and Xuemin Lin, Ed. Reproduction for academic, not-for profit purposes permitted provided this text is included.

Data mining using machine learning algorithms is another technique towards improving CBIR systems, whereby images can be categorised and clustered based on their features. For CBIR, the algorithms can be trained using examples of image features in order to then identify images that are relevant to a query image.

The motivation behind this work is simple. Let us consider the situation where a user only has limited number of examples to query with. Probably the most effective approach is to extract the low-level image features from these limited examples and combine them for retrieving similar images in a collection. In this paper we present a comparison of using an equal-weight linear combination to combine the features that form a region of interest with the classification of several machine learning algorithms.

We also experiment with the retrieval effectiveness of using the whole image, a single-region and multiple-regions from the query example. When we have multiple images or regions, we use a combining function to produce the final ranked list. In this paper, the results that we present use MAXIMUM similarity as the combining function (which is the minimum distance). This was found to be the most effective combining function by Tahaghoghi et al. (2002).

To reduce the problem of object segmentation, we test our approach on a domain where regions are easily separated: a collection of comic strips. In this domain, objects and characters comprise multiple regions of approximately uniform colour. The objects in the comics have relatively consistent size and orientation, guiding our choice of the following region-based and contour-based shape features: the region area; the mean grey level value of the pixels in the region; circularity; and shape boundary. We did not use any texture features in this work since colour is a much more prominent feature in the comic image collection.

The remainder of this paper is organised as follows. In Section 2, we present the background on existing CBIR systems and shape descriptors that we use. In Section 3, we describe the testbed and the queries used in conducting our experiments. In Section 4, we present the QBIE retrieval process. In Section 5, we explain the single-region querying and extraction of the shape features. We explain multiple-region querying in Section 6. In Section 7 we present and discuss experimental results, and conclude in Section 8 with a discussion of our findings and suggestions for future work.

2 Background

Colour, texture and shape are the fundamental elements of object appearance used by humans to differentiate objects. Much research exists on the optimal choice of features and feature representations (Latecki et al. 2002). In this work we focus on combining

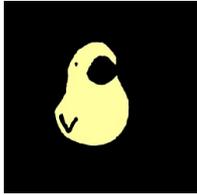
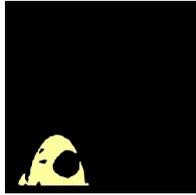
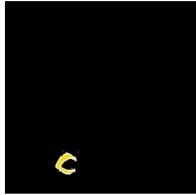
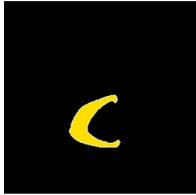
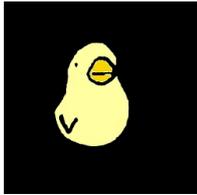
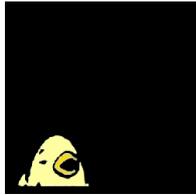
Query Type	Body Part	Colour	Example 1	Example 2	Example 3
Whole Image					
Region 1	Body	Light Yellow			
Region 2	Beak	Dark Yellow			
Multiple-region					

Figure 1: Query specifications for character *Oliver*. This figure is best viewed in colour.

colour and shape features.

Examples of existing CBIR systems that use single-image QBE include Tiltomo,¹ CIRES (Iqbal & Aggarwal 2002), our,² SIMBA³ and SIMPLIcity⁴.

The GNU Image Finding Tool (GIFT)⁵ supports single- and multiple-image QBIE (Squire et al. 2000). Tahaghoghi et al. (2001, 2002) have shown that using multiple query images generally improves retrieval effectiveness.

The need to describe shapes mathematically leads to the two broad methods for shape representation and description — region- and contour-based (Zhang & Lu 2004).

In region-based methods, the features are extracted from the whole region. Such region-based features include area, length and angle of major and minor axes, and moments.

Contour-based methods represent shape by a coarse discrete sampling of its perimeter. Contour-based shape descriptors include perimeter, Hausdorff distance, shape signatures, Fourier descriptors, wavelet descriptors, scale space, autoregressive, elastic matching and shape context (Zhang & Lu 2004).

Region-based shape descriptors are often used to

discriminate regions with large differences (Zhang & Lu 2004), and are usually combined with contour-based features. Shape matching is performed by comparing the region-based features using vector space distance measures, and by point-to-point comparison of contour-based features.

In this work we combine both the region- and contour-based method of shape feature extraction for a region of interest — area, mean, circularity and boundary. Area is the total number of pixels inside a region. Mean is the average grey value within the region. This is the sum of the grey values of all the pixels in the shape divided by the number of pixels.

Circularity (also known as compactness) is a shape feature that is both region- and contour-based. It is calculated using the formula (Costa & Cesar Jr. 2000):

$$circularity = 4\pi \left(\frac{area}{perimeter^2} \right)$$

A circularity value of 1.0 indicates a perfect circle — the most compact shape — while smaller values indicate increasingly elongated shapes.

