ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Content Based Image Retrieval System for Texture Feature Extraction Using Various Machine Learning Algorithm

Koteswara Rao M
Research Scholar, ECE Department
JNTUH Hyderabad, Telangana.
koteshproject@gmail.com

K Veera Swamy
Professor, ECE Department
Vasavi college of Engineering Hyderabad
k.veeraswamy@staff.vce.ac.in

K Anitha Sheela
Professor, ECE Department
JNTUH Hyderabad, Telangana.
kanithasheela@gmail.com

***Abstract*—** A localized content-based image retrieval as a CBIR task where the user is only interested in a portion of the image, and the rest of the image is irrelevant. In this paper, we present a localized CBIR system, ACCIO!, that uses labeled images in conjunction with a multiple-instance learning algorithm to first identify the desired object and weight the features accordingly, and then to rank images in the database using a similarity measure that is based upon only the relevant portions of the image. A challenge for localized CBIR is how to represent the image to capture the content. We present and compare two novel image representations, which extend traditional segmentation-based and salient point-based techniques, respectively, to capture content in a localized CBIR setting.

***Keywords*—** Machine learning, content-based image retrieval, multiple-instance learning, salient points.

I. INTRODUCTION

CLASSIC content-based image retrieval (CBIR) taksingle query image and retrieves similar images. Since the user typically does not provide any indication of which portion of the image is of interest, such a search must rely upon a global view of the image. We define localized content-based image retrieval as a CBIR task, where the user is only interested in a portion of the image, and the rest is irrelevant.

Unless the user explicitly marks the region of interest, localized CBIR must rely on multiple images (labeled as positive or negative) to learn which portion of the image is of interest. The query set contains a set of images either directly provided by the user or obtained using relevance feedback [25] to add labeled feedback images to the original query image. For example, frames from surveillance video could be available for times when suspicious activity occurred (labeled as positive) and others for times when nothing out of the ordinary occurred (labeled as negative). When used in conjunction with an image repository containing unlabeled video frames, ACCIO! could be used to search for frames that have some object in common with those containing suspicious activity. In localized CBIR, the query set is used to identify the portion(s) of the image that are relevant to the user's search as well to determine an appropriate weighting of the features.

Cite this article as: Koteswara Rao M, K Veera Swamy & K Anitha Sheela,"Content Based Image Retrieval System for Texture Feature Extraction Using Various Machine Learning Algorithm", International Journal & Magazine of Engineering, Technology, Management and Research, Volume 6 Issue 4, 2019, Page 155-165.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Many CBIR systems either subdivide the image into predefined blocks [21], [22], [28], or more commonly, partition the image into different meaningful regions by applying a segmentation algorithm [24], [30]. In both cases, each region of the image is represented as a vector of feature values extracted from that region. Other CBIR systems extract salient points [15], [16], [20], [27], [31], [32], which are points of high variability in the features of the local pixel neighborhood. With salient point-based methods, one feature vector is created for each salient point.

One distinction between region-based CBIR systems and localized CBIR is how the image is processed. Single feature vector CBIR systems represent the entire image as one feature vector. For example, a color histogram [5], [11], [18] defined over the entire image is such a representation. In contrast, multiple feature vector CBIR systems represent the image as a collection of feature vectors with one feature vector for either a block in some prespecified image subdivision (e.g., [21], and [22]), the region defined by a segmentation algorithm (e.g., [30]) or a window around each salient point (e.g., [15], [16], [20], [27], [31], and [32]).

Another important distinction is the type of similarity metric used to rank the images. In global ranking methods, all feature vectors in the image representation affect the ranking. While salient point-based methods only use portions of the image around the salient points, if the ranking method uses all salient points, then it is a global method. In contrast, local ranking methods select only portions of an image (or a subset of the salient points) as relevant to rank the images. For example, if a salient point- based method learns which subset S of the salient points are contained in desirable images and ranks images using only the subset S of salient points, then it is a local ranking method. Localized CBIR systems must use local ranking methods..

We present ACCIO!, named after Harry Potter's sum-moning charm, that uses a small query set in conjunction with a multiple-instance (MI) learning algorithm to identify the desired local content, reweight the features, and then rank new images. We also present two novel image representations. The first extends traditional segmenta- tion-based techniques, and the second extends traditional salient point-based techniques. These image representations allow ACCIO! to perform well even when the desired content is complex, defined by multiple parts or objects.

