

An Image Clustering and Retrieval Framework Using Feedback-based Integrated Region Matching

Liping Zhou, Chengcui Zhang

Department of Computer and Information Sciences, the University of Alabama at Birmingham, Birmingham, AL, USA
{zlp, zhang}@cis.uab.edu

Wen Wan

Division of Preventive Medicine, the University of Alabama at Birmingham, Birmingham, AL, USA
wenw@uab.edu

Jeffrey Birch

Department of Statistics, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA
jbbirch@vt.edu

Wei-Bang Chen

Department of Computer and Information Sciences, the University of Alabama at Birmingham, Birmingham, AL, USA
wbc0522@cis.uab.edu

Abstract—Most existing object-based image retrieval systems are based on single object matching, with its main limitation being that one individual image region (object) can hardly represent the user's retrieval target especially when more than one object of interest is involved in the retrieval. In this paper, we present a Feedback-based Image Clustering and Retrieval Framework (FIRM) using a novel image clustering algorithm and integrating it with Integrated Region Matching (IRM) and Relevance Feedback (RF). The performance of the system is evaluated on a large image database, demonstrating the effectiveness of our framework in reflecting users' retrieval interests in object-based image retrieval.

Keywords—information retrieval; machine learning; clustering algorithm; query reweighting

I. INTRODUCTION

Object-based image retrieval has recently become an important research issue in retrieving images on the basis of the semantics of images. However, two critical issues exist in existing object-based image retrieval systems. First, most existing systems retrieve images according to a single object/segment of the user's interest, thus cannot deal with the user query that involves more than one object-of-interest. In addition, image segmentation, known as an extremely difficult process, may produce inaccurate segmentation, resulting in over-segmentation and/or under-segmentation and therefore may negatively impact the image retrieval results. Second, most existing systems require a complex user interface which can display all the segments (regions) of a query image so that users can choose desirable region(s) from these segments.

Integrated Region Matching (IRM) measure [1][3] has been proposed to alleviate the above two problems to some degree. First, IRM does not require a complex user interface to display all image segments in the query image because IRM adopts an overall image-to-image similarity measure

based on the similarity of two sets of objects/segments. Second, IRM effectively reduces the side effect of inaccurate segmentation by incorporating properties of all the regions in the images into one single integrated region matching scheme. However, several challenges remain for IRM, including: (1) how to efficiently index and search in a large-scale image segment/object database, and (2) how to bridge the "semantic gap" between low level object features and high level perceptions of image content consisting of a set of objects. This paper aims to provide an object-based image clustering and retrieval framework with feedback-based integrated region matching to address the challenges aforementioned.

In order to support integrated region-based image retrieval, we need to divide each image into several semantic regions and examine integrated region similarity during image retrieval. However, this further increases the search space by a factor of 7~8 when compared with single-region based image retrieval. Therefore, clustering is used to reduce search space [7]. Given the huge amount of regions/segments in this problem, we first preprocess image regions by grouping them into clusters. In this way the search space can be reduced to a few clusters that are relevant to the regions/objects in the query image. In this study, we propose a new genetic algorithm based clustering algorithm called Modified Genetic Algorithm (MGA) for image region clustering that is more robust compared to the state-of-the-art.

To bridge the semantic gap between low level region/object features (e.g., color, texture, and shape) and high-level user perceptions, relevance feedback [2] is used as a powerful technique in the field of information retrieval and has been an active research area for the past decade. In this paper, we proposed a Feedback-based Integrated Region Matching (FIRM) scheme to address this issue. With the proposed scheme, we consider each image as a set of regions, and the entire image database can be viewed as a

large collection of regions. After clustering, image regions are grouped into clusters. Each cluster represents a specific semantically meaningful concept. In the initial retrieval, the user provides a query image, and the system performs retrieval on the set of image clusters (reduced search space) that involve the regions in the query image using the original IRM scheme [3]. After the initial retrieval, the top 30 images are returned as the results to the user, and the user is asked to provide feedback (positive/relevant or negative/irrelevant) for each returned image. In order to decide whether the returned images and the query image are relevant, users are required to take 3 or 4 major regions in a query image (all the major objects in a query image) into consideration. We assume that the regions in the query image are not equally important to the user query. Some regions capture more of the user's attention during retrieval than the others. This is modeled in our matching scheme as assigning different weights to individual regions. The region weights are then used in the calculation of the similarity between the query image and an image in the database. A learning component is designed and implemented to automatically calculate and adjust the weights of query regions according to the user's feedback. We take the set of region weights as our hypothesis of the user's search interest and continue to run the feedback-retrieval process using the proposed FIRM scheme. Ideally, through several iterations, the user's preferences can be captured through gradual refinement of region weights. In brief, the proposed framework consists of two major parts, including the preprocessing of images (region segmentation, feature extraction and clustering), and relevance feedback and retrieval.

