

## Multidimensional Markov Stationary Feature for Image Retrieval Systems

Md. Saiful Islam<sup>1\*</sup>, Md. Emdadul Haque<sup>2</sup>, and Md. Ekramul Hamid<sup>1</sup>

<sup>1</sup>*Department of Computer Science and Engineering, University of Rajshahi,  
Rajshahi 6205, Bangladesh.*

<sup>2</sup>*Department of Information and Communication Engineering, University of Rajshahi,  
Rajshahi 6205, Bangladesh.*

\*Corresponding author: [msiscse@ru.ac.bd](mailto:msiscse@ru.ac.bd)

### Abstract

Markov Stationary Features (MSF) not only considers the distribution of colors like histogram method does, also characterizes the spatial co-occurrence of histogram patterns. However, handling large scale database of images, simple MSF method is not sufficient to discriminate the images. In this paper, we have proposed a robust content based image retrieval algorithm that enhances the discriminating capability of the original MSF. The proposed Multidimensional MSF (MMSF) algorithm extends the MSF by generating multiple co-occurrence matrices with different quantization levels of an image. Publicly available WANG1000 and Corel10800 databases are used to evaluate the performance of the proposed algorithm. The experimental result justifies the effectiveness of the proposed method.

*Keywords:* Markov stationary feature; Markov chain; MMSF; Content based image retrieval.

### INTRODUCTION

With the proliferation of the world-wide-web, and the rapid increase of the multimedia data such as image and video, there is a strong demand for developing efficient techniques for storing, browsing, retrieval and indexing to exploit the full benefit of the explosive growth [1-4]. In content-based image retrieval system, images from database are automatically indexed by summarizing their low level visual contents such as color, texture, shape or spatial relationship according to user's visual requirement. Color is one of the most flexible and reliable visual features used in image retrieval methods. The feature is almost independent of image size and orientation, which is robust to its background complication.

There are several methods to represent colors of an image including color histogram, Color Coherence Vector(CCV), Color Auto Correlogram (CAC) and Markov Stationary Feature(MSF). G. Pass *et al.* introduce [5]the color histogram which is an effective way to represent the colors of an image. The histogram describes the global color distribution in the image. However, the histogram comparison saturates the discrimination when the database contains a large number of images. Thus, it was inevitable to integrate spatial information of an image with its color information. In CCV[6], proposed by G. Pass *et al.*, the histogram bins are divided into two types: coherent and incoherent. According to CCV, a bin is coherent if it belongs to large color region. Otherwise, the bin is incoherent. The CCV method shows better performance than color histograms when the images in the database have mostly uniform colors and the image is texture dominated.CAC method, introduced by Huang *et al.* [7-8], is used to characterize both the color distribution of

image pixels and the spatial correlation of pairs of colors. Instead of the indirect use of spatial information, the CAC method encodes local spatial structure information directly into histograms. However, CAC method only take cares the information of within-histogram bins. The Hidden Markov chain model [9-10] also uses the spatial information to discriminate the images of database. The model inherits the limitations of CAC.

An effective content-based image feature called Markov Stationary Feature (MSF) was introduced by Li *et al.*[11], for image indexing, retrieval and classification in the context of large scale image database. It characterizes the spatial co-occurrence of histogram patterns by Markov chain models.

In the context of MSF characterization, images in the database are divided into four categories depending upon the discrimination capability of histogram analysis: histogram-level distinguishable, intra-bin distinguishable (*i.e.* within-histogram bin), extra-bin distinguishable (*i.e.* between-histogram bins) and histogram undistinguishable images. To form the so-called MSF, initial and stationary distributions of the homogeneous Markov chain are combined to encode the intra-bin and extra-bin relationship of histogram, respectively. In practice, the MSF method generally outperforms the corresponding earlier content based methods. Subsequently, some research such as Directed Markov Stationary Feature (DMSF)[12], Multi-direction Markov Stationary Feature MDMSF[13], Markov Stationary Features and Vector Quantization Histogram(MSFHQ) [14] and some other methods [15-17] based on the same model have been demonstrated to enhance the performance of the original MSF.

Though the MSF method and its extended versions are far better than the earlier methods, they still suffer difficulty when the underlying image database is heterogeneous. To overcome the problem, we propose Multidimensional Markov Stationary Feature (MMSF) model that extends the existing MSF. It populates more spatial information to its feature space by computing multidimensional co-occurrence matrices of an image with multiple numbers of histogram bins depending on the image quantization levels.

