ISSN: 0976-9102 (ONLINE)
DOI: 10.21917/ijivp.2016.0175

A DISTRIBUTIONAL APPROACH FOR IMAGE RETRIEVAL USING HOTELLING'S T-SQUARE STATISTIC

K. Seetharaman[1] and S. Muthukumar[2]
[1]*Department of Computer Science and Engineering, Annamalai University, India*
E-mail: kseethadde@yahoo.com
[2]*Department of Computer Science and Engineering, Manonmaniam Sundaranar University, India*
E-mail: muthu123@gmail.com

Abstract

This paper proposes a novel method, based on a statistical probability distributional approach, Hotelling's T^2 statistic and Orthogonality test. If the input query image is structured, it is segmented into various regions according to its nature and structure. Otherwise, the image is treated as textured; and it is considered for the experiment as it is. The test statistic T^2 is applied on each region and compares it to the target image. If the test of hypothesis is accepted, it is inferred that the query and target images are same or similar. Otherwise, it is assumed that they belong to different groups. Moreover, the Eigen vectors are computed on each region, and the orthogonality test is employed to measure the angle between the two images. The obtained results outperform the existing methods.

***Keywords*:**

Query Image, Target Image, Hotelling's T^2 Statistic, Canberra Distance, Similarity

1. INTRODUCTION

The advent of the advanced technologies in data or image repository and image acquisition has enabled the computer vision systems with the usage of the digital libraries and the access of the image data through on-line have got exponential growth in the past decade. Many areas such as medical, educational, traffic management, law enforcement, automatic target recognition, and geographic information systems effectively utilize these advances in vision technology. Handling images, namely manipulating, storing, analysing, indexing, matching, retrieval, display, etc., is very complicated when compared to that of text manipulation. Thus, it needs proper image database systems, which can support the aforesaid image manipulations. It is observed from the literature that remarkable progress has been made in both theoretical research and system development. However, still it is a challenging problem for the researchers in the area of visual data mining to design an automatic retrieval system, because real-world images usually contain complicated objects and colour information. Mainly, images are retrieved based on the content of the image such as colour, shape, texture properties, and spatial orientation of the pixels. There are, broadly, two different approaches have been developed for searching and retrieving the images from the image repository: one is based on the image textual annotation and another one is based on image content information.

The method developed based on image contents is coined as content-based image retrieval (CBIR) system, which has attracted many researchers for over a decade. On CBIR system, the researchers concentrate on developing low-level global visual features such as colour, shape, texture properties, and spatial relationship, etc. that are used as a query for the retrieval process [1-4]. The method proposed by Seetharaman and Jeykarthic [5] segments a structured image into various regions according to the nature of the image. The segmented shapes in an image, for a sample, are depicted in Fig.1. Thus, in this paper, the same method is adopted to segment and the features are extracted on each region.

In this paper, a unified scheme is proposed for automatic image retrieval, based on the multivariate parametric tests, namely Hotelling's T^2 statistic for equality of mean vectors of query and target images, and the orthogonality test, i.e. Geometrical Interpretation. In the proposed technique, mean and covariance (first and second moments of the sample points) are used as representatives of both query and target images. The orthogonality test is used to give a geometrical interpretation of two normal populations [6-8]. The orthogonality test is employed based on the Eigen vectors, which are computed from the variance-covariance matrices of the query and target images.

1.1 OVERVIEW OF THE PROPOSED WORK

Firstly, the input query image is examined whether it is structured or textured. If it is structured then it is segmented into various regions according to its nature [9]; otherwise, it is treated as it is for the experiment. In the case of structured images, first shapes are segregated and then the features are extracted colour-wise, i.e. red, green, and blue colour components. The extracted features are formed as a vector. As discussed in [5], the feature vectors are formed as a feature vector database, and a link is established between the features and their corresponding image in the image database.

The Hotelling's T^2 statistic compares the query and target images. If the test of hypothesis is accepted, then it is inferred that the two images are same or similar. Otherwise, they are different. Furthermore, to emphasis the statistical test, Orthogonality test is performed on both query and target images, and the angle between the images are computed, and the angle 20^o is fixed as optimal threshold. The optimal threshold is fixed after a rigorous experiment. The proposed system facilitates the user to select a number of target images according to his/her suits by fixing threshold.