In Section II we present the preprocessing module. The relevance feedback and retrieval approach is explained in Section III. In Section IV, the evaluation of system performance with experimental results is presented. Finally, Section V concludes the paper.

II. PREPROCESSING

In this section, we describe the detailed preprocessing procedure for the image database. First, images are segmented into semantic regions. Then, object-level features are extracted for each image region. Then images are grouped as clusters with the proposed modified genetic algorithm. All the image clusters that involve the regions in the query image form the candidate image pool for subsequent retrieval and user feedback. Through this way, the search space can be effectively reduced.

A. Region Segmentation and Feature Extraction

In this study, we employ a fast yet effective image segmentation method called WavSeg as proposed in our recent work [11] to partition images instead of manually dividing each images into a couple of regions. In WavSeg, a wavelet analysis is used to segment an image into regions in concert with the SPCPE algorithm [12].

Region feature extraction is a built-in component in the proposed framework, which is used to extract low-level, object-based visual features that can be used to describe image content. In this paper, we extract eight visual features

as adopted in Blobworld [9] for each region. The eight visual features include 3 color features, 3 texture features, and 2 shape features for each extracted image region.

B. Clustering of Image Regions

1) *The Proposed Modified Genetic Algorithm (MGA):* Genetic algorithms [6] are iterative optimization procedures that repeatedly apply GA operations (such as selection, crossover, and mutation) to a group of solutions until some criterion of convergence has been satisfied. The GA starts with an initial population of chromosomes and then the population is evolved through generations. During each generation, parent chromosomes are selected; child chromosomes are created; and the fittest child is identified. The above mentioned process iterates until an (near) optimal solution is found.

The process of the original GA is random, so it does not utilize any directional search explicitly. In order to improve the computational efficiency of the GA, we modify the GA by incorporating a new local directional search method into the GA process, namely MGA [13].

The principle of our local search method is to utilize the best offspring among both the offspring and the parent populations from a GA process and find three local directions by only requiring a few extra function evaluations, so the GA process may be guided to further possible improvement. When the best offspring among both the offspring and the parent populations is found, we can trace back to find its parents. This process establishes two directions: one is from the first parent to the current best offspring; and the other is from the second parent to the same offspring. Both directions have obtained improvement, since the best offspring of interest is an improvement over both its parents in terms of evaluation results of an objective function. We individually project the two directions to n axis for a n -dimensional space and compare the components of parent directions in i^{th} axis, if both parent directions are consistent on i^{th} axis (either both positive or both negative), the i^{th} component of the third direction on the i^{th} axis is the parents' common direction; otherwise, the third direction is zero which means the best offspring (the starting searching point) will not move on the i^{th} axis for the third direction. After each move, three new searching points are generated individually on three directions. We evaluate the new searching points using the fitness function and compare their fitness values with the previous searching point. This process iterates until there is no further improvement in evaluation by fitness function, and then we replace the best offspring with the best searching point among three directions.

In general, the procedure of a MGA consists of the following steps:

1. Define an objective/fitness function, and its variables. Configure GA operations (such as population size, parent/offspring ratio, selection method, the number of crossovers, and the mutation rate).
2. Randomly generate the initial population.

3. Evaluate each individual (or chromosome) in the initial population by the objective function.
4. Generate an offspring population, by GA operations (such as selection/mating, crossover, and mutation).
5. Evaluate each individual in the offspring population by the objective function.
6. Check if the best offspring in the offspring population is also the best over the current parental population.
 If ‘No’, go to Step 7.
 If ‘Yes’, find the three local directions aforementioned. Collect data points along the paths with some appropriate moving distance until no improvement is observed in the objective function. Find the best point and replace the best offspring by the best point. Then go to Step 7.
7. Decide which individuals to include in the next population. This step is referred to as “replacement” in that individuals from the current parent population are “replaced” by a new population, whose individuals come from the offspring and/or parent population. In MGA, a ranking replacement is used in which all individuals in the offspring and current parent populations are sorted from best to worst, and only the top M individuals [13] are used to replace the top M individuals in the original parent population and are involved into the next generation.
8. If a stopping criterion (i.e. the number of evaluations reaches a pre-defined value) is met, the procedure is halted. Otherwise, go to Step 4.

