# Image Description Using a Multiplier-Less Operator

Hakan Tuna, İbrahim Onaran, and A. Enis Çetin, *Member, IEEE*

*Abstract*—**A fast algorithm for image classification based on a computationally efficient operator forming a semigroup on real numbers is developed. The new operator does not require any multiplications. The co-difference matrix based on the new operator is defined and an image descriptor using the co-difference matrix is developed. In the proposed method, the multiplication operation of the well-known covariance method is replaced by the new operator. The proposed method is experimentally compared with the regular covariance matrix method. The proposed descriptor performs as well as the the regular covariance method without performing any multiplications. Texture recognition and licence plate identification examples are presented.**

*Index Terms*—**Co-difference matrix, covariance matrix, license plate identification, multiplier-less signal processing, texture recognition.**

## I. INTRODUCTION

**D**ESCRIPTIVE feature extraction from images or image regions are necessary in many image recognition and video analysis problems. Practical applications include intelligent video surveillance systems with object tracking, human and vehicle recognition and license plate recognition features [1], [2].

In this letter, we introduce an operator which forms a semigroup on real numbers and define an image region descriptor based on this operator. We replace the multiplication operation in regular covariance matrix method with the new operator and we call this matrix as the co-difference matrix. We show that the co-difference matrix method performs as well as the covariance matrix method in texture classification and license plate identification applications.

Porikli et.al introduced the covariance matrix method as a new image region descriptor, and showed that covariance method performed better than the previous approaches to the texture recognition problem [3]–[5]. They also developed an object tracking method using the covariance matrix.

Let $\mathbf{f}$ be a $d$-dimensional feature vector for each pixel $I(x, y)$ of a two-dimensional image. The vector $\mathbf{f}$ may contain the intensity, color components, and gradient values of a given pixel. Let us index the image pixels using a single index $k$, and assume that there are $n$ pixels in a given image region. As a result we have $n$ $d$-dimensional feature vectors $(\mathbf{f_k})_{k=1...n}$. The covariance matrix of the image region is defined as

$$\mathbf{C} = \frac{1}{n-1} \sum_{k=1}^{n} (\mathbf{f_k} - \mu)(\mathbf{f_k} - \mu)^T \qquad (1)$$

where $\mu$ is the mean vector of the feature vectors. Since the $d \times d$ covariance matrix $\mathbf{C}$ is symmetric the number of independent parameters are not $d \times d$ but $d(d+1)/2$ .

## II. CO-DIFFERENCE MATRIX

Computational cost of a single covariance matrix for a given image region is not heavy. However, computational cost becomes important when we want to scan a large image at different scales and all locations to detect a specific object. Furthermore, many video processing applications require real-time solutions. In order to decrease the computational cost, we introduce the co-difference matrix as follows

$$\mathbf{S} = \frac{1}{n-1} \sum_{k=1}^{n} (\mathbf{f_k} - \mu) \odot (\mathbf{f_k} - \mu)^T \qquad (2)$$

where the operator $\odot$ acts like a matrix multiplication operator, however, the scalar multiplication is replaced by an additive operator $\oplus$. The operator $\oplus$ is basically an addition operation but the sign of the result behaves like the multiplication operation:

$$a \oplus b = \begin{cases} a + b, & \text{if } a \geq 0 \quad \text{and} \quad b \geq 0 \\ a - b, & \text{if } a \leq 0 \quad \text{and} \quad b \geq 0 \\ -a + b, & \text{if } a \geq 0 \quad \text{and} \quad b \leq 0 \\ -a - b, & \text{if } a \leq 0 \quad \text{and} \quad b \leq 0 \end{cases} \qquad (3)$$

for real numbers $a$ and $b$. We can also express (3) as follows

$$a \oplus b = \text{sign}\,(a \times b)\,(|a| + |b|). \qquad (4)$$

Since $a \oplus b = b \oplus a$ the co-difference matrix is also symmetric as the covariance matrix. Co-difference behaves similar to the covariance function. If two variables tend to vary together, co-difference function produces positive results as the covariance. When two variables tend to vary inversely, co-difference equation gives negative results. On the other hand, computational cost is decreased by replacing the multiplication operation with addition operation.