2. TEST STATISTIC FOR SIMILARITY OF IMAGES

Let X be a random variable that represents the intensity value with Gaussian noise of a pixel at location (k, l) in a colour image.

The pixel $X(k, l) \in \Re^3$ is a linear combination of three colours such as red, green, and blue, i.e. $X(k, l) = [r(k, l), g(k, l), b(k, l)]^T$, where T represents the transformation of the vector. The mean intensity value of each colour is represented by μ_r, μ_g and μ_b respectively, and the variance-covariance matrix is denoted by Σ. The multivariate normal density function of $X(k, l)$ is given by,

$$\frac{1}{\left(\sqrt{2\Pi}\right)^3 |\Sigma|} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right). \tag{1}$$

The density function mentioned in Eq.(1) can be denoted as $n(x/\mu, \Sigma)$ and the distribution law as $N(\mu, \Sigma)$. The i-th diagonal element of the covariance matrix, σ_{ii} is the variance of the i-th component of $X(k, l)$. The mean vector of each colour of the pixels in the image is,

$$\mu = E(X) = E\begin{bmatrix} X_r \\ X_g \\ X_b \end{bmatrix} = \begin{bmatrix} \mu_r \\ \mu_g \\ \mu_b \end{bmatrix} \tag{2}$$

and the variance-covariance matrix is

$$\Sigma = \begin{bmatrix} \sigma_{rr} & \sigma_{rg}\rho & \sigma_{rb}\rho \\ \sigma_{gr}\rho & \sigma_{gg} & \sigma_{gb}\rho \\ \sigma_{br}\rho & \sigma_{bg}\rho & \sigma_{bb}\rho \end{bmatrix} = \begin{bmatrix} \sigma_r^2 & \sigma_{rg}\rho & \sigma_{rb}\rho \\ \sigma_{gr}\rho & \sigma_g^2 & \sigma_{gb}\rho \\ \sigma_{br}\rho & \sigma_{bg}\rho & \sigma_b^2 \end{bmatrix} \tag{3}$$

where, σ_r^2, σ_g^2 and σ_b^2 are the variations among the intensity values of red, green, and blue colours respectively; σ_{rg} represents the interaction between the red and green colours; similarly, σ_{rb} and σ_{gb} represent the interaction between the corresponding colours; ρ represents correlation or interrelation between the corresponding colour pixels. The covariance matrix Σ is a symmetric and positive definite.

2.1 IMAGE MATCHING

The proposed system tests the equality of mean vectors of the query and target images; if they pass the test, then it is proceeded to test the orthogonality of Eigen vectors of the query and target images. Test for equality of mean vectors is employed to test whether the spectrum of energy of two images is same or not. If the test for equality of mean vectors and the orthogonality test pass the tests, then it is concluded that the two images are same; otherwise, it is assumed that they differ.

2.1.1 Test for Equality of Mean Vectors:

Let $x_{(\alpha)}^{(q)}$, $\alpha = 1,2,\ldots., N_q$ be the intensity values of the query image $N(\mu^{(q)}, \Sigma_q)$, and $x_{(\alpha)}^{(t)}$, $\alpha = 1,2,\ldots., N_t$ be the intensity values of the target image $N(\mu^{(t)}, \Sigma_t)$. Also, $x_{(\alpha)}^{(q)}$ is independent of $x_{(\alpha)}^{(t)}$. The intensity values of both query and target images are independent and identically distributed to Gaussian process; and are assumed to be $\Sigma_q = \Sigma_t$.

Hypotheses:

H_0: $\Sigma_q = \Sigma_t$ (Similarity)

H_a: $\Sigma_q \neq \Sigma_t$ (Non-Similarity)

To test the hypothesis that $\mu(q) = \mu(t)$, a special vector, we consider the squared statistical distance $\bar{x}^q - \bar{x}^t$. Now,

$$E\left(\bar{x}^q - \bar{x}^t\right) = E\left(\bar{x}^q\right) - E\left(\bar{x}^t\right) = \mu^{(q)} - \mu^{(t)}. \tag{4}$$

Since $\bar{x}^q$ and $\bar{x}^t$ are independent and thus $\operatorname{cov}\left(\bar{x}^q - \bar{x}^t\right) = 0$, it follows that

$$\operatorname{cov}\left(\bar{x}^q - \bar{x}^t\right) = \operatorname{cov}\left(\bar{x}^q\right) + \operatorname{cov}\left(\bar{x}^t\right) = \frac{1}{n_1}\Sigma + \frac{1}{n_2}\Sigma = \left(\frac{1}{n_1} + \frac{1}{n_2}\right)\Sigma \tag{5}$$

Because S_{pooled} estimates Σ, we see that $\left(\frac{1}{n_1} + \frac{1}{n_2}\right)S_{pooled}$ is an estimate of $\operatorname{cov}\left(\bar{x}^q - \bar{x}^t\right)$.