Blobworld (Carson et al. 1999) supports single-shape — rather than single-image — queries for images containing a specific region of interest. The shape features are represented by area, eccentricity and orientation. The low-level features used for querying are colour, texture and shape location. This

¹ <http://www.tiltomo.com>

² <http://our.imgseek.net>

³ <http://simba.informatik.uni-freiburg.de>

⁴ <http://wang14.ist.psu.edu>

⁵ <http://www.gnu.org/software/gift>

region-based retrieval approach is not comparable to the method proposed here since we do not use texture and shape location as part of the query.

3 The Collection

To compare the effectiveness of the equal-weight linear combination and the machine learning algorithms, we have created an image collection that consists of comic strip panels from the Goats comic.⁶ These include extended information that describes every panel of the comic strip. The description assists us in performing the relevance judgements on our retrieval results.

The collection consists of 452 coloured strips, each containing one to five panels. Dividing the strips into panels gives us a total of 1440 panels. We tested the retrieval effectiveness using 1115 regions extracted from 202 panels. The remaining panels are reserved as a validation set for our future work. From this point onwards, we refer to the individual panels as images.

For five concepts representing the main characters in the Goats comic strips — *Bob* (an alien), *Diablo* (a chicken), *Fineas* (a fish), *Jon* (a person) and *Oliver* (a chick), we randomly chose three images that can be used as query examples. Since one of the selected query images includes two characters, our query set has a total of fourteen images. The three query images for the character *Oliver* are depicted in Figure 1.

Within each query image example, we identified and extracted two regions (Region 1 and Region 2) corresponding to a particular character. For example, we identified the body of *Oliver* as the Region 1 query and the beak as the Region 2 query. Detailed explanations of the queries for other characters and objects appears elsewhere (Awang Iskandar et al. 2006). We used the thirty extracted regions as follows:

1. QBRE for the single and multiple-regions; and
2. Training data for the machine learning algorithms.

Apart from serving as query regions, the query images are used as a training set to study and analyse the best weight combination of shape features for retrieving the region of interest using a single-query region. However, we found that equal weight gave us the best result. Due to this finding, we used the equal-weight linear combination for combining the shape features.

4 Whole Image Example

As the baseline method of retrieving the concepts in an image, we used the GNU Image Finding Tool (GIFT) to retrieve similar images using the colour features extracted from the whole image. While this CBIR system supports both colour and texture features, we limit our experiments to local and global HSV colour features. GIFT ranks images in the collection by decreasing similarity to the query example.

To retrieve images containing a particular concept, we present the images containing the concept as the query. For instance, to execute the queries for *Oliver*, we present the whole image queries depicted in Figure 1 as the examples. We conducted an initial experiment where we compared the retrieval effectiveness of using two or three image examples; we found that using two example images produces better results than three examples. Nevertheless, we chose to use three

example images so that there is more training data for the machine learning algorithms.

Since we have three query images for each concept, we will have three ranked lists. To reduce these to a single list, we adopted the multiple image example combining functions employed by Tahaghoghi et al. (2002), with the slight difference that we use similarity, rather than distance, values. Thus, the combining functions we use are:

- SUM, as the average similarity of the candidate image and the query images;
- MAXIMUM similarity (minimum distance) between the query images and the candidate image; and
- MINIMUM similarity (maximum distance) between the query images and the candidate image.

The final ranked list contains the images arranged by decreasing similarity values.

5 Single-region Example

For a single-region query, we execute queries for each region that was identified for the Region 1 and Region 2 query examples. Each character has three query example regions for Region 1, and another three query example regions for Region 2.

To retrieve the image regions, we first extract all the regions and their shape features from the images in the collection. This involves obtaining the area, mean, circularity and boundary features for all regions in the collection.

We use the colour feature in an initial search of all frames to find any regions that might match the query character. Retrieval based on colour histograms has been shown to outperform retrieval based on shape alone, both in terms of efficiency and robustness (Jain & Vailaya 1996).

To identify the candidate regions from the collection, we implemented a plug-in for the GNU Image Manipulation Program (GIMP). The plug-in selects image regions based on the similarity of their colour to that of the query shape representing a particular comic character. Each of the red, green, and blue colour components has a value in the range [0, 255]; during similarity matching, we allow a variation of up to 20 for each component value.

We then apply the ImageJ program⁷ particle analyser plug-in to the candidate regions to acquire the area, mean value and circularity. To obtain the region's boundary feature, we use the shape context algorithm proposed by Belongie et al. (2000).

Figure 2 depicts the similarity between three candidate regions, CR_n , with the query region, QR . Each candidate region's colour matches the colour of the query region. The similarity of the candidate region shape features, $feature(CR)$, to the query region shape features, $feature(QR)$, is calculated as the complement of the Euclidean distance:

$$Sim_{feature} = 1 - \left(\sqrt{(feature(QR) - feature(CR))^2} \right)$$

where $feature \in \{area, mean, circularity, boundary\}$.

