Evaluation of Bag-of-Colors Descriptor for Land Use Classification

Vedran Jovanović and Vladimir Risojević, *Member, IEEE*

Abstract — **The objective of this paper is to evaluate Bag-of-Colors (BoC) descriptor for land use classification. BoC can be used either as a global or local descriptor. In this paper we present and evaluate both approaches. We analyze the influence of different parameters on classification accuracy and introduce a modification of descriptor extraction process, which significantly influences the classification results and performance.**

***Keywords* — color descriptors, BoC descriptor, Land Use Classification**

I. Introduction

In this paper we consider the problem of land use classification in high-resolution overhead imagery. The main objective of this work is evaluation of recently proposed bag-of-colors (BoC) descriptor [1] for land use classification.

A lot of papers have been written in the field of descriptor computation. Many of them are based on using of local descriptors, in particular Scale Invariant Feature Transform (SIFT) descriptors and its variants [2], [3]. SIFT as a local descriptor is often used in bag-of-visual-words (BoVW) framework, where we get a single fixed-size vector from a set of local descriptors. But most of them does not take the color information into account, and only works with gray-level images.

Color information is a very expressive visual feature, and very important in recognition and detection of different objects. It can be very useful in land use classification, where different types of land cover have different colors (forest and grass are green, water is blue, etc.). One of the first works using color as a visual feature for image representation is [4], where color histogram is introduced. Color descriptors are also included in MPEG-7 standard for description of audio-visual content [5].

The main contributions of this paper are the evaluation of BoC descriptor for land use classification, analysis of the influence of different parameters on classification accuracy and a modification of descriptor extraction process resulting in more effective and efficient descriptor.

V. Jovanović, University of Banja Luka. Faculty of Electrical Engineering, Patre 5, 71000 Banja Luka, Bosnia and Herzegovina (phone: +387-51-221-874, e-mail: vedran.jovanovic@etfbl.net).

V. Risojević, University of Banja Luka, Faculty of Electrical Engineering, Patre 5, 71000 Banja Luka, Bosnia and Herzegovina (phone: +387-51-221-847, e-mail: vlado@etfbl.net).

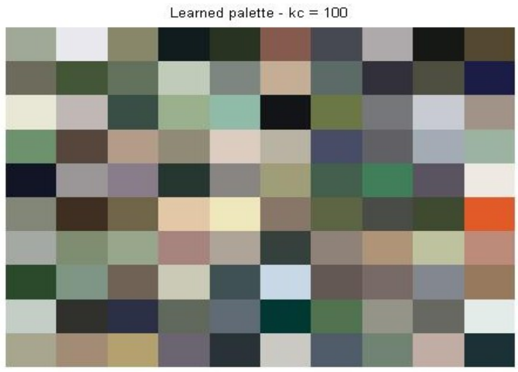

Fig. 1. Color palette obtained after learning on a dataset of aerial images ($N = 8$, $k_c = 100$)

This paper is organized as follows. In Section 2, we review the global and local BoC descriptors. In Section 3, we the used dataset and methodology of testing, and present the experimental results. Section 4 is the conclusion.

II. Bag-of-Colors descriptor

In this paragraph, we describe the process of generating BoC descriptor. BoC can be used either as a global or local descriptor. Global descriptor is computed for the whole image, while local descriptor is computed for a set of image patches.

A. Global BoC descriptor

The first step in the process of computing the descriptor is to generate a codebook of colors (*color palette)*. Color palette is generated in the same way as described in [1]:

1. Divide the dataset into training and test sets,
2. Convert all images into CIE-Lab color space,
3. Quantize each of the L, a and b color components uniformly into N bins,
4. Divide the images into blocks of size 16×16 pixels,
5. For each block find the most frequently occurring color. If there is more than one color corresponding to the maximal number of occurrences, randomly choose a color from the block. After this, each image from the training set is described with 256 colors.
6. Final step in the process of color palette generation is clustering of all colors obtained from the training set images, using k-means algorithm, into a set of k_c colors.