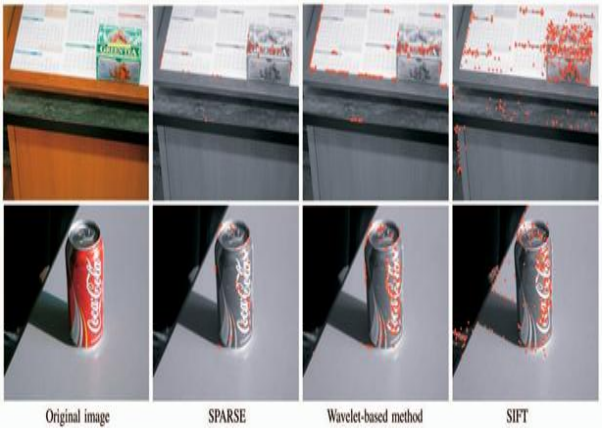

Fig. 1. Salient points detection with SPARSE, the Harr Wavelet-based method, and SIFT.

Salient Points Autoreduction Using Segmentation (SPARSE)

Our SPARSE image representation limits the number of salient points in each segment while maintaining the diversity needed for

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

localized CBIR. SPARSE first applies a salient point detection algorithm to the image. We use a Harr wavelet-based salient point detection method. Begin- ning at the coarsest level of wavelet coefficients, we keep track of the salient points from level to level by finding the points with the highest coefficients on the next finer level among those used to compute the wavelet coefficients at the current level. The saliency value of a salient point is the sum of the wavelet coefficients of its parent salient points from all coarser scales. Note any salient point detection method can be used here instead, with little modification.
Next, a segmentation algorithm is applied to the image.

The segmentation algorithm we use is a clustering-based segmentation method [10] that uses the euclidean distance between six-dimensional feature vectors, with three color features and three texture features, as its similarity measure. The resulting segmentation is used to reduce the number of salient points. Specifically, SPARSE keeps at most k salient points in each segment by keeping those with the highest saliency value. In our implementation, k ¼ 3.

Fig. 1 shows examples of salient points detected using SPARSE. For comparison, we also show the salient points detected by the Harr wavelet-based salient points detection method and the SIFT [20] method. The wavelet-based method selects the top 200 salient points for each image. SPARSE reduces it to at most 96 salient points per image. SIFT selects 392 salient points for the tea box and 288 salient points for the coke can. When using SPARSE, the salient points predominantly gather at complex objects, whereas with the wavelet-based method, the salient points gather at the edges. While the wavelet-based method reduces the number of salient points on the textured region (such as at the printed words on the calendar and tea box), SPARSE further reduces the number of salient points at textured regions.

VSWN: Our Salient Point Representation

We now describe our variably-split window with neighbor (VSWN) representation. Since salient points are often on the boundary of objects, the features assigned to a salient point involve pixels from different objects, which is not good for localized CBIR because only one of these objects might be of interest. If we divide the window, we can better capture the color and texture of an individual object.

For each salient point, VSWN uses the local characteristics of the window around each salient point to split the window in either the horizontal, vertical, or one of the two diagonal directions. The variably split window (VSW) technique adaptively chooses the best split. VSW applies a wavelet transform on the pixels in the window and measures the average coefficients in the HL (vertical), LH (horizontal), and HH (up-right diagonal). We also flip the window around the vertical to compute a flipped-HH coefficient (up-left diagonal). If the LH and HL channels have similar coefficients, then we use the split associated with the larger of the HH and flipped-HH channel. Otherwise, we use the split based on the largest of the four channels. While the best segmentation of the region is unlikely to be one of the four splits considered (since it is an 8 × 8 window), the selected split

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

serves as a sufficiently good approximation. If desired, we could further subdivide each subwindow.

Second, as was the case for segmentation with neighbors, for salient points, it is advantageous to incorporate information about the neighboring subwindow to provide additional context. The two subwindows for each salient point are represented via three color features and three texture features. VSWN augments the feature vector for each subwindow with the difference between its values and the other subwindow's values for each of the six features. We use the feature differences to allow for robustness against global changes in the image, such as brightness changes from variable light or shade. Since we do not know which subwindow might hold the object of interest, we create two 12-dimensional feature vectors for each salient point: one for each subwindow as the object of interest.

