IJCSC 0973-7391 Volume 7 • Number 2 March 2016 - Sept 2016 pp. 140-147 available online at www.csjournalss.com

Satellite Image Retrieval by Fusion of Morphological and LBP Features

Swati Jain, Shailee Patel
Institute of Technology, Nirma University, Ahmedabad

Abstract: Morphological and Local Binary patterns are amongst the preferred Feature Sets for Content based image retrieval of satellite images in visible domain. Although each of these feature sets individually has limited capability in terms of defining the image content. In the proposed work, we have enhanced the retrieval results, using combination of the two i.e. CCH (circular co variance histogram and LBP (Local Binary Pattern), for a given query image, by minimizing the instances where non relevant images are assignedhigher ranks than relevant images, thus weeding them out. This minimization process results into a weight matrix, which is used for obtaining the combined distance of the query image with database images. This combined distance is finally used for ranking the images. The proposed algorithm is experimented with UC Merced LULC image data set. The retrieval results obtained using proposed algorithm is compared with results obtained using individual feature set and also with results obtained using combined feature sets without any weight matrix. The experimentation was performed exhaustively and that demonstrates considerable improvement in the results.

Keywords: Satellite Images • Late Fusion • Early Fusion • Image Retrieval.

1. Introduction

Images have taken a very important place in our life, ubiquitous presence of camera have accelerated the rateof increase of multimedia data like never before. Also in case of exclusive domain, such as scientific imaging and medical imaging the image capturing rate has gone up due to technological advancement. Developing smart image retrieval solutions have attracted lot of attention in all the domains including Satellite Images. Content descriptors commonly known as feature sets are crucial part of the multimedia retrieval systems as they are expected to extract higher as well as lower level content or information of the images.

Over the years various features are defined and designed to capture distinguishing characteristics in an image and enabling content based image retrieval (CBIR) system. These features can be categorised into color, texture, and edge. . Different color features are color coherence vector (CCV)[13], color moment[7], color histogram etc. Edge features concentrate on edges of the local regions. They are more useful for the application that demands the task of object detection. Examples of edge features are edge direction histogram and edge coherence vector[4]. Gabor filters[16] and co-occurrence matrix[5] are most referred and cited texture descriptors. Both these texture descriptors are time tested in gray scale images. One more relatively new category of image descriptors are Morphological Texture. Erchan [1,2] in his publications introduced morphological texture descriptors and later used them to describe satellite images. Morphological covariance as operator is used, in order to find textures. Circular Covariance Histogram (CCH) and Rotation Invariant Point Triplets (RIT) are morphological texture descriptors [2]. The process to calculate the morphological feature set is extremely compute intensive and hence that makes the feature extraction activity extended in terms of time.[8] shows how the parallel implementation can lead to considerable gain in computation time and hence resulting faster feature extraction. In the class of texture based features there are few more feature sets often used in defining images including satellite images,like Local Binary Patterns (LBP)[3] and Local Tetra Pattern (LTrP)[11]. Both the feature sets captures the relationship amongst the neighbouring pixels.

Majorly two approaches in feature fusion prevail, namely Early Fusion and Late Fusion. Early fusion combines the distance score of different feature so also known as feature fusion and, late fusion is termed as score fusion as it combines the class scores[15].

Satellite Images are normally huge in size, taken frequently in various spectrums, so its rate of increase is also high. Authors have proposed several feature sets each capturing a particular dimension of the information in the image. An efficient way to combine these features will facilitate a holistic comparison of the image content. The experimentation of fusion using proposed algorithm is done on the UC Merced Land Use Land Cover (LULC) dataset[14] and retrieval results are compared when features are plainly combined or when only one of the features are considered.

2. Proposed Work

We propose a weight learning algorithm that combines the Morphological and LBP features appropriately to improve the retrieval results. Majorly the feature fusion is employed for image classification with intent to ensure that the cases of true positive are maximized. Here in the proposed work, feature fusion is employed to minimize the cases where distance of irrelevant images is less than the distance of relevant images.