The likelihood ratio test of H_0: $\mu^{(q)} = \mu^{(t)}$ is based on the square of the statistical distance, T^2 is,

$$T^2 = \left(\bar{x}^q - \bar{x}^t - \delta_o\right)' \left[\left(\frac{1}{n_1} + \frac{1}{n_2}\right)S_{pooled}\right]^{-1} \left(\bar{x}^q - \bar{x}^t - \delta_o\right) > c^2 \tag{6}$$

where, the critical distance, c^2, is determined from the distribution of the two-sample Hotelling's T^2 statistic [10].

$$c^2 = \frac{(n_q + n_t + 2)p}{(n_q + n_t - p - 1)} F_{p, n_q + n_t - p - 1}(\alpha),$$

where, α is the level of significance.

$$S_{pooled} = \frac{\sum_{j=1}^{n_q}\left(\bar{x}_j^q - \bar{x}_j^t\right)\left(\bar{x}_j^q - \bar{x}_j^t\right)' + \sum_{j=1}^{n_t}\left(\bar{x}_j^q - \bar{x}_j^t\right)\left(\bar{x}_j^q - \bar{x}_j^t\right)'}{n_q + n_t - 2}$$

$$S_{pooled} = \frac{n_q - 1}{n_q + n_t - 2} S_q + \frac{n_t - 1}{n_q + n_t - 2} S_t \tag{7}$$

$$\bar{x}^q = \frac{1}{n_q}\sum_{j=1}^{n_q} \bar{x}_j^q \tag{8}$$

$$S_q = \sum_{j=1}^{n_q}\left(x_j^q - \bar{x}^q\right)\left(x_j^q - \bar{x}^q\right)' \tag{9}$$

are the sample mean vectors and sum of product of sample variance-covariance matrix of the query image.

$$\bar{x}^t = \frac{1}{n_t}\sum_{j=1}^{n_t} \bar{x}_j^t \tag{10}$$

$$S_t = \sum_{j=1}^{n_t}\left(x_j^t - \bar{x}^t\right)\left(x_j^t - \bar{x}^t\right)' \tag{11}$$

are the sample mean vectors and sum of product of sample variance-covariance matrix of the target image.

***Critical region*:** The query and target images are judged to be same, if $T^2 \leq c^2$, where, c^2 is the upper critical value of the F-distribution with $(n_q + n_t - 2)$ degrees of freedom at significance level α; otherwise, it is inferred that the two images are different.

2.2 ORTHOGONALITY TEST

The sum of the product of the variance-covariance matrices of the query and target images, i.e. S_{pooled} considered here, and let λ be an eigen value of S_{pooled}. If x is a non-zero vector such that $Ax = \lambda x$ then x is said to be an eigen vector of the matrix S_{pooled}, which is associated with the eigen value λ. Hereafter, in this section, the matrix S_{pooled} is called as matrix A, which is a positive definite and square matrix.

Let,

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n-1} & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n-1} & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n-11} & a_{n-12} & \cdots & a_{n-1n-1} & a_{n-1n} \\ a_{n1} & a_{n2} & \cdots & a_{nn-1} & a_{nn} \end{bmatrix} \quad (12)$$

and

$$|A\text{-}\lambda I| = 0 \quad (13)$$

$$|A - \lambda I| = \begin{bmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n-1} & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n-1} & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n-11} & a_{n-12} & \cdots & a_{n-1n-1}-\lambda & a_{n-1n} \\ a_{n1} & a_{n2} & \cdots & a_{nn-1} & a_{nn}-\lambda \end{bmatrix}. \quad (14)$$

The above Eq.(14) has n characteristic roots. The eigen vector can be computed as follows:

$$e = x/\sqrt{xx'} \quad (15)$$

where, x is a column vector. If there are n characteristic roots, there will be n eigen vectors, and that are considered as a feature vector of the query and target images.