TABLE 1
Summary of AUC Values Averaged over the Categories of SIVAL and COREL Natural Scenes Data Sets

	SIVAL		Natural Scenes	
system	8 pos, 8 neg	Single pos	8 pos, 8 neg	Single Pos
Accio!	81.8	53.5	83.6	67.2
Accio! (SPARSE+VSWN)	81.6	56.3	.	.
SIMPLIcity	57.9	55.7	74.8	73.7
SBN	53.9	50.3	73.6	61.4
Accio! (conference version)	74.6	61.0	87.7	74.5

2 EXPERIMENTAL RESULTS

We compare the system-level performance of ACCIO!, SIMPLIcity [14], [30], and SBN [22] on the SIVAL and COREL natural scenes data sets, both with small query sets (2-16 images), for which ACCIO! was designed, and with the traditional CBIR setting of a single positive query image. We also compare the performance of our SPARSE+VSWN salient point-based representation to that of SIFT and the wavelet-based method. On the Flickr data set, we compare our segmentation-based and salient point-based representations. Unless otherwise indicated, ACCIO! results were produced using the segmentation-based representation, where was set roughly equal to the bag size. Thus, ¼ 25 for the segmentation-based representation, and ¼ 75 for the salient point-based representations. For results that used a single query image, we set ¼ jHj.

For the SBN algorithm, we replace DD by the EM-DD algorithm because of its performance gains in both retrieval accuracy and efficiency [38]. Since SIMPLIcity is designed to use a single positive example, we created a variant of it that uses any size query image set. Let P be the set of positive images, and let N be the set of negative images. For image q in the query set and image x in the image repository, let rqðxÞ be the ranking SIMPLIcity gives to image x when the query image is q. (The highest rank image is rank 0.) Our variation of SIMPLIcity ranks the images in decreasing order based on

$$\prod_{q2P}\left(1-\frac{r_q(x)}{n}\right)\cdot\prod_{q2N}\frac{r_q(x)}{n}.$$

We selected this measure since it is similar to the definition of diverse density of a point t, DDðtÞ¼ Q q2P[N PrðtjqÞ. For an image, q 2 P, ð1 — rqðxÞ=nÞ can be viewed as the probability that x is positive given that q is positive. Similarly, for an image, q 2 N ,

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

rqðxÞ=n can be viewed as the probability that x is positive given that q is negative. When given a single positive image, the ranking is the same as that given by the original SIMPLIcity algorithm [14], [30].

As our measure of performance, we use the area under the ROC curve [12] that plots the true positive rate as a function of the false positive rate. The area under the ROC curve (AUC) is equivalent to the probability that a randomly chosen positive image will be ranked higher than a randomly chosen negative image. Unlike the precision-recall curve, the ROC curve is insensitive to the ratio of positive to negative examples in the image repository. Regardless of the fraction of the images that are positive, for a random permutation the AUC is 0.5. For all AUCs reported, we repeat each experiment with 30 random selections of the positive and negative examples and use these to compute the average AUC and the 95 percent confidence intervals for the AUC.

System Performance

Table 1 compares the average performance (over all cate- gories) of ACCIO!, SIMPLIcity and SBN for the SIVAL and the natural scenes data sets. Fig. 2 compares all 25 object categories of SIVAL when the query set contains eight random positive and eight random negative examples. For two categories—"LargeSpoon" and "CandleWith Holder"— SIMPLIcity's segmentation algorithm failed on a few images so results could not be provided. For both representations, in every category, ACCIO!'s performance is statistically better than that of both SIMPLIcity and SBN, with the exception of "LargeSpoon" for SBN and "RapBook" for SIMPLIcity. ACCIO!, using segmentation with neighbors, has an average improvement of 51.7 percent over SIMPLIcity and 41.4 percent over SBN. ACCIO!, using SPARSE+VSWN, has an average improvement of 51.2 percent over SIMPLIcity and 41.0 percent over SBN.