The similarity of a candidate image to the query is determined to be the highest similarity value of any region in that image. The final ranked list is in descending order of similarity values.

⁶ <http://www.goats.com>

⁷ <http://rsb.info.nih.gov/ij>

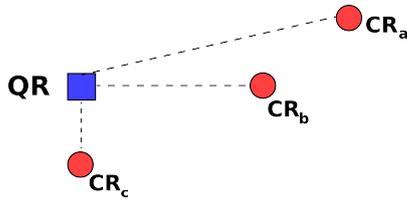


Figure 2: Single-region query points in a two-dimensional query space. The nearest candidate region 1 (CR_c) has the highest similarity value for the query region (QR), followed by CR_b , and lastly CR_a .

To combine all the shape features that form a region, we experiment with two techniques — an equal-weight linear combination and with machine learning algorithms. We analyse the retrieval effectiveness of these techniques in Section 7.

5.1 Equal-weight Linear Combination

The similarity values of each shape feature are combined using a simple linear combination with equal weights to obtain the overall similarity for a candidate region (CR) when querying using a single-region.

$$Sim_{CR} = 0.25 \times (Sim_{area} + Sim_{mean} + Sim_{circularity} + Sim_{boundary})$$

A list of candidate images is then presented to the user, with the images ranked by decreasing similarity of the candidate regions that they contain. When an image contains several candidate regions, its similarity is determined by the region that is the most similar to the identified region in any one of the three query examples.

5.2 Machine Learning Algorithms

Machine learning algorithms need to be trained on many examples — commonly two-thirds of data in a collection (Witten et al. 1999), and so we hypothesise that our simple approach of combining the shape features and retrieving similar regions of interest will perform as well as machine learning when using two or three examples.

To compare the retrieval effectiveness of using equal-weight linear combination and machine learning algorithms, we have experimented with twelve machine learning algorithms provided by the WEKA toolkit (Witten et al. 1999); we explain the machine learning parameters that we use in this section.

We train the machine learning algorithms with fifteen region examples for Region 1 and Region 2 respectively. Each concept has three positive examples and twelve negative examples.

5.2.1 Bayesian Classifiers

Under this classifier, we experimented with four functions:

Bayesian Networks (BN) pre-discretise numeric attribute values and replace missing attributes. Learning the Bayesian network involves two phases — first learn a network structure by using a searching algorithm, then learn the probability tables (Remco 2004).

To search through the network we used the K2 algorithm, that has been proven to be a valuable search method (Cooper & Herskovits 1992). We used a simple estimator to learn the conditional probability tables of the Bayes Network.

Naïve Bayesian (NB) implements the Bayes’s rule of conditional probability using the kernel density estimator. Naïve Bayes classification has the ability to learn using a limited amount of training data for each possible combination of the variables by assuming that the effect of a variable value on a given class is independent of the values of other variables (Lewis 1998) — this assumption is called class conditional independence, and is made to simplify the computation.

A variable attribute is either categorical or numeric. Categorical values are discrete, while numerical values can be discrete or continuous. To learn and classify using Naïve Bayesian, the numerical values are discrete as the classification performance tends to be better (Dougherty et al. 1995). We used a supervised discretisation approach to process numeric attributes as it is shown to perform better when compared to the unsupervised discretisation approach.

Naïve Bayesian Updateable (NBU) is an incremental version of *Naïve Bayesian* that processes one instance at a time. This is an implementation of the Flexible Naïve Bayesian algorithm proposed by John & Langley (1995). We use supervised discretisation to process numeric attributes.

Complement Naïve Bayesian (CNB) builds a classifier that modifies the Naïve Bayesian classifier (Fawcett & Mishra 2003) to balance the amount of training examples and the estimate weight for the decision boundary. We used the default smoothing value of 1.0 to avoid zero values.

5.2.2 Decision Trees

Decision tree learning algorithms are a major type of effective learning algorithm, and represent a supervised approach to classification popularised by Quinlan (1993). A decision tree is a simple structure where non-terminal nodes represent tests on one or more attributes and terminal nodes reflect decision outcomes. We experimented with five variations of decision trees:

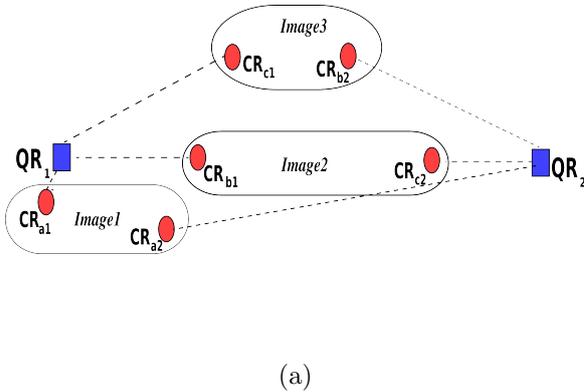
Decision Stump (DS) is a one-level decision tree. It is a weak learner as it is based upon simple binary decisions. Thus; the Decision Stump is normally integrated with boosting and bagging methods (Witten et al. 1999).