Formulation of the problem of finding the appropriate fusion weights is discussed as below. The final distance vector which is combination of all the feature set is obtained by the equation 3. Where W is the vector that is to be learnt, and then final score is obtained as weighted (w) sum of different distances obtained using different feature set. In [10] score fusion is demonstrated, that maps all the distances in one scale. In the given algorithm late fusion is used and hence score values, obtained by classifier, where higher values signifies higher degree of similarity, is used. In our proposed work we have used distance vector instead of scores, and higher values signifies smaller degree of similarity. Let all the images which are relevant be Rj and all the images non relevant as Ri . The idea is to minimize all the cases where the irrelevant image has less distance measure in comparison to relevant images hence trying to minimize them.

$$min_w \frac{1}{2}\|w\|^2 + \sum((w.(R_j - R_i)), s.t. w_n \geq 0 \quad (1)$$

The key idea is to add the distances between query image and all other images, using different features. All these calculated distances are in different scale. The objective is to map all the feature vectors into a common score space. Given a distance measure obtained by a particular feature set and w is the learned weight.

$$D_j = \sum_{j=1}^{n} d_j^j . w_i^j \quad (2)$$

In eq 2, Di is the sum of the product of distance of ith image w.r.t the query image and their respective jth image using weights. Distance d_j^i (distance for the ith image andjth feature), n is the number of feature sets considered. For a given feature set it is the vector of N (N is the j number of images in the dataset). This di for different values of j is not in the same scale and range, but in all the feature sets smaller distance value signifies higher degree of similarity to the query Image. As the score values are incomparable, fusion cannot be done directly. So to calculate the final distance $D_j = \sum_{j=1}^{n} d_j^j . w_i^j$ is not appropriate, hence equation (1) is used. To calculate the weight matrixw_j^iwe consider Si as the set of images which are relevant and Sj as the set of non-relevant images. We minimize the cases where irrelevant images have distances less in comparison to relevant images with the query image.

Distance of all the images is calculated with respect to the query image, using one feature set at a time, and all the distances are added to find the final distance value. Pair i,j is marked, where i belongs to set relevant images and j belongsto the set of irrelevant images, such that the distance calculated for irrelevant images is less than the distance calculated for relevant images. The set I contains all such marked pairs. Then j we calculate a pair wise comparative matrix zi which is the difference of the Euclidean distance of images in set I. Since the database used is classified images hence all images belonging to the same class as that of query image is considered as relevant and all other images are considered irrelevant. It is expected that distance for relevant images is smaller than the distance obtained for irrelevant, and all the cases that violate are marked in I. Algorithm 1 is the variant of algorithm proposed in [10] where score fusion is replaced by distance measure.

Algorithm 1 describes the steps to calculate the weights for the distance scores. The algorithm is based on modified Newton method[10,6,9].

In step 5 equation is solved linearly to obtain w. Size of I and hence Z is MxM where M is the number of image in the dataset. Line 6 removes all the negative values and hence results into negation of false negative values of the feasible set. The next two lines obtain value of wt+1. The loop is repeated n times n is found empirically such that the value of w becomes close to constant.

Algorithm 1 : Calculation of weight

1: *Initialize Array:*w^t*=1*
2: *Repeat n times, step 3 to 8*
3: $I^t \leftarrow \{(i,j) \mid w^t(x_i - x_j) > 0\}$
4:$Z_{i,j} \leftarrow x_i - x_j, \forall (i,j) \in I^t$
5: $\overline{w} \leftarrow w \mid \sum_{i,j \in I^n} (I + 2Z_{ij}^T Z_{ij})\mathrm{w} = 0$
6: $\overline{w}$ $\leftarrow$ all positive values of w
7: $\alpha_t \leftarrow argmin_{0 \geq \alpha \geq 1} f(w^t + \alpha(\overline{w} + w^t))$ (finding alpha between 0-1 in an interval of 0.2)
8: $w^{t+1} \leftarrow w^t + \alpha(\overline{w} - w^t)$
9: Output: w^{t+1}