Fig. 2 also compares the SPARSE+VSWN and the segmentation with neighbors representation of ACCIO!. We see similar performance in 17 of 25 categories. In five categories, the segmentation-based approach performs statistically better, and in three categories, SPARSE+VSWN is statistically better. Therefore, overall, their results are comparable. Segmentation with neighbors encodes its neighbors in a manner that preserves orientation, which is advantageous for some tasks (e.g., when distinguishing a waterfall from a river). Furthermore, encoding four neighbors, instead of just one, captures more contextual information for each segment. On the other hand, SPARSE+VSWN has several advantages over segmentation with neighbors.

The reduction in dimensionality from 30 to 12 improves the time complexity. Also, VSWN's use of a single neighbor that is both mirror invariant and rotation invariant allows it to perform better on categories in which the images experience significant rotation (90 degree and 180 degree). While the salient points currently encode only the same information as the segmentation-based method, salient points can be encoded with a multitude of additional features not easily derived from segmentation methods such as

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

the orientation histogram used by SIFT. Addition- ally, salient points by their nature can capture much finer detail in the image than segmentation.

Since SIMPLIcity was designed for a single positive query image, we also considered when the query set contains only a single positive image (not shown). On

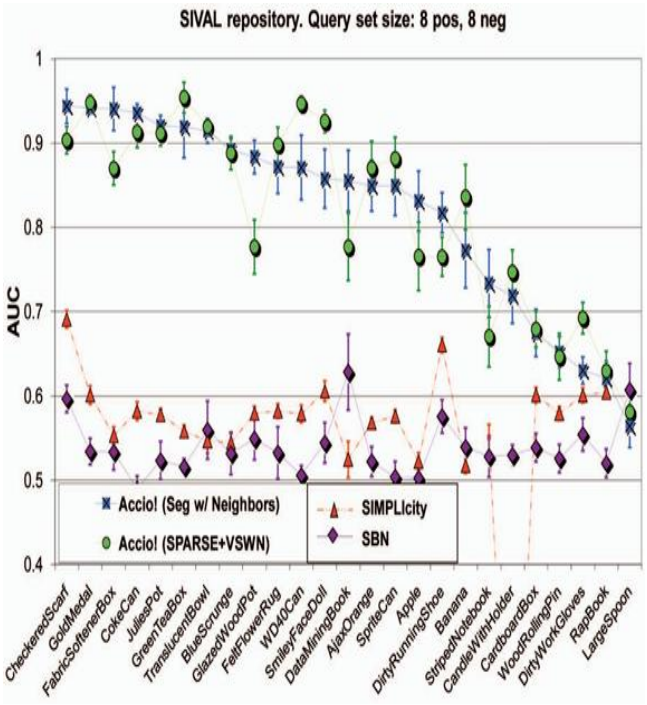

Fig. 2. CBIR systems results for the SIVAL data set when the query set has eight positive and eight negative examples.

average ACCIO! obtains a 4.3 percent improvement in performance over SIMPLIcity and a 15.6 percent improve- ment over SBN. For our alternative SPARSE+VSWN representation, ACCIO! obtains a 12.4 percent improvement in performance over SIMPLIcity and a 24.5 percent improvement over SBN. The version of SIMPLIcity we created to make use of a query set with multiple images did improve performance over having a single query image in 12 of the 23 categories.

Fig. 3 shows the performance of ACCIO! when using the segmentation with neighbors representation on the Flickr data set for varying query set sizes. Likewise, Fig. 4 shows the performance of ACCIO! when using SPARSE+VSWN. As the size of the training data increases, in general, we both get better retrieval performance, and also the variation in performance is reduced. For some categories, increasing the training size has a very small impact (e.g., waterfall, samurai helmet), yet for others (e.g., American flag and Pepsi can), the impact is quite large. When there is some aspect of a category that is very distinctive, then the smaller training set can be effective. However, when the object is defined by fairly typically occurring colors/textures (e.g., the color red) or in a category with a lot of variation (e.g., fire flame), having a larger query set can really help performance.