J48 is an implementation of Quinlan’s C4.5 decision tree model. We set the confidence factor for pruning to 0.25 where it works reasonably well in most cases⁸ and the minimum number of instances per leaf to 1 since this will create a more specialised tree.

Logistic Model Trees (LMT) are a combination of a tree structure and logistic regression models (Lavrač et al. 2003, Landwehr et al. 2005) to produce a single decision tree. Logistic Model Trees give explicit class probability estimates rather than just a classification.

We set the minimum number of instances at which a node is considered for splitting to the default value of 15.

⁸Decision Trees for Supervised Learning, http://grb.mnsu.edu/grbts/doc/manual/J48_Decision_Trees.html



Rank	SUM	MAXIMUM	MINIMUM
1	Image2	Image1	Image3
2	Image3	Image2	Image2
3	Image1	Image3	Image1

Figure 3: (a) Multiple-region querying points in a two-dimensional query space. (b) The ranked list when combining functions are used to determine the similarity between the candidate regions (CR_{a1} , CR_{b1} , CR_{c1} , CR_{a2} , CR_{b2} , and CR_{c2}) in the images and the query regions (QR_1 and QR_2).

Naïve Bayesian Tree (NBTree) is a fusion of decision tree and Naïve Bayesian. This algorithm creates a decision tree as the general structure, and deploys Naïve Bayes classifiers at the leaves (Kohavi 1996) to overcome the uniform probability distribution problem of decision trees.

Random Forest (RF) is built by bagging ensembles of random trees. The trees are built based on a random number of features at each node and no pruning is performed. Random Forest refers to the procedure of creating a large number of trees and voting for the most popular class among the trees (Breiman 2001). We set the number of trees to be generated to 10, and the random number seed to be used to 1 so that the time taken for the training phase is minimised.

REPTree is a fast decision tree learner. It builds a decision tree using information gain and prunes it using reduced-error pruning with back-fitting. Missing values are dealt with by splitting the corresponding instances into pieces, same as in C4.5 (Witten & Frank 2005). We set no restriction for the maximum tree depth, and set the minimum total weight of the instances in a leaf to 2. We also set the number of folds to 3 and used 1 seed for randomising the data.

5.2.3 Rules

We experimented with only the zeroR rule, which is the simplest rule algorithm, and use it as a baseline to compare with the other machine learning algorithms.

5.2.4 Functions

We experimented with the Simple Logistic (SL) function, which builds linear logistic regression models. We also used the LogitBoost function with simple regression functions as base learners for fitting the logistic models.

6 Multiple-region Example

Almost all studies on shape retrieval focus on retrieving matching shapes using a single shape or region as the query example in the query specification (Carson et al. 1999, Gevers & Smeulders 2000, Belongie et al. 2000). In this work, we use six regions as the query examples to retrieve similar images containing the region that is the most similar to any of the query regions.

To handle multiple-region querying, we combine the query answers of Region 1 and Region 2 that was retrieved for the single-region query. Example of multiple-region queries for *Oliver* is depicted in Figure 1 — where *Oliver* has six query regions. This is the same for other concepts.

Using the machine learning algorithms, we trained with thirty regions, where a concept has six positive examples and twenty-four negative examples.

To explore the effect of using multiple-regions in the query, we applied the combining functions employed by Tahaghoghi et al. (2002) to merge two query regions. Hence, we refine the combining function to suit multiple-region example as:

- SUM of the similarity values of the candidate regions in the image;
- MAXIMUM similarity for any of regions in the image; and
- MINIMUM similarity for any of the regions in the image.

We compute a similarity value for each candidate region in the image, and then apply the combining function to reduce these to a single similarity value for the image.

To illustrate multiple-region querying, Figure 3(a) depicts two query regions — QR_1 and QR_2 and three images that each contain two candidate regions — Image1(CR_{a1}, CR_{a2}), Image2(CR_{b1}, CR_{c2}) and Image3(CR_{c1}, CR_{b2}) from the collection as points in a two-dimensional query space. CR_{n1} and CR_{n2} (where $n = a, b, c, \dots, z$) are the candidate region for QR_1 and QR_2 respectively. Among the three images, which image contains the best match for the combined query of regions QR_1 and QR_2 ? Three simple solutions would be to pick:

- Image1, since one of the candidate region (CR_{a1}) contained in the image is close to QR_1 ;
- Image2, since the candidate regions contained in this image has the highest total similarity from both query regions; and
- Image3, since the candidate regions contained in this image are equally similar to either QR_1 and QR_2 .