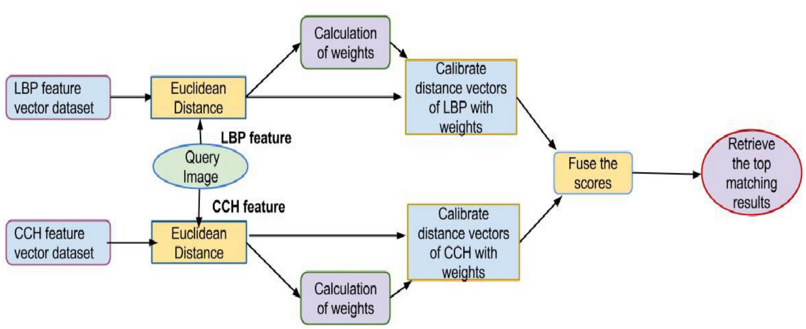

Fig. 1 Flow of Implementation

IJCSC 0973-7391 Volume 7 • Number 2 March 2016 - Sept 2016 pp. 140-147 available online at www.csjournalss.com

3. Performance Measure

Gain Ratio:The gain ratio is obtained as:

$$Gainratio = \frac{n_1}{N} / \frac{n_2}{N} (3)$$

where N is the total number of images retrieved, n1 is the number of relevant images retrieved out of total retrieved images when the fusion algorithm is applied; while n2 is the relevant images retrieved out of N images by just adding the Euclidean distances of two features for a given query image.

Rank: We also used normalized average rank as a performance measure as proposed in [12]. In case of gain ratio the scenario beyond the N cannot be captured. For instance, when N=30, this will not be able to capture when all the relevant images are ranked just from 31^{st}. Hence normalized rank is considered for performance evaluation, which tries to find the average of ranks given to all the relevant images in the data set and trying to capture the deviation from the ideal condition. Ideal condition is defined as, when all N_r relevant images are ranked from 1 to N_r by the retrieval algorithm. The value obtained is normalized, hence scaling it between 0-1. Value 0 meaning no deviation from the ideal condition hence smaller value represents better performance. The normalized average rank is given as below.

$$Rank = \left(\frac{1}{NN_R}\right) \sum_{i=1}^{N_R} R_i - \frac{N_R(N_R - 1)}{2}$$

R_i is the rank at which the i^{th} relevant image is retrieved. Value 0 represents perfect performance 0 and as it approaches 1 it signifies the performance degradation. Table 2 describes the average rank values obtained by evaluating the performance of fusion algorithm.

3.1 Results

The database has total 2100 images, 21 class, with 100 images of each class. All the 2100 images are considered as query image one at a time and the retrieval performance over the class is shown in the results. Table 1 summarizes the performance in terms of GainRatio and Table 2 represents the average rank value. The results are representing the number of relevant (belonging to the same class as the query image in the database) images out of top 30 retrieved images.

The results of the GainRatio obtained are shown in Table 1. The table shows relevant images out of top 30 images retrieved when only mentioned feature is used, and later when features are combined with weights and without weights in the following columns. An exhaustive experimentation is performed to ensure that the conclusions drawn are accurate. We took all 2100 image in the database as a query image one by one, and then averaging is performed in each class to obtain the average number of images retrieved in one class. The table contains average value of each category for all 21 categories of LULC data set. The next three columns represents the average number of relevant images from the top 30 retrieved images using CCH, and LBP feature sets for retrieval. The next heading CCH + LBP summarizes the retrieval result, when considering CCH+LBP features without weights, with weights and then gain ratio obtained when compared with result of without weights and results obtained using individual features. Without weight means the number of relevant images retrieved when the Euclidean distance of CCH and LBP are combined to get the results while with weights means the relevantimages retrieved when the proposed fusion algorithm is employed to learn the weights. Looking at the value of the gain ratio obtained we can conclude clearly that the algorithm shows substantial increase in performance in almost every category, compared to when individual features are used and when distance averaged out without weight. The gain goes as high as 4.4 meaning increase in number of retrieved image is four times or 300%. It also shows results as low as only 1.1 times in case of Agriculture, Forest and Chaparral, on investigating the images it was observed that the images of Agriculture and Forest are very similar visually, though annotated as different class. The count in table strictly represents the number of images belonging to same class as defined by the data set and no visible similarity across the annotated class is considered. In case of Chaparral the count is already very high so less scope of improvement.