We now compare the performance of these two represen- tations when there are eight positive and eight negative examples. ACCIO! using segmentation with neighbors per- forms statistically better on "Snowboarding," "Sushi," and "Persian Rug." Conversely, though ACCIO! using SPAR- SE+VSWN does not perform statistically better in any of the

image categories, there is a noticeable improvement over segmentation with neighbors on the "American Flag," "Fire Flame," "Pepsi Can," and "Coca-Cola Can" categories. Since the "American Flag," "Pepsi Can," and "Coca-Cola Can" images all contain a specific complex object of interest, there are

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

a large number of salient points in each image corresponding to that object. Since the object is specific, the colors and textures defining the object are very well defined, with variations due only to lighting, shading, and noise. Therefore, the sets of salient points for these objects serve as effective identifiers enabling the SPARSE+VSWN method to perform well on such categories. The "Fire Flame" category does not target as specific an object, but the nature of fire does lend itself well to a small set of easily distinguishable colors and textures.

The "Snowboarding," "Sushi," and "Persian Rug" categories do not contain a specific object but include objects with common sets of color schemes and textures, which can be effectively captured by whole-segment feature characteristics. As long as the segmentation algorithm can effectively segment out the object regions, and there are sufficient training examples to characterize both the most common variations in the object's color and texture and the spatial relationship between the object(s) of interest, ACCIO! using segmentation with neighbors generally performs well. For example, in "Snowboarding," segmentation with neighbors searches for regions containing snow bordered by one or more neighboring regions containing trees, mountains, or sky. While these categories can be character- ized by a variety of different color schemes and textures, with effective segmentation and sufficient training data segmentation with neighbors achieves good performance for these categories.

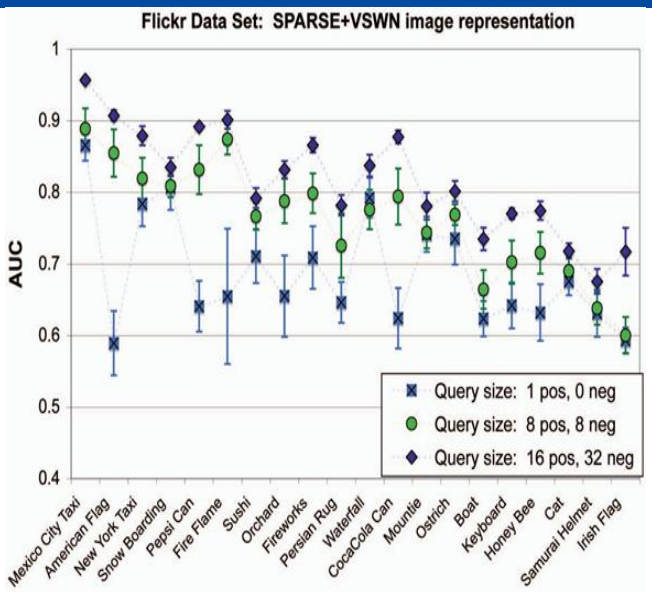

Fig. 4. Results from our SPARSE+VSWN representation on the Flickr data set.

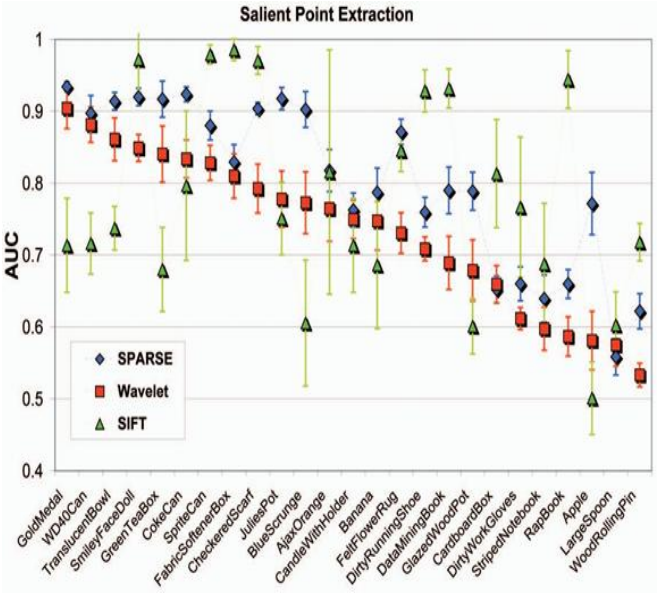

Fig. 5. Comparing salient point methods on the SIVAL data set, where the query set has eight positive and eight negative examples.