In processing a multiple-region query, we obtained the similarity between a candidate region and the query region that was calculated for the single-region query. Then, we applied the MAXIMUM combining function to reduce the multiple similarity values to a

single similarity value. This is performed to all the images in the collection that contain the candidate regions. The user is then presented with a ranked list of images, sorted by decreasing similarity value.

Revisiting Figure 3(a), applying the functions to Image1, Image2 and Image3 would return a ranked list of the best region similarity corresponding to QR_1 and QR_2 as in Figure 3(b).

7 Result Analysis

We evaluate the retrieval effectiveness using the standard recall-precision measure. We analyse the average precision (AP) — the average of precision at each relevant document retrieved and Precision at five (P@5) documents retrieved.

Table 1 presents the retrieval results of whole image, Region 1, Region 2 and multiple-region using equal-weight linear combination and machine learning algorithms. The results presented here are based on the retrieval effectiveness of using the MAXIMUM combining function, where Tahaghoghi et al. (2002) in previous work found that it was the best combining function when one or more images are used as query example.

QBIE-MAX denotes the results obtained using whole image query example. Results obtained using Region 1 and Region 2 as the query region with the linear combination of evidence and MAXIMUM combining function are denoted as QR_1 -QBRE-LCE-MAX and QR_2 -QBRE-LCE-MAX respectively. QR_{1and2} -QBRE-LCE-MAX denote the multiple-region querying using both Region 1 and Region 2.

Region 1 was chosen as the main region that represents the concept. However, when comparing the retrieval effectiveness between Region 1 and Region 2, from Table 1 we observed that Region 2 is better than Region 1 at distinguishing the concept Bob and Jon. This indicates that humans do not necessarily pick the best region, which further motivates us in using multiple regions in the query.

Using multiple-region queries outperformed the retrieval effectiveness of using only the whole image, or Region 1, or Region 2 as examples. This shows that having more examples to represent the concept improves the retrieved results.

Not surprisingly, among the machine learning algorithms, the ZeroR rule performed worst in all types of query examples. This may be due to the fact that the ZeroR rule simply predicts the majority class in the training data (Witten et al. 1999).

Query-by-region-example with the equal-weight linear combination gave better AP compared to the machine learning algorithms. However, in some cases the machine learning algorithms performed better in retrieving the first five images (P@5). Among the machine learning algorithms, we observed that Bayes Network, Random Forest, Logistic Model Tree, Naïve Bayesian, and REPTree achieved better AP compared to other machine learning algorithms.

An example of the first five comic frames retrieved for the character “Oliver” are shown in Figure 4. Visual inspection of the retrieved comic frames shows that the multiple-region queries retrieve more relevant comic frames.

8 Conclusion and Future Work

We have shown that using a single-region query example is better than using the whole image as the query example. However, the multiple-region query examples outperformed the single-region query example and also the whole-image example queries. This

indicates that using more examples of the region of interest improves the retrieved results.

We have also compared using the equal-weight linear combination with various machine learning algorithms, and have shown that an equal-weight linear combination of shape features is simpler and at least as effective as using a machine learning algorithm.

Currently, we are exploring using the equal-weight linear combination approach with the machine learning approaches discussed in this paper to recognise fifteen objects such as door, cigarette, table, chair, screwdriver, and bag.

To further compare our findings, we will use the support vector machine (SVM) technique that is known to be resilient to over-fitting (Burges 1998).

We aim to integrate the multiple-region QBE approach described here to automatically generate the relationship between image regions and the semantic concepts in an image ontology as a top-down approach towards bridging the semantic gap in CBIR (Hare et al. 2006). We plan to develop an image ontology query language as a stepping stone to this end. We also plan to adapt the multiple-region QBE with equal-weight linear combination of shape features to more complex domains such as photographic images.

Acknowledgement

This research was undertaken using facilities supported by the Australian Research Council, an RMIT VR11 grant and a scholarship provided by the Malaysian Ministry of Higher Education. We thank Jonathan Rosenberg, author of the Goats Comic for the permission to publish the comic images.

References

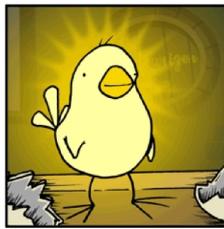
- Awang Iskandar, D. N. F., Thom, J. A. & Tahaghoghi, S. M. M. (2006), Querying comics using multiple shape examples, Technical Report TR-06-5, Royal Melbourne Institute of Technology University.
- Belongie, S., Malik, J. & Puzicha, J. (2000), Shape context: A new descriptor for shape matching and object recognition, *in* ‘Neural Information Processing Systems’, pp. 831–837.
- Breiman, L. (2001), ‘Random forests’, *Machine Learning* **45**(1), 5–32.
- Burges, C. J. C. (1998), ‘A tutorial on support vector machines for pattern recognition’, *Data Mining and Knowledge Discovery* **2**(2), 121–167.
- Carson, C., Thomas, M., Belongie, S., Hellerstein, J. M. & Malik, J. (1999), Blobworld: A system for region-based image indexing and retrieval, *in* ‘Proceedings of the Third International Conference on Visual Information Systems’, Springer.
- Cooper, G. F. & Herskovits, E. (1992), ‘A Bayesian method for the induction of probabilistic networks from data’, *Machine Learning* **9**(4), 309–347.
- Costa, L. & Cesar Jr., R. M. (2000), *Shape Analysis and Classification: Theory and Practice*, CRC Press, Inc., Boca Raton, FL, USA.
- Dougherty, J., Kohavi, R. & Sahami, M. (1995), Supervised and unsupervised discretization of continuous features, *in* ‘Proceedings of the International Conference on Machine Learning’, pp. 194–202.