2) *Applying MGA on Image Region Clustering*: In our application, one image is defined as a set of image regions, and each image region is represented as an 8-dimensional vector in the dataset. The image region clustering problem is defined as the task of dividing an original image region dataset into a desired number of groups so that the Euclidean distance between each data point and its corresponding cluster centroid is minimized.

In order to apply the proposed MGA algorithm to cluster image regions, a chromosome encoding scheme is proposed. In the proposed encoding scheme, the cluster centroids are encoded as genes into the chromosome and the length of each chromosome is the number of cluster centroids, i.e., the number of desired clusters. In our test image database, there are 8,900 images which are further segmented into 82,556 regions. The value of each gene in a chromosome is its index in the set of regions which ranges from 1 to 82,556, and it must be unique in the same chromosome. One limitation of many existing clustering algorithms is that they assume the number of clusters is known. However, in practice, the number of clusters may not be known. In this study, we use an approach to achieve pseudo-unsupervised clustering which aims at determining an appropriate number of clusters without any prior knowledge about it. In

particular, we proceed by repeating the clustering for different numbers of groups, and choosing the best partition. We finally decide to divide the whole set of image regions into 1,000 clusters since it results in a good balance between good fitness and efficiency. In other words, the original image dataset is represented by a dataset which contains 82,556 points (image regions) centered around 1,000 cluster centroids.

The sum of distance of each point to its cluster centroid, known as the total distance measurement of the clustering, is calculated as in (1). We use the Euclidean distance to calculate the distance between two images as in (2).

$$Totaldist(C) = \sum_{j=1}^k \sum_{p_i \in C_j} Dist(p_i, C_j) \quad (1)$$

$$Dist(I, J) = \left(\sum_t |feature_t(I) - feature_t(J)|^2 \right)^{\frac{1}{2}} \quad (2)$$

In (1), C_j is the j^{th} cluster centroid of the input dataset (the set of regions); p_i represents the i^{th} data point (image region) in cluster j ; each data point (region) is described as an 8-feature vector. C_j is an existing data point (image region) in the dataset chosen as the centroid for cluster j . Equation (2) is used to calculate the distance $Dist(I, J)$ between region I and region J , where $feature_t(I)$ is the t^{th} feature of region I ($t = 1 \dots 8$).

In this study, the fitness function is defined as the inverse value of the total distance which is calculated as shown in (3). Our MGA algorithm aims at approximating the maximum of the fitness function.

$$fitness = \frac{1}{Totaldist(C)} \quad (3)$$

Other implementation choices include using random pairing for selection, applying blending crossover in which the number of crossover points depends on the number of dimensions of a specific objective function and is set to 2 in our case. Random uniform mutation is utilized and the mutation rate is set to 0.04. The type of replacement over both parent and offspring populations is ranking in which the population size is 50 and M is set to be the same as the population size (see Step 7 in Section II.B.1 and [13]).

3) *Comparison between MGA and CMA-ES*: In order to evaluate the performance of the proposed MGA, we compare it with Covariance Matrix Adaptation Evolution Strategy (CMA-ES) [8] in tackling the image region clustering problem. The quality of clustering is evaluated in terms of the total distance measurement as mentioned in the previous section, and a smaller total distance value indicates better clustering quality. The CMA-ES is considered as an attractive option for non-linear optimization in a discontinuous search space. In our application, data points (image regions) are scattered in a multidimensional space and therefore are discrete. For an input dataset with 82,556 data points (regions), our goal is to cluster these points into 1,000 groups where the sum of distance of each point to its corresponding cluster centroid is minimized. Table I

demonstrates the comparison results of MGA and CMA-ES. The comparison is based on 2,400 evaluations of the fitness function. This particular number is chosen because both algorithms start to converge at around 2,000 evaluations. The results presented in Section IV are also based on the same number of evaluations. We observe that MGA has better performance than that of the CMA-ES based on the total distance measurement.

TABLE I. RESULTS OF COMPARISON BETWEEN MGA AND CMA-ES

Group #	Clustering Algorithm	The Total Distance
1000 Groups	MGA	67,872.1
	CMA-ES	69,926.4

III. RELEVANCE FEEDBACK AND RETRIEVAL

In this section, an innovative region-matching scheme is proposed for relevance feedback and retrieval. It defines that the similarity of two images is the weighted sum of the similarities of their matched regions. To calculate the similarity, we first need to find out all the matched pairs of regions in two images. The relevance feedback suggests which regions in a query image better capture a user's attention. Then the weight of regions is updated by learning through relevance feedback. The principle is that the highest weight should be assigned to the region where the user is the most interested. Compared with the original Integrated Region-Matching scheme, our region-matching scheme further improves the retrieval accuracy.