Table 1 Gain Ratio of precision values in top 30 retrieved Images

Category wise Generalized Fusion retrieval results							
	CCH	**LBP**	**CCH+LBP**				
Categories			**Without Weights**	**With Weights**	**Gain w.r.t without weights**	**Gain w.r.t CCH**	**Gain w.r.t LBP**
Agriculture	14	10	10	13	1.3	0.9	1.3
Airplane	9	11	8	14	1.8	1.6	1.3
Baseball	5	8	7	19	2.7	3.8	2.4
Beach	6	13	13	23	1.8	3.83	1.8
Building	5	7	6	13	2.2	2.6	1.9
Chaparral	22	23	21	23	1.1	1.1	1
Dense Residential	6	9	8	18	2.2	3	2
Forest	18	18	18	21	1.2	1.2	1.2
Freeway	5	11	10	17	1.7	3.4	1.5
Golf course	8	11	7	21	3	2.6	1.9
Harbor	14	12	12	15	1.3	1.1	1.25
Intersection	6	7	6	18	3	3	2.6
Medium Residential	8	10	9	23	2.6	2.9	2.3
Mobile home park	8	12	11	19	1.7	2.4	1.6
Overpass	8	6	8	19	2.4	2.4	3.2
Parking lot	12	9	10	16	1.6	1.3	1.8
River	9	8	7	19	2.7	2.1	2.4
Runway	10	18	16	17	1.1	1.7	0.9
Sparse Residential	7	10	9	14	1.6	2	1.4
Storage Tanks	6	7	7	17	2.4	2.8	2.4
Tennis Court	6	7	7	17	2.4	2.8	2.4

IJCSC 0973-7391 Volume 7 • Number 2 March 2016 - Sept 2016 pp. 140-147 available online at www.csjournalss.com

The main aim of feature fusion is to combine different feature sets in such a way that the resultant distance measure can characterize the image more efficiently. Hence, it comes intuitively to combine features belonging to different categories in order to get more efficient fusion results. This is the reason for taking the combination of CCH and LBP as they belong to different class of the feature.

Table 2 contains average ranks calculated for both the situations, first, when distances obtained by two features are simply added and second, when added using equation 2, and w is learned using Algorithm 1. Looking at table it can be concluded that the algorithm shows improvement in value of ranks.

There are few exceptions which can be observed in the CCH, LBP combination where the rank obtained with fusion is higher compared to without weights. Such exceptions are observed for the categories like Airplane, Dense Residential, Medium Residential, Spares Residential and TennisCourt. These five categories show ranks higher values of rank but when there retrieval is analysed we can see that the images obtained in retrieval are visibly similar to the query image.

Table 2 Performance evaluation of two feature combinations (CCH and LBP) using normalized average rank

	Rank CCH-LBP	
Category	**Without weights**	**Withweights**
Agriculture	0.6566	0.5792
Airplane	0.31	0.3374
Baseball diamond	0.3755	0.3445
Beach	0.3054	0.284
Building	0.4183	0.3624
Chaparral	0.091	0.076
Dense Residential	0.3029	0.3031
Forest	0.1896	0.159
Freeway	0.6409	0.356
Golf course	0.3247	0.2946
Harbor	0.4622	0.4061
Intersection	0.3235	0.3157
Medium Residential	0.2446	0.2478
Mobile HomePark	0.3802	0.3493
Overpass	0.4922	0.2911
Parking Lot	0.2921	0.2379
River	0.2944	0.1928
Runway	0.4685	0.3412
Sparse Residential	0.3126	0.3412
Storage Tanks	0.3385	0.2791
Tennis Court	0.3069	0.3577

The retrieved images are annotated under different categories but their visual similarity is more, for instance if images from dense residential and sparse residential are to be differentiated with human intervention they are sure to be varying. Our count of relevant images strictly is restricted to annotated class as query image, and not considering visible similarity across the class. For example if we take the first image from dense residential category as shown in fig 1as a query image, then negative performance is obtained. By just adding Satellite Image Retrieval by Fusion of Morphological and LBP Features the Euclidean distances of CCH and LBP the rank obtained is 0.232486 while after applying fusion the rank increases to 0.246195 which because of lot of interference in dense residential, building and sparse residential images.