#### MARKOV STATIONARY FEATURE (MSF)

The MSF extends the histogram features by characterizing the spatial co-occurrence of histogram patterns utilizing the Markov chain models and improves the distinguishable capability to the extra-bin distinguishable level. A brief discussion of the MSF is given bellow.

Suppose, an image  $I$  is quantized into  $K$  levels, thus the set of histogram bins of the image is  $S = \{s_1, s_2, \dots, s_k\}$ . The co-occurrence matrix containing spatial information is defined as  $C = \{c_{ij}\} \in \mathbb{R}^{k \times k}$  with each element

$$c_{ij} = \#(x_1 = s_i, x_2 = s_j | |x_1 - x_2| = d) \quad (1)$$

where,  $d$  indicates  $L_1$  distance between two (adjacent if  $d=1$ , that is in our case) pixels  $x_1$

and  $x_2$ . Each element  $c_{ij}$  accumulates the number of co-occurrence for bin  $s_i$  and  $s_j$ . When

the patterns  $s_i$  and  $s_j$  have large spatial co-occurrence, the possibility that  $s_i$  transit to  $s_j$  is

high. Note here the co-occurrence matrix  $C$  is a nonnegative symmetric matrix and can be

interpreted from a statistical view [11]. Markov chain model is adopted to characterize the spatial relationship of histogram bins, which treats bins as states in that model. The transition matrix, essential component of the chain, is statistically derived from the spatial co-occurrence matrix, defined as  $\mathbf{P} = [p_{ij}] \in \mathbb{R}^{K \times K}$ , where

$$p_{ij} = \frac{c_{ij}}{\sum_{j=1}^K c_{ij}}. \quad (2)$$

The  $P_{ij}$  is precisely the  $(ij)^{\text{th}}$  element of the transition matrix  $\mathbf{P}$ . The matrix  $\mathbf{P}$  must satisfies the basic properties of a Markov chain, namely,

$$\begin{aligned} p_{ij} &\geq 0 \text{ for all } i, j \in \{1, \dots, K\}, \text{ and} \\ \sum_{j=1}^K p_{ij} &= 1 \text{ for all } i \in \{1, \dots, K\}. \end{aligned} \quad (3)$$

It should be noted that, every individual image can be represented by a Markov chain what should be modeled as an individual transition matrix. Thus, comparing two Markov chains (i.e. two images) corresponds to the comparing two transition matrices. However, the transition matrix is space expensive because it requires  $M(K^2)$  spaces for  $K$  states i.e. bins. At the same time, time complexity of comparing two transition matrices is  $O(K^2)$  what is impractical for a large image database. Thus, it is desirable to build up a compact yet robust feature representation from the transition matrix rather than to become the required feature itself (i.e. the full transition matrix).

Based on the basic properties of the transition matrix and considering the Markov chain's two potential conditions [18], namely, irreducibility and a periodicity, Li *et al* [11] formulates a space efficient ( $2K$  instead of  $K^2$  in feature dimension) yet robust feature, exploiting the concept of Chapman-Kolmogorov equation. The so called feature representation is known as Markov stationary features  $[\pi_0, \pi]$ , which combines two  $K$ -dimensional vectors: initial distribution denoted by  $\pi_0$  and stationary distribution denoted by  $\pi$ . The initial distribution, also known as autocorrelogram, encodes intra-bin transitions of histogram bins of the underlying image, can be obtained as,

$$\pi_0 = \frac{c_{ii}}{\sum_{i=1}^K c_{ii}}. \quad (5)$$

On the other hand, the stationary distribution encodes extra-bin transitions. With a regular Markov chain, the stationary distribution of the transition matrix denoted by  $\pi = (\pi_1, \pi_1, \dots, \pi_k)$ , satisfies

$$\pi = \pi \mathbf{P}. \quad (6)$$

However, when the chain is irregular, unique solution of (6) may not exist. Thus, for a general case (i.e. for both regular and irregular chain), the fundamental limit theorem of Markov chain [19] would be a better solution as

$$\mathbf{A} = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^n \mathbf{P}^i, \text{ where } \mathbf{P}^0 \text{ is a zero matrix.} \quad (7)$$