Now, we have learned a color palette which is more adjusted to real-world images. An example of the learned color palette learned is shown in Fig. 1. We can see that colors such as green and brown have bigger contribution in

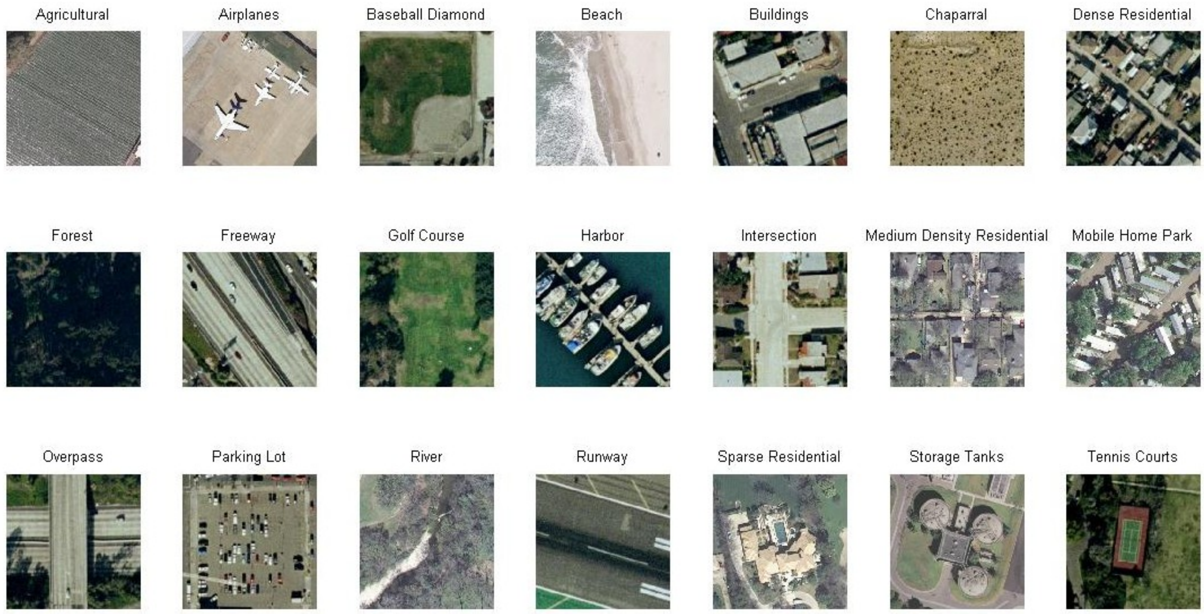

Fig. 2. Examples of images from the UC Merced Land Use Dataset

the palette than red, blue or yellow. This is a consequence of unbalanced color statistic of aerial images, where, for example, green color (parks, trees, forest,...) is more probable than red (rooftops).

After we learn the palette, the next step is computation of the BoC descriptor. In order to compute the global BoC descriptor for an image, we compute the color histogram of the image with respect to the palette. This is achieved by computing the Euclidean distance of the color of each image pixel from the colors in the palette, whereupon we increment the histogram bin corresponding to the closest palette color. The obtained descriptor is k_c-dimensional.

B. Local BoC descriptor

In order to obtain the local BoC descriptor, initial procedure is similar to the process of generating a global descriptor. The first step is generation of a palette, and it is computed in the same way as described before. In the second step, we compute the histograms of the patches from all images, with respect to the palette. These patches can be obtained in many ways. They can be extracted by interest point detector, regularly sampled, etc. In this paper, the patches are regularly sampled from the image. Now, each image is represented with a bag of local histograms. After this, the set of local histograms gathered from training images is clustered using k-means algorithm, producing the codebook with M elements.