Table 1: Average Precision and Precision at five results (P@5) for individual concepts. **Bold** values indicate the best retrieval effectiveness. *Italic* values indicate the best retrieval effectiveness for each concept within the various query types.

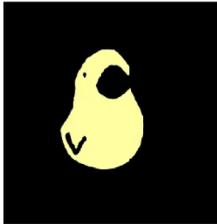
Concepts	Bob		Diablo		Fineas		Jon		Oliver	
	AP	P@5								
Whole Image										
QBIE-MAX	0.4506	0.4	0.2462	0.4	0.4392	0.6	0.2169	0.6	0.2852	0.4
Region 1										
<i>QR₁</i> -QBRE-LCE-MAX	<i>0.7601</i>	1.0	<i>0.7107</i>	1.0	<i>0.7792</i>	1.0	<i>0.3588</i>	<i>0.4</i>	<i>0.4583</i>	0.4
BN	0.6320	0.6	0.4106	0.0	0.7255	0.8	0.2700	0.0	0.4260	0.0
NB	0.0625	0.2	0.2871	0.2	0.7040	0.8	0.1618	0.2	0.4192	0.2
NBU	0.0625	0.2	0.2706	0.2	0.7099	0.8	0.0846	0.2	0.4206	0.0
CNB	0.1528	0.2	0.0000	0.0	0.5966	0.8	0.1072	0.0	0.1073	<i>0.6</i>
DS	0.2208	0.6	0.0000	0.0	0.7108	0.8	0.0000	0.0	0.0000	0.0
J48	0.0000	0.0	0.4106	0.0	0.7108	0.8	0.2700	0.0	0.4300	0.0
LMT	0.0000	0.0	0.4106	0.0	0.7080	1.0	0.2354	0.2	0.4294	0.0
NBTree	0.7227	0.6	0.4430	0.0	0.6829	0.8	0.2765	0.0	0.4300	0.0
RF	0.5581	0.8	0.4361	0.0	0.7134	0.8	0.2770	0.0	0.4360	0.0
REPTree	0.6027	0.8	0.0000	0.0	0.7108	0.8	0.2700	0.0	0.4313	0.0
zeroR	0.0000	0.0	0.0000	0.0	0.7108	0.8	0.0000	0.0	0.0000	0.0
SL	0.3616	0.6	0.3836	0.0	0.7191	1.0	0.1090	0.0	0.4252	0.0
Region 2										
<i>QR₂</i> -QBRE-LCE-MAX	<i>0.7761</i>	<i>0.8</i>	<i>0.6300</i>	1.0	<i>0.7021</i>	1.0	<i>0.5837</i>	0.8	<i>0.1886</i>	<i>0.4</i>
BN	0.4927	0.6	0.2790	0.4	0.5343	0.2	0.3513	0.6	0.0000	0.0
NB	0.6382	<i>0.8</i>	0.1775	0.2	0.5378	0.2	0.3741	0.6	0.0000	0.0
NBU	0.6382	<i>0.8</i>	0.2276	0.2	0.5518	0.2	0.3741	0.6	0.0593	<i>0.4</i>
CNB	0.4927	0.6	0.0000	0.0	0.4781	0.2	0.3018	0.6	0.0000	0.0
DS	0.4930	0.6	0.0000	0.0	0.5343	0.2	0.0000	0.0	0.0000	0.0
J48	0.4927	0.6	0.3861	0.2	0.5343	0.2	0.3513	0.6	0.0000	0.0
LMT	0.6625	<i>0.8</i>	0.3125	0.4	0.5851	0.8	0.3348	0.4	0.0000	0.0
NBTree	0.6667	<i>0.8</i>	0.2092	0.2	0.5315	0.2	0.3583	0.6	0.0946	0.2
RF	0.5222	<i>0.8</i>	0.4598	0.4	0.5343	0.2	0.3550	0.6	0.0613	0.2
REPTree	0.4927	0.6	0.0000	0.0	0.5343	0.2	0.3513	0.6	0.0000	0.0
zeroR	0.0000	0.0	0.0000	0.0	0.5343	0.2	0.0000	0.0	0.0000	0.0
SL	0.5818	<i>0.8</i>	0.1678	0.4	0.5343	0.2	0.3488	0.6	0.0000	0.0
Multiple-region										
<i>QR_{1and2}</i> -LCE-QBRE-MAX	0.9821	1.0	0.7856	1.0	0.8142	1.0	0.6568	0.8	0.4757	0.4
BN	0.9036	1.0	0.4367	0.0	0.7050	0.8	0.3832	0.2	0.4260	0.0
NB	0.6382	0.8	0.3532	0.4	0.6923	0.8	0.4250	0.6	0.4192	0.2
NBU	0.6382	0.8	0.3943	0.4	0.6949	0.8	0.3900	0.6	0.4277	0.0
CNB	0.5389	0.8	0.0000	0.0	0.6576	0.8	0.3077	0.4	0.1073	0.6
DS	0.5488	0.4	0.0000	0.0	0.6950	0.8	0.0000	0.0	0.0000	0.0
J48	0.4364	0.6	0.4803	0.2	0.6950	0.8	0.3832	0.2	0.4300	0.0
LMT	0.6479	0.8	0.4546	0.0	0.7636	1.0	0.4140	0.4	0.4273	0.0
NBTree	0.8229	1.0	0.4267	0.0	0.6783	0.8	0.3974	0.2	0.3915	0.2
RF	0.7831	0.8	0.4921	0.2	0.6999	0.8	0.3896	0.2	0.4282	0.0
REPTree	0.7572	0.8	0.0000	0.0	0.6950	0.8	0.3832	0.2	0.4313	0.0
zeroR	0.0000	0.0	0.0000	0.0	0.6950	0.8	0.0000	0.0	0.0000	0.0
SL	0.7727	1.0	0.4320	0.0	0.6864	0.8	0.4268	0.8	0.4252	0.0

