SaFe: A General Framework for Integrated Spatial and Feature Image Search

John R. Smith
Dept. of Electrical Engineering
Columbia University
New York, N.Y. 10027 USA
jrsmith@ctr.columbia.edu

Shih-Fu Chang
Dept. of Electrical Engineering
Columbia University
New York, N.Y. 10027 USA
sfchang@ctr.columbia.edu

Abstract - We present a system for querying for images by the spatial and feature attributes of regions[1]. The system enables the user to find the images that contain an arrangement of regions similar to that diagrammed in a query image. We propose a general framework which allows for different types of features (e.g., color, texture, shape, motion) to be integrated with spatial information in the query process. We demonstrate that integrated spatial and feature querying improves image search capabilities over previous content-based image retrieval methods.

INTRODUCTION

In this paper, we present the general framework and a prototype system for querying for images by spatial and feature attributes. The spatial and feature (SaFe) system integrates content-based techniques with spatial query methods in order to search for images by arrangements of regions. SaFe has been deployed on-line in an application for querying in a large collection of unconstrained images (more than 650,000 images). Our contribution is the use of fully automated region and feature extraction and indexing, and the integration of spatial and feature image querying. These capabilities of SaFe distinguish it from other recent image retrieval systems (Virage [1], QBIC [2] and Photobook [3]) which do not provide this enhanced functionality.

The spatial and feature image query paradigm provides a powerful method for image retrieval. However, it is extremely complex in that it requires that several disparate image query techniques be combined. First, the feature query component requires the assessment of the feature similarities of regions. Second, the spatial query component requires the assessment of the similarities in spatial locations and sizes of regions. Third, the system requires the comparison of images consisting of multiple regions. Last, the system requires that the spatial relationships, such as "above," "below," "near," and so forth, be resolved. As depicted in Figure 1, the overall comparison of images

[1] for on-line demo see http://disney.ctr.columbia.edu/safe

utilizes both the feature and spatial attributes of the regions in computing their similarity.

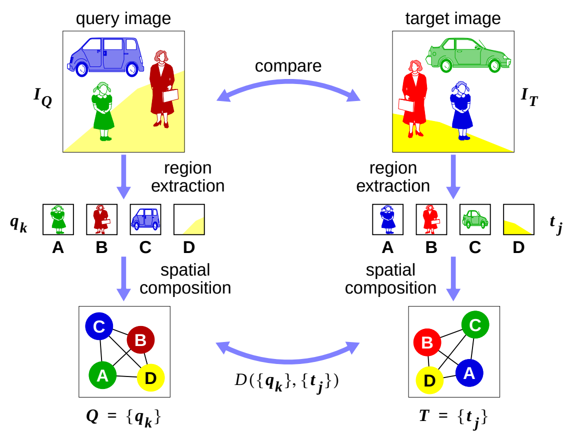

Figure 1: Integrated spatial and feature query. Images are compared by comparing the spatial arrangements of regions.

REGION ATTRIBUTE QUERY

The region attributes are stored in a region relational table (`REGION`), as depicted in Table 1. The spatial attributes of a region are determined by the spatial centroid of the region (x, y), the `area` of the region, and the width and height (w, h) of the minimum bounding rectangle (MBR) that encapsulates the region. The feature attributes ($\mathbf{f}$) of the region correspond to the visual features of the region, such as its color, texture, shape, motion and so forth.

IMID	REGID	f	x	y	area	w	h
0001	0001	$\mathbf{f}_0$	18	63	430	30	15
0001	0002	$\mathbf{f}_1$	34	45	968	65	32
0002	0001	$\mathbf{f}_2$	76	54	780	53	42
0003	0001	$\mathbf{f}_3$	55	12	654	43	55

Table 1: The `REGION` relation stores region attributes.

The region attribute queries are performed using a parallel strategy. For example, given the query region: $Q = \{\mathbf{f}_q, (x_q, y_q), \mathtt{area}_q, (w_q, h_q)\}$, the query is processed by computing queries individually for each attribute. The region attribute match lists are then `Join`ed to obtain the best set of regions. We have also explored a pipeline strategy for the region attribute queries. The query is processed by first computing the query on the feature attribute. This output is then filtered by location, then by size and spatial extent. The pipeline strategy avoids the computation of the attribute `Join` required in the parallel strategy. However, a special indexing structure may only be used one attribute, i.e., feature value. The pipeline strategy cannot effectively utilize special indexes for the other attributes that follow in the pipeline, i.e., R-tree for region MBRs, or quad-tree on region centroids.

IMAGE QUERY

The overall image query strategy consists of `Joining` the results of the queries on the individual regions in the query image. The `Join` of the region match results identifies the candidate target images. For these images, the image match score is computed by adding the weighted region scores. In the final stage, the relative spatial locations that may have been specified in the query are evaluated using query-time 2-D string projection. A 2-D string comparison determines whether candidate target images satisfy the constraints of the relative region placement. The image match process is illustrated in Fig. 2.