The last step is the computation of the histogram of codeword occurrences. For each local histogram, we compute the Euclidean distance between the local histogram and each of the codewords and increment the bin corresponding to the closest codeword. At the end of this process each image is represented with a single vector of length M.

III. Experiments

In this part of paper, we describe the dataset on which we tested our classifier, methodology of testing, and the results we obtained.

A. Dataset and methodology

Dataset we chose for testing of BoC descriptor is UC Merced Land Use (UCMLU) dataset, which consists of 2100 images, manually divided into 21 classes, with 100 images in each. The classes correspond to different land cover and land use types: agricultural, airplane, baseball diamond, beach, buildings, chaparral, dense residential, forest, freeway, golf course, harbor, intersection, medium density residential, mobile home park, overpass, parking lot, river, runway, sparse residential, storage tanks and tennis courts. The examples of images from all classes are shown in Fig. 2. Resolution of each image is 30 cm per pixel, and sizes of all images are 256×256 pixels.

We compute image descriptors as described in the previous section. Since both global and local BoC descriptors are essentially histograms we experimented with different normalization techniques, e.g. inverse-document frequency (IDF), power-law, L1 and L2 normalizations. We obtained the best classification results using only L1 normalization, which was therefore used in the experiments reported in this paper.

For classification we use Support Vector Machine (SVM) classifier. The aim of SVM is to construct the decision hyperplane in high-dimensional space such that the distance to the nearest training data point of any class is maximal. In order to benefit from using a nonlinear kernel but still maintain the efficiency of using linear SVM, we perform use χ^2 kernel mapping of the data and use linear SVM [6].

TABLE 1: CLASSIFICATION ACCURACIES (%) OBTAINED USING GLOBAL BOC DESCRIPTOR

		N – Number of quantization levels per color component						
		4	8	16	32	64	128	256
k_c – Palette size	50	65,2 ± 1,4	69,4 ± 2,0	71,6 ± 1,1	72,8 ± 1,5	73,1 ± 1,7	74,0 ± 2,0	73,4 ± 1,7
	100	63,3 ± 2,8	72,2 ± 2,0	74,1 ± 1,5	76,7 ± 1,5	76,0 ± 1,9	75,9 ± 2,1	77,1 ± 1,7
	500	63,9 ± 2,1	73,9 ± 1,4	80,0 ± 2,1	82,0 ± 2,1	83,8 ± 1,6	84,0 ± 1,9	84,5 ± 1,8
	1k	63,8 ± 1,7	72,5 ± 1,9	81,5 ± 2,6	84,4 ± 1,9	85,5 ± 1,9	85,7 ± 2,5	85,8 ± 1,8
	2k	64,9 ± 2,8	72,4 ± 2,1	82,1 ± 2,6	85,6 ± 2,1	85,9 ± 1,7	85,7 ± 1,7	85,7 ± 1,8
	5k	67,0 ± 2,0	71,6 ± 2,6	81,5 ± 1,8	85,3 ± 1,8	**86,5 ± 1,8**	85,9 ± 2,2	86,2 ± 1,6

TABLE 2: CLASSIFICATION ACCURACIES (%) OBTAINED USING MODIFIED GLOBAL BOC DESCRIPTOR

		N – Number of quantization levels per color component						
		4	8	16	32	64	128	256
k_c – Palette size	50	72,8 ± 1,5	72,2 ± 1,9	73,4 ± 1,5	72,9 ± 1,6	72,6 ± 1,3	74,5 ± 2,4	73,6 ± 1,7
	100	77,1 ± 2,1	75,8 ± 1,9	76,5 ± 1,7	76,3 ± 1,9	77,3 ± 1,9	77,1 ± 2,1	76,1 ± 2,5
	500	84,0 ± 2,2	84,5 ± 2,5	84,3 ± 2,2	83,2 ± 1,6	83,8 ± 1,7	84,3 ± 2,0	85,0 ± 2,1
	1k	85,4 ± 2,0	85,0 ± 2,5	85,2 ± 2,3	85,0 ± 1,8	85,1 ± 1,9	85,5 ± 2,1	84,8 ± 2,1
	2k	85,8 ± 1,6	86,1 ± 2,0	85,7 ± 1,7	85,8 ± 1,7	85,7 ± 1,7	86,0 ± 1,5	85,9 ± 1,9
	5k	86,3 ± 1,9	86,5 ± 1,9	86,0 ± 1,7	86,2 ± 1,5	86,2 ± 1,5	**86,5 ± 1,7**	86,1 ± 1,7