The operator $\oplus$ satisfies totality, associativity and identity properties, therefore it is a monoid function. In other words, it is a semigroup with identity property. We successfully used similar statistical methods in [6]. Another similar statistical function is the Average Magnitude Difference Function (AMDF) which is widely used in speech processing to determine the periodicity of voiced sounds.
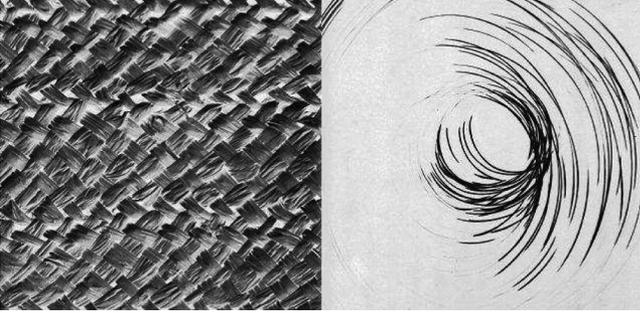
Fig. 1. Sample images from the Brodatz texture database. It contains non-homogeneous textures as well as homogeneous texture images.

TABLE I
SUCCESS RATES OF COVARIANCE AND CO-DIFFERENCE
METHODS IN BRODATZ TEXTURE DATABASE

|  | Covariance Matrix | Co-Difference Matrix |
|---|---|---|
| K=5 | %95.9 | %94.1 |
| K=10 | %96.3 | %96.8 |
| K=20 | %96.3 | %96.8 |



Fig. 2. Sample images from the license plate database.

## III. EXPERIMENTAL RESULTS AND CONCLUSIONS

We use the well-known Brodatz texture database for texture classification experiments. We compare the results of regular covariance method [3] and the proposed co-difference method.

The Brodatz texture database which we used in our experiments consists of 111 texture images. The size of each image is $640 \times 640$. Classification is a challenging task because of the non homogeneous texture images in the database. Sample images from the database are shown in Fig. 1. In our experiments, we divide each texture image into $320 \times 320$ sized four sub-images. The first half of these images are used for training and the remaining ones are used for testing.

We use intensity and the norms of first and second order derivatives of intensity values of pixels in both x and y directions for feature space as follows $\mathbf{f} = [|I| \, |I_x| \, |I_y| \, |I_{xx}| \, |I_{yy}|]$.

Therefore every pixel in a given image region is mapped to a $d = 5$-dimensional feature vector. Then the covariance and the co-difference of these features are calculated using both (1) and (2), respectively. As a result, we end up with $5 \times 5$ dimensional covariance and co-difference matrices, representing each region.

To represent each texture image, we choose 100 regions from random locations in the image. Each region is a square box with random sizes which varies from $16 \times 16$ to $128 \times 128$. We calculate the covariance matrix of each region. Thus, every texture image is represented with 100 covariance matrices extracted from random regions of these images.

We use the K-nearest neighbor (k-NN) algorithm for the classification task. The k-NN algorithm is a supervised learning

method which classifies samples according to majority of the closest training samples in the feature space. We use a generalized eigenvalue based distance metric to compare covariance and co-difference matrices which was introduced in [7], [8] and used in [3] as a part of the k-NN method:

$$D(\mathbf{C}_1, \mathbf{C}_2) = \sqrt{\sum_{k=1}^{d} \log^2 \lambda_i(\mathbf{C}_1, \mathbf{C}_2)} \qquad (5)$$

where $\lambda_i(\mathbf{C}_1, \mathbf{C}_2)$ are the generalized eigenvalues of covariance matrices $\mathbf{C}_1$ and $\mathbf{C}_2$.