To mitigate the approximation error due to small  $n$ , it is a good idea to average the rows of the matrix  $\mathbf{A}$  further as below,

$$\pi \approx \frac{1}{K} \sum_{i=1}^K \vec{a}_i, \text{ where } \mathbf{A} = [\vec{a}_1, \dots, \vec{a}_K]^T.$$

### PROPOSED MULTIDIMENSIONAL MSF

In the original MSF scheme discussed above, pixels of image (with  $K$  level quantization) in all directions (i.e. 8-neighborhood) are counted to generate a single dimensional co-occurrence matrix ( $K$ -by- $K$  in size). The two components (i.e. initial and stationary distribution) of the MSF are then derived based on the co-occurrence matrix. Practically, the MSF shows better retrieval performance on a large scale image database ranging from homogeneity to limited heterogeneity of image storages. However, when the underlying image database contains moderate or high degree heterogeneous images (i.e. histogram-level distinguishable, intra-bin distinguishable, extra-bin distinguishable and histogram undistinguishable images with different size, orientation, color and light condition) the single dimensional MSF method does not show the desirable performance. This is because, with an original (i.e. single dimensional) MSF, it is possible to capture only limited spatial information that is not sufficient to discriminate the images accurately.

In this paper we propose the Multidimensional Markov Stationary Feature (MMSF) algorithm to address the problem of database heterogeneity by populating more spatial information in the feature components. In this method, an image  $I$  is quantized into  $nK$  (e.g.  $K=10$ ) levels, thus the set of histogram bins are  $S_{nK} = \{s_1, s_2, \dots, s_{nK}\}$ , where  $n = 1, 2, \dots$ . Here,  $N$  indicates the dimension of our proposed MMSF method. The corresponding co-occurrence matrices are calculated as (for,  $K=10$ ),

$$c_{ij}^{10n} = \#(x_1 = s_i, x_2 = s_j | (s_i, s_j) \in (S_{10n})). \quad (9)$$

For example, if  $N=3$ , there will be three different co-occurrence matrices  $c_{ij}^{10}, c_{ij}^{20}$ , and of sizes 10-by-10, 20-by-20 and 30-by-30, respectively.

After computing the  $N$  different co-occurrence matrices, the Markov chain model (discussed in the previous section) is adopted for every co-occurrence matrix to characterize the spatial co-occurrence relation of corresponding set of histogram bins. Corresponding initial distribution (using (5)) and stationary distribution (using (6)) are then calculated and combined to form one dimensional MSF. The resultant MSFs for every co-occurrence matrices contribute to form our proposed MMSF feature space ( $K=10$ , in our case), as

$$\begin{aligned} \mathbf{x} &= [\pi_0^{10}, \pi^{10}, \pi_0^{20}, \pi^{20}, \pi_0^{30}, \pi^{30}], \text{ when } N = 3 \text{ (in our case)} \\ \mathbf{x} &= [\pi_0, \pi], \text{ where } \pi_0 = [\pi_0^{10}, \pi_0^{20}, \pi_0^{30}] \text{ and } \pi = [\pi^{10}, \pi^{20}, \pi^{30}]. \end{aligned} \quad (10)$$

### IMPLEMENTATION OF MMSF

HSV is one of the most prominent color space for low level image processing because using HSV color space the CBIR system results in better agreement with color perception than when using RGB. Thus, the RGB image of the underlying database is converted into HSV image first. Then, the individual H component (i.e. H image) of the HSV image is quantized into  $nK$  levels. The corresponding spatial co-occurrence matrices (e.g.  $C_{ij}^{10}, C_{ij}^{20}$ , and  $C_{ij}^{30}$ ) for  $K=10$  and  $N=3$  are calculated using (9). The initial

distributions (e.g.  $\pi_0^{10}, \pi_0^{20}, \pi_0^{30}$ ) are then computed from the co-occurrence matrices individually using (5). On the other hand, the equation (2) is employed individually for

each co-occurrence matrix to compute the corresponding transition matrices (e.g.  $p_{ij}^{10}$ ,  $p_{ij}^{20}$ , and  $p_{ij}^{30}$ ), respectively. Then, the stationary distributions (e.g.  $\pi^{10}$ ,  $\pi^{20}$ ,  $\pi^{30}$ ) are calculated individually from the transition matrices using (7) and (8). The resultant initial distributions and stationary distributions are used to form the targeted feature space according to the (10). The block diagram of our proposed system is shown in Fig. 1.