Oliver Query 1

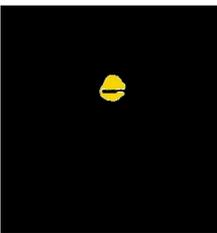
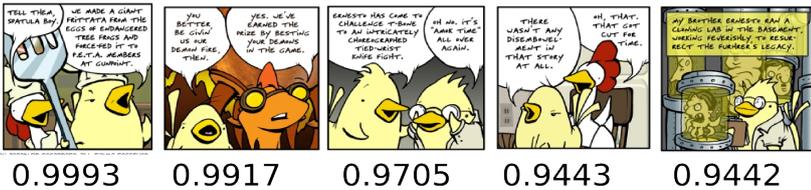
Whole Image Query



Single-region Query



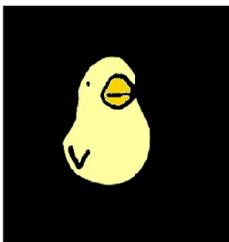
Region 1



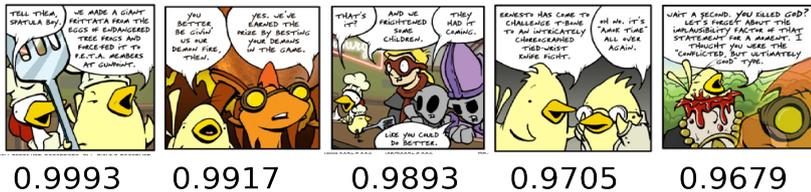
Region 2



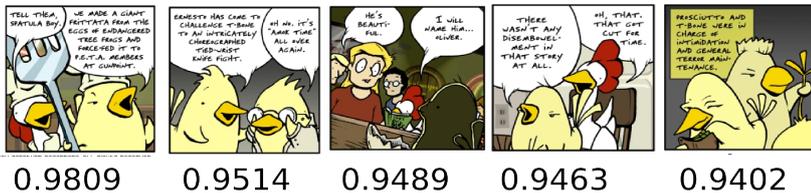
Multiple-region Query



Maximum



Sum



Minimum

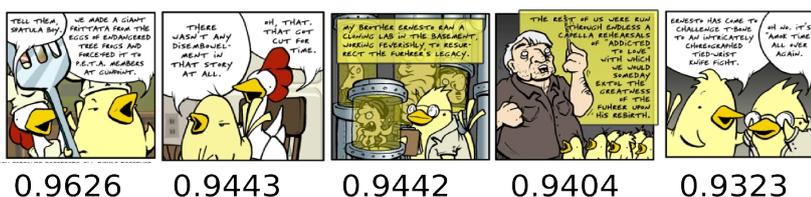


Figure 4: Comic frames retrieved for the character “Oliver” using (top) the whole image as the query example, (middle) two single-region queries, and (bottom) a multiple-region query. This figure is best viewed in colour.