Comparison of Salient Point-Based Representations

In this section, we compare the performance of ACCIO! when using salient point-based representations. Fig. 5 compares the three

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

salient point extraction methods: Wavelet, SIFT, and SPARSE+VSWN. In these experiments, both SPARSE and Wavelet use the same salient point extraction and representation methods (three color and three texture dimensions). The primary difference between them is where the salient points are placed in the image. SIFT uses both a different feature extraction method, placing the salient points differently, and a more complex feature representation. SPARSE outperforms Wavelet in 23 of 25 categories, 16 of which are statistically significant. The use of SPARSE can also improve the algorithm efficiency by reducing the number of feature vectors per bag and hence the number of computations.

The SIFT feature vector has 128 dimensions that describe the local gradient orientation histogram around a salient point. Results using SIFT were generated using five random selections of the training data (as opposed to 30) since the high dimensionality makes it very compu- tationally intensive. SIFT performs 5.9 percent better than Wavelet over all categories. However, SPARSE outper- forms SIFT by 3.2 percent (over all categories), despite its relatively simpler feature representation.

We also compared the performance obtained from varying the salient point extraction method. The average AUC values (across all categories) of SIVAL are 81.6 for SPARSE+VSWN, 80.3 for SPARSE, 77.0 for VSWN, 73.5 for Wavelet, and 77.9 for SIFT. Both SPARSE and VSWN can help improve retrieval performance, and when used together, they improve performance further.

Fig. 6 independently compares the effect of using SPARSE and VSWN with the standard wavelet-based salient points on SIVAL. Adding VSWN to Wavelet leads to a 4.8 percent improvement when averaged over all categories with statistically significant improvements in nine categories. Adding SPARSE to Wavelet leads to a 9.3 percent improve- ment when averaged over all categories. When both SPARSE and VSWN are added to Wavelet, an 11 percent increase in performance occurs.

3 Conclusions And Future Work

We have presented ACCIO!, a localized CBIR system that does not assume that the desired object(s) are in a fixed location or have a fixed size. We have demonstrated that ACCIO! outperforms existing systems for localized CBIR on both a natural scenes image repository and SIVAL, our new benchmark data set. Our experimental results when using the Flickr data set demonstrate that ACCIO! can successfully be used for real image retrieval problems where the user is interested in general object categories.

We introduce the SPARSE technique, which uses segmentation as a filter to reduce the total number of salient points while still maintaining diversity. Finally, we introduce the VSWN salient point representation, which splits salient points on region boundaries into two salient points, characterizing the separate objects at the boundary. There are many directions for future work.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

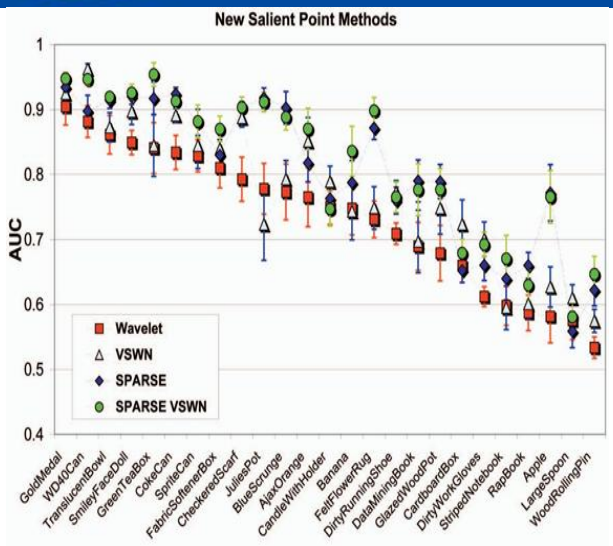

Fig. 6. Comparing salient point methods on SIVAL data set for a query set of eight positive and eight negative examples.

We believe that ACCIO! can be improved further by making further improve- ments to GEM-DD, by employing improved segmentation algorithms, and perhaps by the careful introduction of some features to represent shape. For SPARSE, an important area of future work is to perform experiments to determine the sensitivity to k, the number of salient points per segment, and develop methods to select the best value for k.