We measure the distances between the instance covariance matrix to be classified and the covariance matrices in the train database. Similarly, two co-difference matrices $\mathbf{S}_1$ and $\mathbf{S}_2$ are compared to each other using the same metric $D(\mathbf{S}_1, \mathbf{S}_2)$. The k nearest samples from the train database is chosen and the query instance is assigned to the class most common amongst these k samples from the train database. If $k = 1$, then the query instance is assigned to the class of its nearest neighbor.

The choice of k depends on the data. Large values of k with respect to the number of samples decrease the probability of misclassifying and decrease the effect of noise. However it makes the classification boundary less distinct.

*1) Classification Results:* Brodatz texture database is a challenging database with lots of non-uniform texture images. To compare the proposed co-difference matrix method with the original covariance matrix method

Both covariance matrices and co-difference matrices are extracted from randomly selected regions and added to the training set. Then the same procedure is also applied in the query set. For different values of K, textures are classified by using the k-NN algorithm in both methods. Classification results for different k values are summarized in Table I.

In [3], slightly higher classification results than the results in Table I are obtained for the covariance method in Brodatz texture database. This is possible because each texture is represented by covariance matrices extracted from randomly selected regions. As a result, classification results may not turn out the same.

*2) Identification of License Plate Regions:* In order to compare our co-difference matrix method with the covariance method [9], we constructed a license plate database from the Internet.

License plate database consists of plate images gathered from an internet page which contains galleries of used cars for sale. This dataset contains Turkish license plate samples and is a challenging dataset. License plate images taken from this website have different illumination, and they are at different scales and the pictures are taken from different angles. The negative samples for training and query datasets are obtained randomly from car pictures with darkened or removed license plate regions. The training and test set contains 99 positive and 120 negative images, respectively. The license plate database and the software is available at www.ee.bilkent.edu.tr/~signal. Sample images from the database are shown in Fig. 2.

The feature vector used in covariance matrix and co-difference matrix computation is 7 dimensional in this problem: $\mathbf{f} = [\, |x| \, |y| \, |I| \, |I_x| \, |I_y| \, |I_{xx}| \, |I_{yy}| \,]$ where $x$ and $y$ coordinate
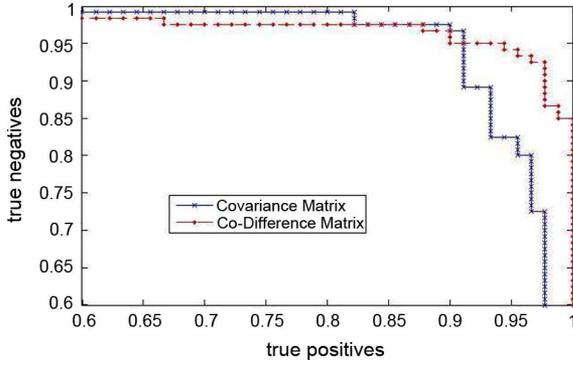
Fig. 3. ROC curve of the covariance matrix method and co-difference matrix method.

|  | Covariance Matrix | Co-Difference Matrix |
|---|---|---|
| Sum | $\frac{3M^2N+NM-M^2-M}{2}$ | $\frac{4M^2N+2NM-M^2-M}{2}$ |
| Multiplication | $\frac{M^2+M}{2}N$ | 0 |
| Sign Comparison | 0 | $\frac{M^2+M}{2}N$ |
| Division | $\frac{M^2+M}{2}$ | $\frac{M^2+M}{2}$ |

values of pixels are normalized to $[0\ 1]$ in order to gain scale robustness against scale, $I$ is the intensity value, $I_x$, $I_y$, $I_{xx}$, $I_{yy}$ corresponds to the first and second order derivatives of intensity values along the x and y directions, respectively. Since $cov(x,x)$, $cov(y,y)$ and $cov(x,y)$ values are always constant for all images we end up with $[7 \times (7+1)/2] - 3 = 25$ different covariance or co-difference values.