We randomly split images from each class into training and test sets, where 80% of them were assigned to training and 20% to test set. Now, we get 1680 training images and 420 test images. After that, we extract the palette (for global) and codebook of histograms (for local descriptor) using only training images, whereupon we calculate BoC descriptor as final representation of images. Then, we normalize descriptors using L1 normalization. We use ten different random training/test set splits, and report mean values and standard deviations of obtained classification accuracies.

B. Global descriptor

The results obtained using global BoC descriptor on UCMLU dataset are shown in Table 1. In order to determine how the descriptor parameters affect the accuracy of classification, we varied the number of quantization levels per color component and the size of the palette. As we can see from Table 1. both parameters have a significant influence on classification accuracy. Increasing the number of colors in the palette brings greater diversity among the shades of colors, and therefore increases the discriminative power of the descriptor. Likewise, increasing the number of quantization levels increases the number of different colors in the palette. Also, with larger values of N, we lose less information in images and thereby increase discriminative power of the descriptor. On the other hand, increasing the value of N influences the time needed to learn a color palette since the time needed to choose colors for learning the palette increases when N increases.

TABLE 3: CLASSIFICATION ACCURACIES (%) OBTAINED USING FIXED HISTOGRAM

		Accuracy (%)
k_c – Palette size	64	56,5 ± 2,4
	512	67,1 ± 2,2
	4096	75,3 ± 2,3

Now we will introduce our modified process of palette generation, and we will show the results of classification obtained using our modified BoC descriptor. Difference in process of palette generation of this descriptor is in the way how the image is quantized. In the first case, the whole range of values for each color component was used in quantization (in MATLAB for Lab color space this range is [0, 100] for L component and [-128, 127] for a and b color components), while in the second case, we calculate the color range for quantization dynamically, using the extreme intensity values of each color component separately: for image i, quantization range is calculated as [min(L_i), max(L_i)] for L, [min(a_i), max(a_i)] for a, and [min(b_i), max(b_i)] for b color component of CIE-Lab color space. The results obtained using the modified method are shown in Table 2. As we can see, parameter k_c significantly influences the classification accuracy, as in the first case. However, in this case, parameter N almost does not have impact on classification accuracy. Independence of classification accuracy to parameter N is reflected in the fact that each image brings its own range of colors, which differs from image to image, and therefore,

TABLE 4: CLASSIFICATION ACCURACIES (%) OBTAINED USING LOCAL BOC DESCRIPTOR

		M – Length of histogram codebook				
		100	500	1000	2000	5000
k_c – Palette size	10	69,5 ± 1,4	78,5 ± 1,1	79,8 ± 0,9	81,1 ± 1,2	81,5 ± 1,1
	50	72,5 ± 2,3	84,0 ± 1,3	84,7 ± 1,2	86,9 ± 1,6	**87,1 ± 1,0**
	100	70,2 ± 1,5	82,0 ± 2,0	83,4 ± 2,1	85,6 ± 1,6	86,9 ± 0,9
	500	62,8 ± 1,3	69,9 ± 1,5	74,1 ± 1,8	80,2 ± 1,3	82,8 ± 1,2
	1k	54,2 ± 2,2	64,8 ± 2,2	62,1 ± 2,2	74,0 ± 2,3	71,1 ± 2,3
	5k	40,5 ± 2,1	49,0 ± 2,3	48,0 ± 2,0	53,2 ± 1,0	56,0 ± 2,4