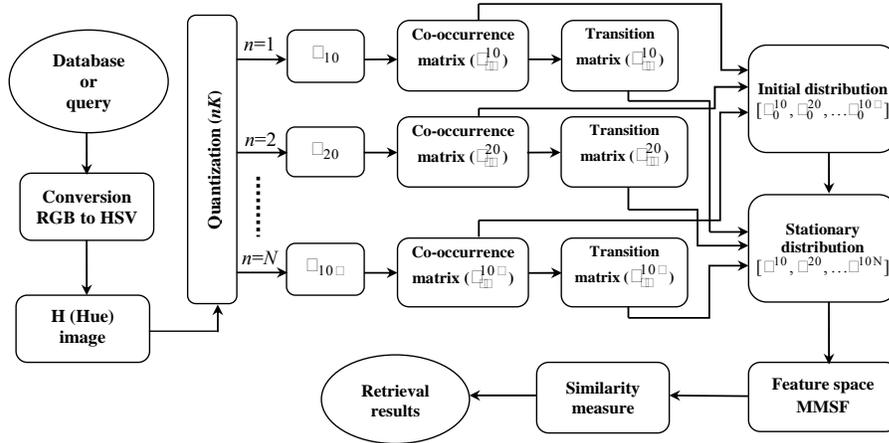


Fig. 1. Block diagram of the proposed MMSF system.

#### SIMILARITY MEASUREMENT

Assuming that no textual captions or other manual annotations of the images are given, the proper representation of the visual features of an image is then the proper description of the image content, such as MMSF. To find images that are visually similar to a given query, it needs a measure that can determine how similar or dissimilar [20] the different images are, comparing to the given query image.

For similarity measures between a query and underlying database images different wide variety of well-known methods are exist including Euclidean distance, city block distance, Earth Mover's distance (EMD), Mahalanobis distance, Chi square distance. For matching the MMSF features between the query and database images, we have selected *Chi square* distance measure in our simulation results.

For any two feature vectors  $h_q$  and  $h_d$ , the *chi square* distance between the corresponding images  $I_q$  (i.e. query image) and  $I_d$  (i.e. database images) can be defined as,

$$D(I_q, I_d) = \frac{1}{2} \sum_{k=1}^K \frac{[h_q(k) - h_d(k)]^2}{h_q(k) + h_d(k)} \quad (11)$$

The above distance formula is used for initial and stationary distribution for each image, which then be summed for totaling. The similarity result for each matching is then stored in an array in order to display the top matches from the database according to the ranking of similarity.

### RESULTS AND DISCUSSION

The proposed MMSF can be used in different kind of image processing and analysis problems. In our work, we have selected content based image retrieval (CBIR) system to demonstrate the resultant feature.

To compare the general performance of histogram, CAC, original MSF and our proposed MMSF method with CBIR, we have used two standard metrics, namely, *recall* and *precision*, which are defined as,

$$\text{Recall} = \frac{\text{No.of Relevant Images Retrieved}}{\text{Total No.of Relevant Images}} \quad (12)$$

$$\text{Precision} = \frac{\text{No.of Relevant Images Retrieved}}{\text{Total No.of Images Retrieved}} \quad (13)$$

As the results original MSF always shows superior to the result of Coherence Vector and Spatiogram [11] we exclude both of them. To make a fair comparison, all the concerned algorithms adopted the same setup and the same quantization method.

In this paper, we conducted experiments using two different worldwide recognized databases named WANG1000 and Corel10800. Because of the size and heterogeneity, it is axiomatic that the databases meet all the requirements to measure the performance of any content based image retrieval system.

For WANG1000 [21], the images of the database are pre-classified into 10 different categories each with 100 in size. The Fig.2 shows some sample images of each categories of the WANG1000 database. The image retrieval result with our proposed method is shown in Fig.3, where the top left image is a query image and the rest are retrieved relevant images.