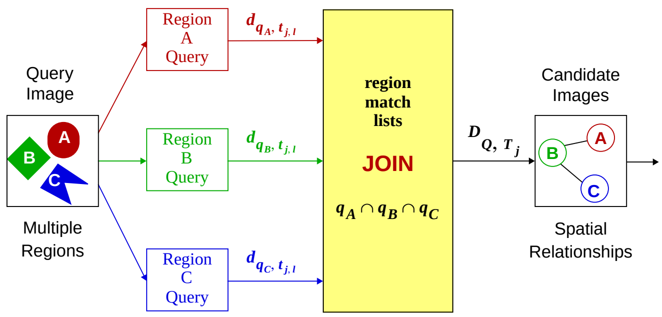

Figure 2: Overall strategy for computing image matches by `Joining` the individual regions queries and resolving spatial relationships at the last stage.

Region relative locations

Querying by absolute locations is not easily extended to include relative locations of regions. We solve the problem of the combinatorial explosion by delaying projection and evaluation of relative spatial relationships until the final stage of the query. The query-time 2-D string projection is generated from the image match lists as follows: the region centroids (x, y) are first quantized. The quantization step size, which is the size of the grid in the query formulation area, is set by the user. The regions are then projected in the x- and y- directions and are sorted to produce the 2-D string. Given the 2-D strings for the query image Q and target images T_j, the spatial relation constraints in Q are evaluated in the T_j's by 2-D string comparison. The 2-D string comparison is carried out using a straight-forward implementation of the 2-D string compare function described in [4].

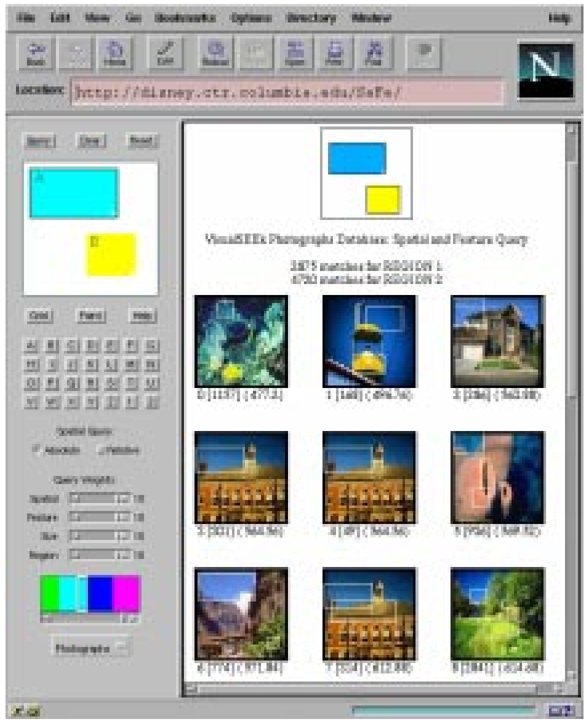

Figure 3: SaFe user interface and returned image matches.

Color photograph query

The integrated spatial and color queries are formulated graphically using the interface depicted in Figure 3. The user sketches regions, positions them on the query grid and assigns them properties of color, size and absolute location. The user assigns boundaries for region location and size. The user also chooses absolute and/or relative spatial querying. In the following query we illustrate the power and flexibility of the system over non-spatial techniques.

Automatic region extraction In order to extract the color regions from images, we developed a color set back-projection technique. We briefly describe the technique here (see [5] for more details). The back-projection of a color set is accomplished as follows: given image $I[x, y]$ and color set $\mathbf{c}$, let k be the index of the color at image point $I[x, y]$, then generate image $B[x, y]$ by $B[x, y] = \max_{j=0\ldots M-1}(A_{j,k}c[j])$, where $A_{j,k}$ measures the similarity of colors j and k. After back-projecting the model color set, image $B[x, y]$ is filtered and analyzed to reveal a set of spatially localized color regions that match color set $\mathbf{c}$.