REFERENCES

[1] S. Andrews, T. Hofmann, and I. Tsochantaridis, "Multiple Instance Learning with Generalized Support Vector Machines," Artificial Intelligence, pp. 943-944, 2002.

[2] K. Barnard, P. Duygulu, N. de Freitas, and D. Forsyth, "Matching Words and Pictures," J. Machine Learning Research, vol. 3, pp. 1107- 1135, 2003.

[3] J. Bi, Y. Chen, and J. Wang, "A Sparse Support Vector Machine Approach to Region-Based Image Categorization," Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR '05), pp. 1121- 1128, 2005.

[4] Y. Chen and J. Wang, "Image Categorization by Learning and Reasoning with Regions," J. Machine Learning Research, pp. 913- 939, 2004.

[5] I. Cox, M. Miller, S. Omohundro, and P. Yianilos, "PicHunter: Bayesian Relevance Feedback," Proc. Int'l Conf. Pattern Recognition (ICPR '96), pp. 361-369, 1996.

[6] A. Dempster, N. Laird, and D. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm," J. Royal Statistics Soc., vol. 39, pp. 1-38, 1977.

[7] T. Dietterich, R. Lathrop, and T. Lozano-Pe´rez, "Solving the Multiple-Instance Problem with Axis-Parallel Rectangles," Artifi- cial Intelligence, vol. 89, nos. 1-2, pp. 31-37, 1997.

[8] P. Duygulu, K. Barnard, N. de Freitas, and D. Forsyth, "Object Recognition as Machine Translation: Learning a Lexicon for a Fixed Image Vocabulary," Proc. Seventh European Conf. Computer Vision (ECCV '02), pp. IV:97-IV:12, 2002.

[9] M. Fischler and R. Bolles, "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography," Comm. ACM, vol. 24, pp. 381-395, 1981.

[10] D. Forsyth and J. Ponce, Computer Vision: A Modern Approach.
Prentice Hall, 2003.

[11] T. Gevers and A. Smeulders, "Image Search Engines: An Over- view," Emerging Topics in Computer Vision, G. Medioni and S.B. Kang, eds., Prentice Hall, 2004.

[12] J. Hanley and B. McNeil, "The Meaning and Use of the Area under a

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Receiver Operating Characteristic (ROC) Curve," Radiology, vol. 143, no. 1, pp. 29-36, 1982.

[13] X. Huang, S.-C. Chen, M.-L. Shyu, and C. Zhang, "User Concept Pattern Discovery Using Relevance Feedback and Multiple Instance Learning for Content-Based Image Retrieval," Proc. Eighth Int'l Conf. Knowledge Discovery and Data Mining (KDD '02), pp. 100-108, 2002.

[14] J. Wang, J. Li, and G. Wiederhold, "IRM: Integrated Region Matching for Image Retrieval," Proc. Eighth ACM Int'l Conf. Multimedia, pp. 147-156, 2000.

[15] Y. Ke and R. Sukthankar, "PCA-SIFT: A More Distinctive Representation for Local Image Descriptors," Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR '04), vol. 2, pp. 506- 513, 2004.

[16] L. Ledwich and S. Williams, "Reduced Sift Features for Image Retrieval and Indoor Localisation," Proc. Australasian Conf. Robotics and Automation (ACRA), 2004.

[17] H.-K. Lee and Y.-S. Ho, "A Region-Based Image Retrieval System Using Salient Point Extraction and Image Segmentation," Proc. Third IEEE Pacific Rim Conf. Multimedia (PCM '02), pp. 209-216, 2002.

[18] F. Long, H. Zhang, and D. Feng, "Fundamentals of Content-Based Image Retrieval," Multimedia Information Retrieval and Manage- ment-Technological Fundamentals and Applications, D. Feng, W. Siu, and H. Zhang, eds., Springer, 2003.

[19] E. Loupias, Salient Points Detection Using Wavelet Transform, http://telesun.insa-lyon.fr/~loupias/points/demo.html, 2008.

[20] D.G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," Int'l J. Computer Vision, vol. 60, no. 2, pp. 91-110, 2004.

[21] O. Maron and T. Lozano-Pe´rez, "A Framework for Multiple- Instance Learning," Proc. Conf. Advances in Neural Information Processing Systems (NIPS '98), pp. 570-576, 1998.