A. The Proposed Region-Matching Scheme

According to the principle of integrated region matching, we first attempt to match regions in two images. The most similar region pair is matched first. The matching allows one region of an image to be matched to several regions of another image. The distance between two regions can be easily calculated as the Euclidean distance in terms of the features extracted. During the matching process, the similarities of region pairs are calculated, and proper weights are assigned to them. The overall integrated similarity score between two images is computed as a weighted sum of the similarity between region pairs. We define the overall distance of two images in (4), where I_1 is the query image which has m regions and I_2 represents an image with n regions in the candidate pool; i and j denote the i^{th} region of I_1 and j^{th} region of I_2 , respectively. The matrix S is referred to as the significance matrix, in which s_{ij} represents the weight assigned to the region pair i and j . D is referred to as the distance matrix in which d_{ij} is the Euclidean distance between two regions i and j . w_i is the weight of region i in the query image (I_1). The initial value of w_i is set to 1, and w_i is automatically updated through relevance feedback learning process.

$$D(I_1, I_2) = \sum_{i=1}^m w_i \sum_{j=1}^n s_{ij} d_{ij} \quad (4)$$

In the initial retrieval, we assume one query image I_q has m regions, and a candidate image I_c has n regions. The significance of region i in I_q and region j in I_c are denoted as s_i^q and s_j^c , respectively. We initialize the significance of a region as the area percentage of that region in a given image, assuming that important objects in an image tend to be those bigger ones. s_{ij} can be derived from s_i^q and s_j^c , subject to the following constraint as given in (5).

$$\sum_{j=1}^n s_{ij} = s_i^q \quad \sum_{i=1}^m s_{ij} = s_j^c \quad (5)$$

$$s_i^q = s_i^q - \min(s_i^q, s_j^c) \quad (6)$$

$$s_j^c = s_j^c - \min(s_i^q, s_j^c) \quad (7)$$

According to the principle of IRM, the significance values of regions i and j are then updated with (6) and (7) with the purpose to make sure that the best matched region pair has the highest significance value. Then we can calculate the integrated similarity score with (4).

We choose to use the six biggest regions (or less if there are no more than six regions) in the query image and each candidate image in similarity comparison. One reason is that the size of the matrix will increase polynomially with the number of regions in an image, which leads to a significant increase of the time-complexity in calculating the overall distance. Another reason is that some region resulting from poor segmentation covers only a small part of a concrete object and is not suitable to be used to represent one complete semantic object in that image. In this case, ignoring these smaller regions in the retrieval actually reduces noise in the dataset.

B. Learning through Relevance Feedback

In the proposed relevance feedback framework, we collect the user's positive feedback as samples to construct the training dataset. Through feedback, an image labeled as positive indicates that it has regions that match the target regions in the query image according to the user's preferences. This information is used to update the weights of the regions in the query image. The weight of a region in the query image represents the significance of that region in calculating the integrated similarity between the two images, where the significance of a region in the query image can be reflected by the degree of matching between the region and the matched region in the candidate image. The degree of matching between a pair of regions can be determined by the distance between them. The less the distance, the better match they are.

$$w_i = w_i' \times \prod_p \frac{1}{\min_k (d_{ik})} \quad (8)$$

Equation (8) formulates the re-weighting scheme where w_i is the current weight of i^{th} region in the query image; d_{ik} is the distance between region i in the query image and region k in a positive image (p) labeled by the user. In the initial query, Equation (4) is used to return the top 30 most similar images from the ranked list for user feedback. By incorporating users' feedback into the retrieval and integrated region matching, we can dynamically update the significance (weight) of each query region. In our framework, the feedback-retrieval process will run for several iterations until the user is satisfied with the returned results.

IV. PERFORMANCE EVALUATION

In this section, we evaluate the performance of the proposed system by applying the proposed method on a Corel image database consisting of 8,900 images (82,556 image segments) from 100 categories. The number of clusters is selected to be 1,000 according to the discussion in Section II. By utilizing the image clustering, the size of the search space is reduced to between 50 and 1,700 (images). The proposed relevance feedback and image retrieval method is based on the candidate images in the reduced image set.