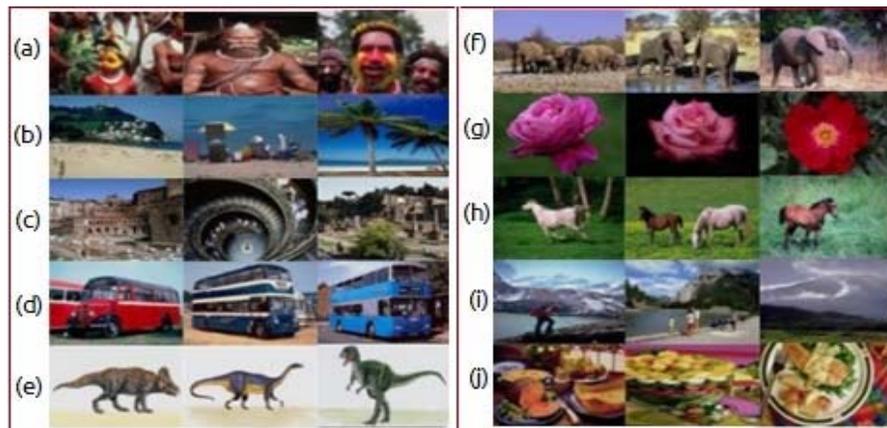


Fig. 2. Sample images of WANG1000 database for different categories: (a) African, (b) Beach, (c) Building, (d) Bus, (e) Dinosaur, (f) Elephant, (g) Flower, (h) Horse, (i) Mountain, and (j) Food.



Fig.3. Retrieval Result of the proposed method obtained for the query image-36 from WANG1000 database.

The comparative retrieval results of the original histogram, color autocorrelogram, original MSF and MMSF methods in terms of precision-recall curves for a single query image (e.g. African) are shown in Fig.4. The average precision-recall curves of the relevant methods for some 20 randomly selected query images from the same database are summarized in Fig.5.

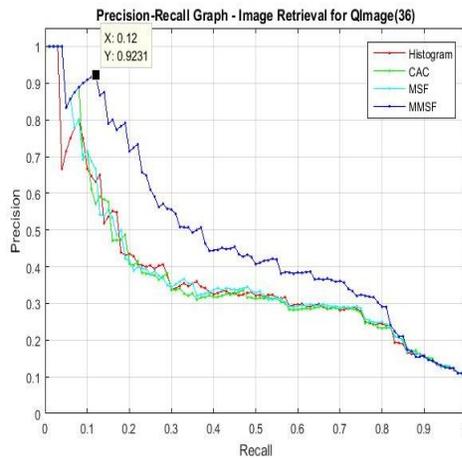


Fig. 4. Precision-recall curves for Histogram, CAC, original MSF, and proposed MMSF method against WANG1000 database for the query image-36.

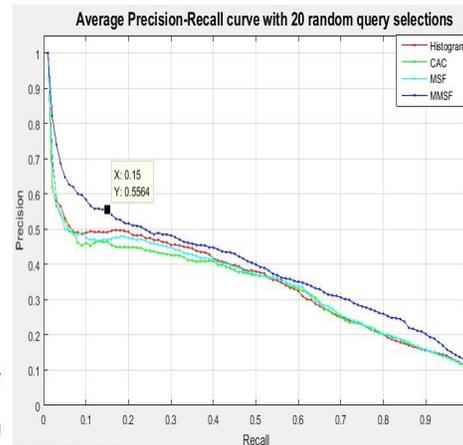


Fig. 5. Average Precision-recall curves of Histogram, CAC, original MSF, and proposed MMSF for some 20 randomly selected images from WANG1000 database.

Fig.4 illustrates that the proposed MMSF shows better performance compare to the existing techniques. The average precision-recall curves shows that the proposed method outperforms the other techniques, which validates the effectiveness of our method. Note that MMSF exploits the spatial co-occurrence information of the different number of color quantization of the same image which improves the performance.

For the Corel10800 [22], the images of the database are pre-classified into 80 different categories with different sizes ranging from 100 to 500 images in each, like, autumn, aviation, bonsai, bus, castle, cloud, dog, elephant, fitness, iceberg, primates, ship, stalactite, steam-engine, tiger, train, texture\_1, texture\_2, and so on. The comparative results in terms of precision-recall curves for a single query image (here the query image no. is 2311) are shown in Fig. 6. To evaluate the effectiveness of the proposed method over the other existing methods we select a category named "bus" containing 100 images. The average precision-recall curves with some 20 randomly selected query images from that category are shown in Fig. 7.