[22] O. Maron and A. Ratan, "Multiple-Instance Learning for Natural Scene Classification," Proc. 15th Int'l Conf. Machine Learning (ICML '98), pp. 341-349, 1998.

[23] K. Mikolajczyk and C. Schmid, "A Performance Evaluation of Local Descriptors," Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR '03), pp. 257-264, 2003.

[24] R. Rahmani, S. Goldman, H. Zhang, J. Krettek, and J. Fritts, "Localized Content-Based Image Retrieval," Proc. Seventh ACM SIGMM Int'l Workshop Multimedia Information Retrieval (MIR '05), pp. 227-236, 2005.

[25] Y. Rui, T. Huang, M. Ortega, and S. Mehrotra, "Relevance Feedback: A Power Tool in Interactive Content-Based Image Retrieval," IEEE Trans. Circuits and Systems for Video Technology, vol. 8, no. 5, pp. 644-655, 1998.

[26] C. Schmid, R. Mohr, and C. Bauckhage, "Evaluation of Interest Point Detectors," Int'l J. Computer Vision, vol. 37, no. 2, pp. 151-172, 2000.

[27] Q. Tian, N. Sebe, M.S. Lew, E. Loupias, and T.S. Huang, "Image Retrieval Using Wavelet-Based Salient Points," J. Electronic Imaging, special issue on Storage and Retrieval of Digital Media, vol. 10, no. 4, pp. 849-935, 2001.

[28] Q. Tian, Y. Wu, and T. Huang, "Combine User Defined Region-of- Interest

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

and Spatial Layout for Image Retrieval," Proc. IEEE Conf. Image Processing (ICIP '00), vol. 3, pp. 746-749, 2000.

[29] P. Torr and A. Zisserman, "Mlesac: A New Robust Estimator with Application to Estimating Image Geometry," Computer Vision and Image Understanding, vol. 78, no. 1, pp. 138-156, 2000.

[30] J. Wang, J. Li, and G. Wiederhold, "SIMPLIcity: Semantics- Sensitive Integrated Matching for Picture Libraries," IEEE Trans. Pattern Analysis and Machine Intelligence, pp. 947-963, 2001.

[31] J. Wang, H. Zha, and R. Cipolla, "Combining Interest Points and Edges for Content-Based Image Retrieval," Proc. IEEE Int'l Conf. Image Processing (ICIP '05), pp. 1256-1259, 2005.

[32] C. Wolf, J.-M. Jolion, W. Kropatsch, and H. Bischof, "Content Based Image Retrieval Using Interest Points and Texture Features," Proc. 15th IEEE Int'l Conf. Pattern Recognition (ICPR '00), vol. 4, p. 4234, 2000.

[33] C. Yang and T. Lozano-Pe´rez, "Image Database Retrieval with Multiple Instance Techniques," Proc. 16th Int'l Conf. Data Eng. (ICDE '00), pp. 233-243, 2000.

[34] H. Zhang, R. Rahmani, S. Cholleti, and S. Goldman, "Local Image Representations Using Pruned Salient Points with Applications to CBIR," Proc. 14th ACM Int'l Conf. Multimedia, pp. 287-296, 2006.

[35] H. Zhang, J. Fritts, and S. Goldman, "An Improved Fine-Grain Hierarchical Method of Image Segmentation," technical report, Washington Univ., 2005.

[36] H. Zhang, J. Fritts, and S. Goldman, "A Fast Texture Feature Extraction Method for Region-Based Image Segmentation," Proc. IS&T/SPIE's 16th Ann. Symp. Image and Video Comm. and Processing, pp. 957-968, 2005.

[37] Q. Zhang and S. Goldman, "EM-DD: An Improved Multiple- Instance Learning Technique," Neural Information Processing Systems, pp. 1073-1080, 2001.

[38] Q. Zhang, S. Goldman, W. Yu, and J. Fritts, "Content-Based Image Retrieval Using Multiple Instance Learning," Proc. 19th Int'l Conf. Machine Learning (ICML '02), pp. 682-689, 2002.

[39] Z. Zhou and M. Zhang, "Ensembles of Multi-Instance Learners," Proc. 14th European Conf. Machine Learning (ECML '03), pp. 492-502, 2003.