In our experiments, we adopt Average Normalized Modified Retrieval Rank (ANMRR) [10] and Accuracy as the standard performance measures. ANMRR measure is the averaged MPEG-7 Normalized Modified Retrieval Rank (NMRR) over the query set, as defined in (9), where q is the current query; $NG(q)$ is the number of positive images in the top 30 returned images in our case; $Rank^*(k)$ is the retrieval rank capped by the $Rank^{\max}$ which is defined as the upper limit of the retrieval rank (30 in our case).

$$NMRR(q) = \frac{1}{NG(q)} \frac{\sum_{k=1}^{NG(q)} Rank^*(k) - 0.5 * [1 + NG(q)]}{Rank^{\max} - 0.5 * [1 + NG(q)]} \quad (9)$$

The accuracy is defined as the percentage of relevant images out of the retrieved paper list. However, as a disadvantage, using accuracy to measure retrieval performance cannot faithfully reflect the rank of the returned images (i.e., ordinal position of a relevant/positive image in the retrieved image list). It is obvious that a good retrieval system should return all the relevant images at the top of the list. A true positive image becomes a missed hit or false negative if it does not appear in the short list. Therefore, it is important to consider the rank of a retrieved relevant image since it directly reflects the retrieval performance. In our study, we adopt ANMRR measure to evaluate the rank of the retrieved relevant image.

In our experimental setting, we randomly select 50 images from 12 categories as query images. Five rounds of relevance feedback (RF) are performed for each query image - Initial (no feedback), First, Second, Third, and Fourth. Accuracy and ANMRR are individually calculated for the top 30 retrieved images. In the following subsections, we evaluate the effectiveness of the proposed

MGA and the relevance feedback method individually by using Accuracy and ANMRR. In addition, the proposed framework FIRM is compared with an existing approach called Dynamic Region Matching (DRM) [4].

A. The Effectiveness of MGA in Reducing the Search Scope

In this experiment, we study the performance of the proposed feedback-based integrated region matching scheme (FIRM) with and without the use of MGA in reducing the search scope. The motivation of this experiment is to show that without significantly sacrificing the performance of image retrieval, MGA can effectively narrow down the search scope, and therefore, reduce the time complexity. We present the experimental results in Table II. FIRM indicates that MGA is used in the proposed scheme, while FULL represents the proposed matching scheme without MGA (full-scope search). The lower ANMRR values indicate better performance, but higher accuracy indicates better retrieval results.

TABLE II. THE EFFECTIVENESS OF MGA

#	ANMRR		Accuracy	
	FIRM	FULL	FIRM	FULL
1	0.255	0.325	0.395	0.436
2	0.247	0.318	0.399	0.441
3	0.235	0.320	0.392	0.451
4	0.230	0.300	0.400	0.445
5	0.229	0.299	0.408	0.446

#: 1: Initial; 2: First; 3: Second; 4: Third; 5: Fourth

Through this comparison, full-scope search achieves better retrieval accuracy at the cost of significantly higher time complexity, while FIRM trades accuracy for efficiency. In our current implementation, with full search, it takes about 4 minutes to perform one round of query, while it only takes about 20 seconds to perform the same query with FIRM and MGA. This indicates that the full search becomes impractical when the size of the image database rapidly grows. Besides, we can observe from Table II that the performance of both FIRM and FULL are gradually improved through 5 rounds of RF. Moreover, the performance of FIRM with MGA is very close to that of FULL at the end of 4th RF iteration.

B. The Effectiveness of Relevance Feedback

We further compare the proposed framework (FIRM) with the original IRM (IRM) without relevance feedback by using the above evaluation criteria. We illustrate the comparison results in Fig. 1. The experimental results indicate that FIRM outperforms IRM, demonstrating the effectiveness of integrating IRM with relevance feedback.

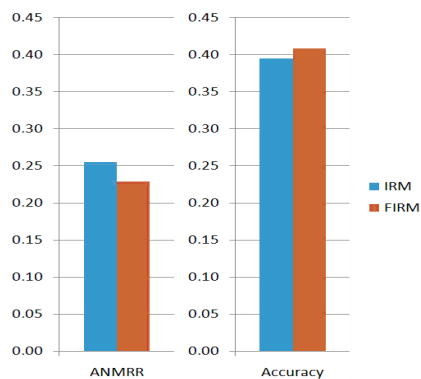


Figure 1. Comparison of IRM with FIRM.