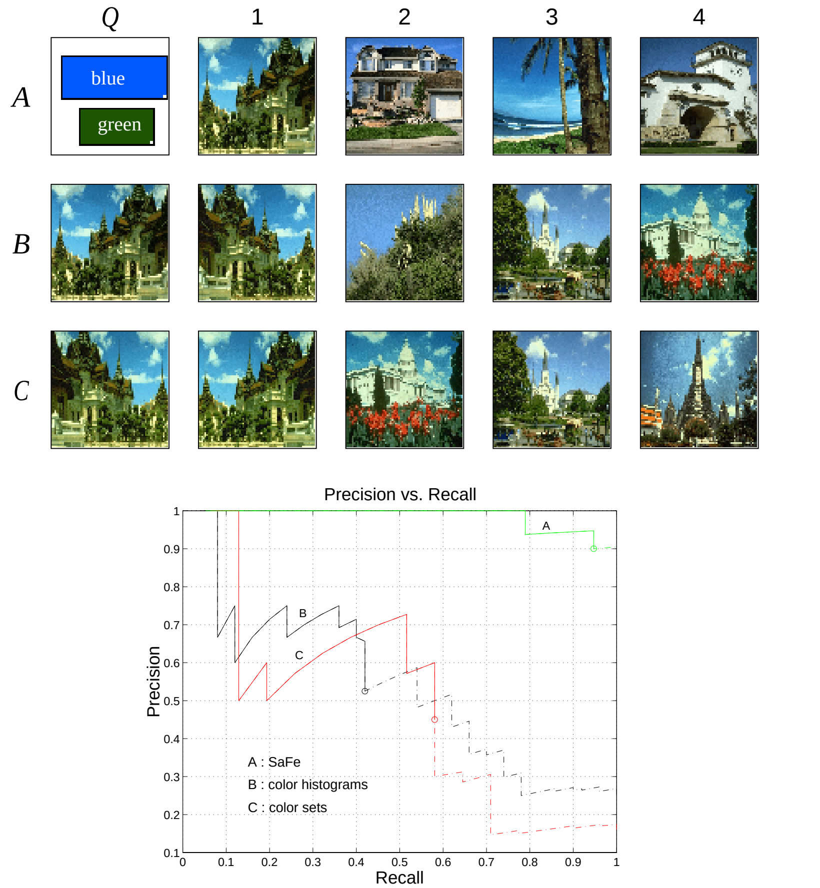

Figure 4: Example color photographic image queries, Q = query images. Best four matches are listed from left to right, where $\mathcal{A}$ = SaFe query, $\mathcal{B}$ = color histograms, $\mathcal{C}$ = color sets.

Example image query We illustrate an example image query to find images that depict blue skies and greenery. Prior to the trials, the test-set of 3,100 images (see [5]) was inspected and each image was manually assigned a relevance of 0, 0.5 or 1. The queries and results are depicted in Fig. 4. The SaFe query (denoted by $\mathcal{A}$) specifies two regions (upper is blue and lower is dark green) and their spatial layout (top left of Figures 4). The best matches (illustrated on top from left to right) have a similar arrangement of similarly colored regions. For the color histogram (denoted by $\mathcal{B}$) and color set (denoted by $\mathcal{C}$) queries, the best match image from $\mathcal{A}$ is used as the seed image. Query $\mathcal{B}$ is computed using the histogram quadratic distance metric from [2]. Query $\mathcal{C}$ is computed using the binary set quadratic distance metric from [5]. The plot in Fig. 4 shows that the retrieval effectiveness for the SaFe query $\mathcal{A}$

improves substantially over the global color feature methods given by $\mathcal{B}$ and $\mathcal{C}$.

SUMMARY

We described a new paradigm and prototype system for image searching which integrates spatial (region absolute and relative locations, and size) and feature querying (visual features, i.e., color). Content-based image query systems ([1, 2, 3]) do not provide both types of querying. Since the discrimination of images is only partially provided by global features, the system utilizes image regions and their features, sizes, spatial locations, and relationships in order to compare images. The integration of content-based and spatial querying provides for a highly functional query system which allows for wide variety of complex spatial and feature queries. We have recently combined the SaFe image query system with the WebSEEk[2] image and video search engine [6] to allow for integrated spatial and feature querying of images from the World-Wide Web.

References

[1] J. R. Bach, C. Fuller, A. Gupta, A. Hampapur, B. Horowitz, R. Humphrey, R. C. Jain, and C. Shu. Virage image search engine: an open framework for image management. In *Symposium on Electronic Imaging: Science and Technology – Storage & Retrieval for Image and Video Databases IV*, volume 2670, pages 76 – 87. IS&T/SPIE, 1996.

[2] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, and P. Yanker. Query by image and video content: The QBIC system. *IEEE Computer*, 28(9):23 – 32, September 1995.

[3] A. Pentland, R. W. Picard, and S. Sclaroff. Photobook: Tools for content-based manipulation of image databases. In *Proceedings of the SPIE Storage and Retrieval Image and Video Databases II*, February 1994.

[4] S.-K. Chang. *Principles of Pictorial Information Systems Design*. Prentice-Hall, Inc, Englewood Cliffs, NJ, 1989.

[5] J. R. Smith. *Integrated Spatial and Feature Image Systems: Retrieval, Analysis and Compression*. PhD thesis, Columbia University, New York, NY, February 1997.

[6] J. R. Smith and S.-F. Chang. WebSEEk: a content-based image and video search engine for the world-wide web. *IEEE Multimedia*, Summer 1997. `http://disney.ctr.columbia.edu/webseek`.

[2] for on-line demo see `http://disney.ctr.columbia.edu/webseek`