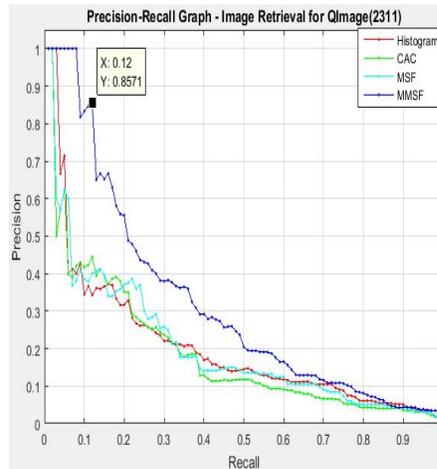


Fig. 6. Precision-recall curves for Histogram, CAC, original MSF, and proposed MMSF method against Corel10800 database for the query image-2311.

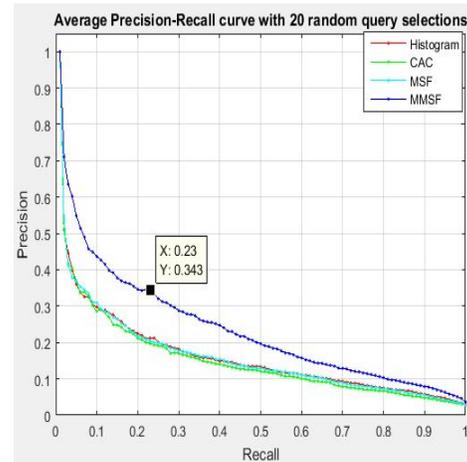


Fig. 7. Average Precision-recall curves of Histogram, CAC, original MSF, and proposed MMSF for some 20 randomly selected images from "bus" category of the Corel10800 db.

From the comparative results shown in Fig. 4, 5, 6 and 7, the following observations can be made: 1) the proposed MMSF method always shows better performance compared to the existing algorithms for both of the databases, 2) when the database grows larger and the degree of heterogeneity increases, the performance (shown in Fig. 6 and 7) of the existing techniques is quite frustrating.

#### CONCLUSION

Color histogram and color correlogram features are widely used in content-based image analysis for recognizing, searching, indexing and classifying images. Currently, Markov

stationary feature methods based on homogeneous Markov chain model are also widely used for content-based image and video analysis. However, handling large scale image database with higher degree of heterogeneity using the previous methods including the original MSF do not show satisfactory performance to discriminate images properly. To overcome the difficulty this paper proposes the MMSF for content-based image retrieval. MMSF extracts multidimensional spatial co-occurrence information in the visual descriptor (*i.e.* feature space) to improve the retrieval performance to a certain extent. According to the simulation results the proposed method outperforms the existing methods in almost all the cases.

#### **Acknowledgment**

This work is supported by the Information and Communication Technology Division, Ministry of Posts, Telecommunication and Information Technology of Bangladesh, under the ICT-fellowship program.