we cluster a large number of different colors. In the first example, that number was always lower or equal to N^3. The advantage of our modified method is not only in improving the classification accuracy, but also in improving the efficiency of the palette computation. In Tables 1. and 2. we see that in order to achieve classification accuracy of 86%, for ordinary BoC method we need to choose $N \geq 64$, while for modified method it is obtained with $N \geq 4$. As we mentioned before, choosing bigger N increases the time for calculation of color histogram. In Table 3. the classification accuracies obtained using fixed color histogram are shown. We can see that the results obtained using BoC are consistently better. Finally, in Fig. 3 the confusion matrix obtained using global BoC descriptor is shown.

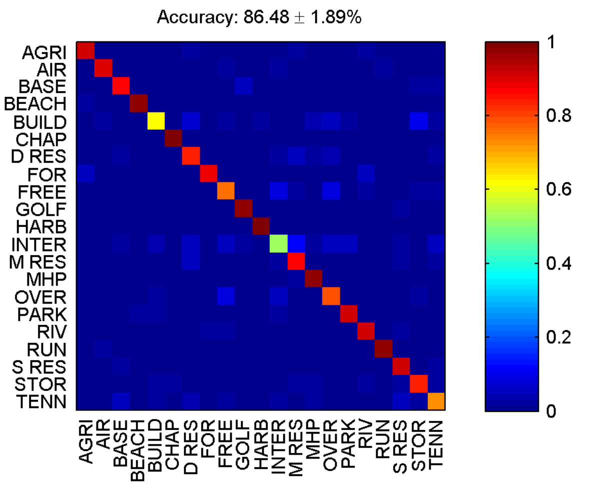

Figure 3. Confusion matrix for global BoC descriptor ($N = 8$, $k_c = 5000$)

C. Local descriptor

Results obtained using local BoC descriptor are shown in Table 4. We used the patch size of 16x16 pixels, which was moved over the image with the step of 4 pixels. For quantization, we chose parameter $N = 8$. As we can see in Table 4. classification accuracies increase when length of local BoC descriptor increases, which is equivalent to the results obtained using global descriptor. However, contrary to global descriptor, the maximum of classification accuracy is obtained using relatively small number of colors. The reason for this behavior lays in the fact that the size of the patch is much smaller than the size of palette in the first case, and choosing a large number for palette size will lead to rarely filled local histograms. Clustering of rarely filled vectors introduces large quantization errors, which reduces the classification accuracy.

IV. CONCLUSION

In this paper we evaluated BoC descriptor for aerial image classification. The obtained classification accuracies indicate that this descriptor based on color information only can be very powerful tool for land use classification. Since the generation of BoC descriptor uses only color information, in order to increase performance and accuracy, it can be combined with other descriptors such as SIFT descriptor, texture descriptors, etc.

LITERATURE

[1] C. Wengert, M. Douze, H. Jegou, "Bag-of-colors for improved image search," in *Proc. 19th ACM International Conference on Multimedia*, 2011, pp. 1437-1440.

[2] Y. Yang and S. Newsam, "Bag-of-visual-words and spatial extensions for land-use classification," in *Proc. ACM SIGSPATIAL GIS*, 2010, pp. 270-279.

[3] R. Arandjelović, A. Zisserman, "Three things everyone should know to improve object retrieval," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, 2012, pp. 2911-2918.

[4] M. J. Swain and D. H. Ballard, "Color indexing," *Int. J. Comp. Vis.*, vol.7, no.1, pp. 11-32, 1991.

[5] B. S. Manjunath, et al. "Color and texture descriptors," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 6, pp. 703-715, 2001.

[6] A. Vedaldi and A. Zisserman, "Efficient additive kernels via explicit feature maps," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.34, no.3, pp. 480-492, 2012.