4. Conclusion

In this paper, a weighted fusion scheme is presented, for Morphological and LBP features distance measures, to enhance the performance of the retrieval results of the satellite images. The presented work effectively demonstrates that how feature set representing image content of different perspective can be combined effectively to give retrieval a holistic perception. The fusion results obtained shows substantial growth in the relevant retrieved images compared to the individual feature performance or averaging approach of fusion. A considerable increase in the number of relevant images is obtained in top 30 positions. Improvement in the average normalized rank is also obtained which considers the rank of all the relevant images in the database. Deviation in performances in some instances demonstrates the inter class visible similarity amongst images in few classes of the images. The conclusions drawn are based on exhaustive experimentation and considering all 2100 images as query and averaging the performance over a class. The experimentation are done on the recently proposed and published feature set like CCH, LBP and LTrP. Similar experiments can be performed for various other class of the features as well. Also the distance metric used is Euclidean distance for simplicity, but a more appropriate distance measure can be explored.

References

[1] Aptoula, E.: Extending morphological covariance. *Pattern Recognition* 45(12), 4524–4535 (2012)

[2] Aptoula, E.: Remote sensing image retrieval with global morphological texture descriptors. *Geoscience and Remote Sensing, IEEE Transactions on* 52(5), 3023–3034 (2014)

[3] Banerji, S., Verma, A., Liu, C.: Novel colorlbp descriptors for scene and image texture classification. *In 15th International Conference on Image Processing, Computer Vision, and Pattern Recognition, Las Vegas, Nevada, pp*. 537–543. Citeseer (2011)

[4] Gao, X., Xiao, B., Tao, D., Li, X.: Image categorization: Graph edit direction histogram. *Pattern Recognition* 41(10), 3179 – 3191 (2008)

[5] Haralick, R.M.: Statistical and structural approaches to texture. *Proceedings of the IEEE* 67(5), 786–804 (1979)

[6] Keerthi, S.S., DeCoste, D.: A modified finite newton method for fast solution of large scale linear svms. *In: Journal of Machine Learning Research, pp*. 341–361 (2005)

[7] Kodituwakku, S., Selvarajah, S.: Comparison of color features for image retrieval. *Indian Journal of Computer Science and Engineering* 1(3), 207–211 (2004)

[8] Kumar, S., Jain, S., Zaveri, T.: Parallel approach to expedite morphological feature extraction of remote sensing images for cbir system. *In: Geoscience and Remote Sensing Symposium (IGARSS), 2014 IEEE International, pp*. 2471–2474 (2014)

[9] Lin, C.J., Weng, R.C., Keerthi, S.S.: Trust region newton method for logistic regression. The Journal of Machine Learning Research 9, 627–650 (2008)

[10] Ma, T., Oh, S., Perera, A., Latecki, L.: Learning nonlinear calibration for score fusion with applications to image and video classification. *In: Computer Vision Workshops (ICCVW), 2013 IEEE International Conference on, pp*. 323–330 (2013)

[11] Murala, S., Maheshwari, R., Balasubramanian, R.: Local tetra patterns: a new feature descriptor for content-based image retrieval. *Image Processing, IEEE Transactions on* 21(5), 2874–2886 (2012)

[12] Mller, H., Mller, W., Squire, D.M., Marchand-Maillet, S., Pun, T.: Performance evaluation in content-based image retrieval: overview and proposals. *Pattern Recognition Letters* 22(5), 593 – 601 (2001)

[13] Pass, G., Zabih, R., Miller, J.: Comparing images using color coherence vectors. *In: Proceedings of the fourth ACM international conference on Multimedia, pp.* 65–73. ACM (1997)

[14] Yang, Y., Newsam, S.: Bag-of-visual-words and spatial extensions for land-use classification. *In: Proceedings ofthe 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems, GIS '10, pp*. 270–279. ACM, New York, NY, USA (2010)

[15] Ye, G., Liu, D., Jhuo, I.H., Chang, S.F.: Robust late fusion with rank minimization. *In: Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on, pp.* 3021–3028. IEEE (2012)

[16] Zhang, D., Wong, A., Indrawan, M., Lu, G.: Contentbased image retrieval using gabor texture features. *In: IEEE Pacific-Rim Conference on Multimedia, University of Sydney, Australia* (2000).