#### **References**

- [1] A. N. Tikle, C. Vaidya and P. Dahiwale, "A Survey of Indexing Techniques for Large Scale Content-Based Image Retrieval," IEEE Int. Conf. on Electrical, Electronics, Signals, Communication and Optimization (EESCO), pp. 1-5, Visakhapatnam, India, Jan. 24-25, 2015.
- [2] K. Juneja et al., "A Survey on Recent Image Indexing and Retrieval Techniques for Low-Level Feature Extraction in CBIR Systems," IEEE Int. Conf. on Computational Intelligence & Communication Technology (CICT), pp. 67-72, Ghaziabad, Pakistan, Feb. 13-14, 2015.
- [3] Keyuri M. Zinzuvadial, Bhavesh A. Tanawala, and Keyur N. Brahmabhatt, "A Survey on Feature Based Image Retrieval Using Classification and Relevance Feedback Techniques," Int. J. of Innovative Research in Computer and Communication Engineering, vol. 3, no. 1, pp. 508-513, Jan. 2015.
- [4] Mussarat Yasmin, Sajjad Mohsin, and Muhammad Sharif, "Intelligent Image Retrieval Techniques: A Survey," J. of Applied Research and Technology, vol. 12, no. 1 pp. 87-103, Feb. 2014.
- [5] G. Pass, and R Zabih, "Histogram Refinement for Content-Based Image Retrieval," 3rd IEEE Workshop on Applications of Computer Vision, pp. 96-102, Sarasota, Florida, Dec. 2-4, 1996.
- [6] G. Pass, R Zabih, and J. Miller, "Comparing Images Using Color Coherence Vectors," 4th ACM Int. Conf. on Multimedia, pages 65-73, Boston, USA, Nov. 18-22, 1996.
- [7] J. Huang et al., "Image Indexing Using Color Correlograms," IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR '97), pp. 762-768, San Juan, Puerto Rico, June 17-19, 1997.
- [8] J. Huang, et al, "Spatial Color Indexing and Applications," Int. J. of Computer Vision ( IJCV), vol. 35, no. 3, pp. 245-268, Dec. 1999.
- [9] L. Hsin-Chih, W. Ling-Ling, and Y. Shi-Nine, "Color image retrieval based on hidden Markov models," IEEE Trans. Image Processing, vol. 6, no. 2, pp. 332-339, Feb. 1997.
- [10] L. Jia, A. Najmi, and R. M. Gray, "Image classification by a two dimensional hidden Markov model," IEEE Trans. Image Processing, vol. 48, no. 2, pp. 517-533, Feb. 2000.
- [11] J. Li, et al, "One Step Beyond Histograms: Image Representation Using Markov Stationary Features," IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR '08), pp. 1-8, Alaska, USA, June 23-28, 2008.

- [12] B. Ni, S. Yan, and A. Kassim, "Directed Markov Stationary Features for Visual Classification," IEEE Int. Conf. on Acoustics, Speech, and Signal Processing, pp. 825-828, Taipei, Taiwan, Apr. 19-24, 2009.
- [13] F. Lee, et al, "Face Recognition Algorithm Using Multi-directional Markov Stationary Features and Adjacent Pixel Intensity Difference Quantization Histogram," 7th Int. Conf. on Systems and Networks Communications (ICSNC 2012), pp. 113-117, Lisbon, Portugal, Nov. 18-23, 2012.
- [14] Qiu Chen, et al, "Face Recognition Using Markov Stationary Features and Vector Quantization Histogram," IEEE 17th Int. Conf. on Computational Science and Engineering (CSE), pp. 1934-1938, Chengdu, China, Dec. 19-21, 2014.
- [15] Y. Song, Xiao Chen, S, Qu, "Content based Image Retrieval with Color Invariants," 2nd Int. Conf. on Computer Science and Electronics Engineering (ICCSEE 2013), pp. 760-762, Hangzhou, China, Mar. 22-23, 2013.
- [16] C. Zhang, J. Liu, H. Lu and S. Ma, "Web image mining using concept sensitive Markov stationary features," IEEE Int. Conf. on Multimedia and Expo (ICME 2009), pp. 462-465, NY, USA, 28 Jun - 03 July, 2009.
- [17] Feifei Lee, et al, "A Robust Face Recognition Algorithm Using Markov Stationary Features and Adjacent Pixel Intensity Difference Quantization Histogram," IEEE 7th Int. Conf. on Signal-Image Technology and Internet-Based Systems (SITIS), pp. 334-339, Dijon, France, Nov. 28 - Dec. 01, 2011.
- [18] Olle Häggström, "Markov Chains," Finite Markov Chains and Algorithmic Applications, Cambridge, UK, Cambridge University Press, 2002, pp. 09-10.
- [19] Charles M. Grinstead and J. Laurie Snell, "Markov Chains," in Introduction to Probability, Rhode Island, USA, American Mathematical Society, 1997, pp. 433-438.
- [20] Yossi Rubner, Carlo Tomasi, "Perceptual Metrics for Image Database Navigation," e-Book, 1st edition, 2001, ISBN 978-1-4757-3343-3.
- [21] James Z. Wang, The WANG databases used in SIMPLiCity paper for research comparison, Jun 2007. Accessed Aug. 15, 2015. <http://wang.ist.psu.edu/docs/related/>.
- [22] Dacheng Tao, The COREL Database for Content Based Image Retrieval, July 2009. Accessed Dec. 16, 2015. <https://sites.google.com/site/dctresearch/Home/content-based-image-retrieval>.