- Enser, P. (1993), 'Query analysis in a visual information retrieval context', *Journal of Document and Text Management* **1**(1), 25–52.
- Fawcett, T. & Mishra, N., eds (2003), *Tackling the Poor Assumptions of Naïve Bayes Text Classifiers*, AAAI Press.
- Gevers, T. & Smeulders, A. W. M. (2000), 'Pic-ToSeek: Combining Color and Shape Invariant Features for Image Retrieval', *IEEE transactions Image Processing* **9-1**, 102 – 119.
- Hare, J. S., Lewis, P. H., Enser, P. G. B. & Sandom, C. J. (2006), Mind the gap: Another look at the problem of the semantic gap in image retrieval, in Chang, E. Y., Hanjalic, A. and Sebe, N., ed., 'Proceedings of Multimedia Content Analysis, Management and Retrieval 2006 SPIE', Vol. 6073, pp. 607309–1.
- Iqbal, Q. & Aggarwal, J. K. (2002), Cires: A system for content-based retrieval in digital image libraries, in 'Proceedings of International Conference on Control, Automation, Robotics and Vision (ICARCV)', pp. 205–210.
- Jain, A. & Vailaya, A. (1996), 'Image Retrieval using Color and Shape', *Pattern Recognition Society* **29**(8), 1233–1244.
- John, G. H. & Langley, P. (1995), Estimating continuous distributions in Bayesian classifiers, in 'Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence', Morgan Kaufmann, San Mateo, pp. 338–345.
- Kohavi, R. (1996), Scaling up the accuracy of Naïve-Bayes classifiers: a decision-tree hybrid, in 'Proceedings of the Second International Conference on Knowledge Discovery and Data Mining', pp. 202–207.
- Landwehr, N., Hall, M. & Frank, E. (2005), 'Logistic model trees', *Machine Learning* **59**(1-2), 161–205.
- Latecki, L., Melter, R. & Gross, A. (2002), 'Special issue: Shape Representation and Similarity for Image Databases', *Pattern Recognition* **35**(1), 1–2.
- Lavrac, N., Gamberger, D., Todorovski, L. & Blockeel, H., eds (2003), *Logistic Model Trees*, Vol. 2837 of *Lecture Notes in Computer Science*, Springer.
- Lew, M. S., Sebe, N., Djeraba, C. & Jain, R. (2006), 'Content-based multimedia information retrieval: State of the art and challenges', *ACM Transaction on Multimedia Computing, Communication and Applications* **2**(1), 1–19.
- Lewis, D. D. (1998), Naïve (Bayes) at forty: The independence assumption in information retrieval., in C. Nédellec & C. Rouveirol, eds, 'Proceedings of ECML-98, 10th European Conference on Machine Learning', number 1398, Springer Verlag, Heidelberg, DE, Chemnitz, DE, pp. 4–15.
- Quinlan, J. R. (1993), *C4.5: Programs for Machine Learning*, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.
- Remco, R. B. (2004), 'Bayesian network classifiers in weka'. (Working paper series. University of Waikato, Department of Computer Science. No. 14/2004). Hamilton, New Zealand: University of Waikato.
URL: http://researchcommons.waikato.ac.nz/cms_papers/36
- Smeulders, A., Worring, M., Santini, S., Gupta, A. & Jain, R. (2000), 'Content-based image retrieval at the end of the early years', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **22**(12), 1349–1380.
- Squire, D. M., Müller, W., Müller, H. & Pun, T. (2000), 'Content-based Query of Image Databases: Inspirations from Text Retrieval', *Pattern Recognition Letters* **21**(13–14), 1193–1198. (special edition for SCIA'99).
- Tahaghoghi, S. M. M., Thom, J. A. & Williams, H. E. (2001), Are two pictures better than one?, in 'Proceedings of the 12th Australasian Database Conference (ADC'2001)', Vol. 23, Gold Coast, Australia, pp. 138–144.
- Tahaghoghi, S. M. M., Thom, J. A. & Williams, H. E. (2002), 'Multiple-example queries in content-based image retrieval', *Proceedings of the Ninth International Symposium on String Processing and Information Retrieval (SPIRE'2002)* pp. 227–240.
- Vasconcelos, N. & Kunt, M. (2001), Content-based retrieval from image databases: Current solutions and future directions, in 'Proceedings of International Conference on Image Processing', pp. 6–9.
- Witten, I. H. & Frank, E. (2005), *Data Mining: Practical Machine Learning Tools and Techniques*, Morgan Kaufmann Series in Data Management Systems, second edn, Morgan Kaufmann.
- Witten, I. H., Frank, E., Trigg, L., Hall, M., Holmes, G. & Cunningham, S. J. (1999), Weka: Practical machine learning tools and techniques with Java implementations, in N. Kasabov & K. Ko, eds, 'Proceedings of the ICONIP/ANZIIS/ANNES'99 Workshop on Emerging Knowledge Engineering and Connectionist-Based Information Systems', pp. 192–196. Dunedin, New Zealand.
- Zhang, D. & Lu, G. (2004), 'Review of Shape Representation and Description Techniques', *Pattern Recognition Society* **37**, 1–19.