We employ a three layer neural network algorithm for classification task. Non-constant and non-repeating values of covariance and co-difference matrices are fed to the neural network, and the neural network outputs a result between $-1$ and $1$ in order to decide if the region corresponds to a license plate or not. In order to obtain ROC curves, we ordered the query samples according to the output values of the neural network. We divide this ordered sequence from every possible location. Then the part with higher values are labeled as positive results and the part with lower values are labeled as negative results. At each different division, the number of true positives and true negatives are computed and marked on the ROC graph. The results are depicted in Fig. 3.

Experimental results show that the proposed co-difference matrix descriptor gives very similar results to the covariance matrix descriptor.

The computational cost of the co-difference method is lower than the covariance method because it does not require any multiplications. This is especially important in real time applications in which the entire image or video frame has to be scanned at several scales to determine matching regions and ASIC implementations [10]–[12].

Table II describes the computational cost of the covariance method and the co-difference method for an image region having N pixels. Each pixel has M features. Therefore the resulting covariance and co-difference matrices are M by M.

We use $M^2 + M/2N$ additions and sign comparisons in co-difference method instead of the same number of multiplications in the covariance method. Both methods also require $M(N-1)$ additions and $M$ divisions for mean computation.

REFERENCES

[1] E. Bala and A. E. Çetin, "Computationally efficient wavelet affine invariant functions for shape recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 8, pp. 1095–1099, 2004.

[2] Y. Dedeoğlu, B. U. Töreyin, U. Güdükbay, and A. E. Çetin, "Silhouette-based method for object classification and human action recognition in video," *Lecture Notes in Computer Science.*, vol. 3979, pp. 64–77, 2006.

[3] Ö Tüzel, F. Porikli, and P. Meer, "Region covariance: A fast descriptor and for detection and classification," in *Proc. Image and Vision Computing*, Auckland, New Zealand, 2004.

[4] F. Porikli, "Making silicon a little bit less blind: Seeing and tracking humans," in *SPIE OE Mag.*, Newsroom ed. : , 2006.

[5] F. Porikli, Ö. Tüzel, and P. Meer, "Covariance tracking using model update based means on riemannian manifolds," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2006.

[6] T. Akgül, S. Mingui, R. J. Sclahassi, and A. E. Çetin, "Characterization of sleep spindles using higher order statistics and spectra," *IEEE Trans. Biomed. Eng.*, vol. 47, pp. 997–1009, 2000.

[7] W. Förstner and B. Moonen, A Metric for Covariance Matrices Dept. Geodesy and Geoinformatics, Stuttgart Univ., Stuttgart, Germany, 1999.

[8] J. N. L. Brümmer and L. R. Strydom, "An euclidean distance measure between covariance matrices of speechcepstra for text-independent speaker recognition," in *Proc. 1997 South African Symp. Communications and Signal Processing*, 1997, pp. 167–172.

[9] F. Porikli and T. Koçak, "Robust license plate detection using covariance descriptor in a neural network framework," in *IEEE Int. Conf. Advanced Video and Signal Based Surveillance, AVSS*, 2006, p. 107.

[10] K. Benkrid, "A multiplier-less FPGA core for image algebra neighbourhood operations," in *Proc. 2002 Int. Conf. Field-Programmable Technology,* , pp. 294–297.

[11] H. Jeong, J. Kim, and W. K. Cho, "Low-power multiplierless dct architecture using image correlation," *IEEE Trans. Consumer Electron*, vol. 50, no. 1, pp. 262–267, Feb. 2004.

[12] T. D. Tran, "The bindct: fast multiplierless approximation of the dct," *IEEE Signal Process. Lett.*, vol. 7, no. 6, pp. 141–144, Jun. 2000.