C. Compare the Proposed Framework FIRM with DRM

In this experiment, we compare the performance of FIRM with an existing approach, Dynamic Region Matching (DRM) [4] which also uses relevance feedback to update the weights of regions. This experiment aims to evaluate the learning algorithms used in the relevance feedback process. We apply the above three methods on the same test images and compare their performance with the two evaluation measures. Table III shows the comparison results in terms of ANMRR and Accuracy measures, respectively.

TABLE III. COMPARISON BETWEEN FIRM AND DRM

#	ANMRR		Accuracy	
	FIRM	DRM	FIRM	DRM
1	0.255	0.255	0.395	0.395
2	0.247	0.319	0.399	0.220
3	0.235	0.264	0.392	0.192
4	0.230	0.269	0.400	0.193
5	0.229	0.279	0.408	0.200

#: 1: Initial; 2: First; 3: Second; 4: Third; 5: Fourth

According to our experimental results, we can observe that the proposed framework (FIRM) has better performance in terms of ANMRR and Accuracy values. The reason that DRM performs worse than FIRM is that the weights of regions in the query image are either increased or decreased by a fixed factor through relevance feedback, and therefore, some regions in the query image may quickly overpower the others through iterations of feedback-retrieval. Consequently, the performance gain obtained through feedback is limited.

V. CONCLUSIONS AND FUTURE WORK

In this paper, we present a Feedback-based Image Clustering and Retrieval Framework (FIRM) which improves the region-based image retrieval accuracy and efficiency by using a novel image clustering algorithm and integrating it with IRM and RF. However, the current RF method only applies the user's positive feedback as samples to construct the training dataset without considering negative feedback. Additionally, to enhance our performance evaluation, comparing the current retrieval

results with those based on image features extracted on a whole image base will be conducted in the future.

ACKNOWLEDGMENT

The research of Chengcui Zhang is supported in part by NSF DBI-0649894 and UAB ADVANCE program.

REFERENCES

- [1] J. Z. Wang, J. Li, and G. Wiederhold, "SIMPLiCity: semantics-sensitive integrated matching for picture libraries," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2001, pp. 947-963.
- [2] C. Zhang and X. Chen, "OCRS: an interactive object-based image clustering and retrieval system," in *Proceedings of the Fourth International Workshop on Multimedia Data Mining (MDM/KDD2005)*, in conjunction with the ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Chicago, USA, 2005.
- [3] K. S. Babu and V. Nagesh, "Optimal solution for image retrieval using integrated region matching," in *Proceedings of 2nd National Conference on Challenges & Opportunities in Information Technology (COIT-2008) RIMT-IET*, Mandi Gobindgarh, 2008, pp. 223-226.
- [4] R. R. Ji, H. X. Yao, and D. W. Liang, "DRM: dynamic region matching for image retrieval using probabilistic fuzzy matching and boosting feature selection," *Signal, Image and Video Processing*, 2(1), 2008, pp. 59-71.
- [5] C. J. Burges, "A tutorial on support vector machines for pattern recognition," *Data Min. Knowl. Discov.*, 2(2), 1998, pp. 121-167.
- [6] R. L. Haupt and S. E. Haupt, *Practical genetic algorithms*, John Wiley and Sons, Inc., New York, 2004.
- [7] Q. Ding and J. Gasvoda, "A genetic algorithm for clustering on image data," *International Journal of Computational Intelligence*, 1(1), 2005, pp. 75-80.
- [8] N. Hansen, "The CMA evolution strategy: a tutorial," DOI=<http://loutaro.fb10.tu-berlin.de/user/niko/cmatutorial.pdf>.
- [9] C. Carson, S. Belongie, and H. Greenspan, J. Malik, "Blobworld: image segmentation using expectation-maximization and its application to image querying," *IEEE Trans Pattern Anal Mach Intell.*, 24(8), 2002, pp. 1026-1038.
- [10] L. Cieplinski, "Color descriptors and their applications," Springer-Verlag GmbH, 2001, pp. 11-20.
- [11] C. Zhang, S.-C. Chen, M.-L. Shyu, and S. Peeta, "Adaptive background learning for vehicle detection and spatio-temporal tracking," in *Proceedings of the 4th IEEE Pacific-rim conference on multimedia*, 2003, pp. 1-5.
- [12] S.-C. Chen, S. Sista, M.-L. Shyu, and R. L. Kashyap, "An indexing and searching structure for multimedia database systems," in *Proceedings of the IS&T/SPIE conference on storage and retrieval for media databases*, 2000, pp. 262-270.
- [13] W. Wan, *Semi-parametric techniques for multi-response optimization*, the Virginia Polytechnic Institute & State University, 2007.