In order to match and retrieve the target image, furthermore, the orthogonality test is performed. It is expressed in Eq.(16).

$$\theta = \cos^{-1}\left(\frac{\vec{v}^{q}.\vec{v}^{t}}{|\vec{v}^{q}|.|\vec{v}^{t}|}\right) \quad (16)$$

where, $\vec{v}^{q}$ and $\vec{v}^{t}$ represent the eigen vectors the query and target images. The angle, θ between the query and target images are computed, and it is fixed at 20^{o}, which is optimal threshold.

3. IMAGE AND FEATURE DATABASES CONSTRUCTION

An image database is constructed using manifold images collected from Corel 10K and 20K image databases [11], CalTech image database [12], Holidays image database [13], and from other sources such as images collected from the internet and images captured by a digital camera. Feature vectors such as Eigen vectors are computed using the expressions given in Eq.(10), Eq.(11), and Eq.(12) for the images in the image database, and also the number of shapes is extracted. The method used in [5] is adopted in this paper for image database and feature vector database designing, construction, and indexing.

In order to implement the proposed method discussed in section 2, 477 colour images of size 512×512 pixels have been collected from various sources that is 152 images from the Holidays image database, 176 images from the Corel image database, and 149 images from the CalTech image databases. The remaining 58 images with size 128×128 are photographed with a digital camera; 43 images with size 128×128 have been downloaded from the Internet. The textured images collected from Brodatz, Coral and VisTex image databases are divided into 16 non-overlapping sub-images of size 128×128. To examine that the proposed system is invariant for rotation and scaling, the images are rotated by 90^o, 180^o and 270^o, and scaled. Thus, totally there are ((16×(152+176+149))+58+43)×3 (rotated by 90^o, 180^o and 270^o)+7733 (scaled) = 30932 images. Based on this image collection, an image database and their feature vector database are constructed as discussed in the previous section. For a sample, some of them have been presented in this paper.

4. MEASURE OF PERFORMANCE

To validate and verify the performance of the proposed method, the Mean Average Precision score (MAP) [14] is used. The average precision for a single query q is the mean over the precision scores at each relevant item:

$$MAP = \frac{1}{N}\sum_{q\in n} AvgP(q) \quad (17)$$

where, N is the number of retrieved documents.

$$AvgP(q) = \frac{\sum_{k=1}^{n} P(k)\times rel(k)}{N_r} \quad (18)$$

where, $rel(k)$ is indicating function equal to 1, if the item at rank k is a relevant document, zero otherwise. $P(k)$ is the precision at cut-off k in the list.

5. EXPERIMENTAL RESULTS

In order to implement the proposed methods discussed in section 2, the image database constructed in section 3 is considered. The input query image, if structured, is segmented into various regions according to its nature as presented in Fig.1; otherwise, it is treated as textured, and the entire image is considered as it is. The number of regions of the query and target images matches, the experiment is proceeded to perform the Hotelling's T^2 test statistic. The T^2 test statistic is applied on each region of the query image.

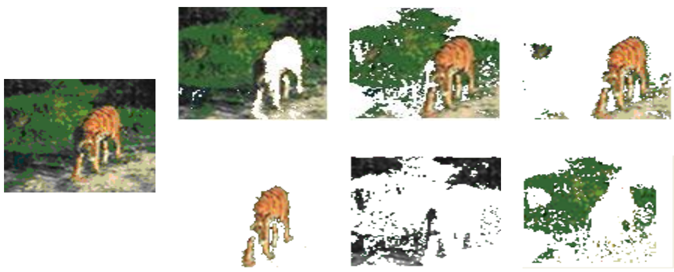

Fig.1. Segmented shapes of tiger image

To examine the proposed test statistic, a number of images are considered from the Holidays image database as query, and included in the experiment. For a sample, due to space constraint, a few of them are presented in this paper. The experiment is conducted at the levels of significance for the query image given in column 1 of the Fig.2. The method retrieves the images in column 2 at 0.01 level of significance; while the level of significance α is fixed at 0.05, the method retrieves the images in column 3; at 0.08 level of significance, the images in column 4 are retrieved; the images presented in columns 5 and 6 are retrieved, while the significance level is at 0.12.

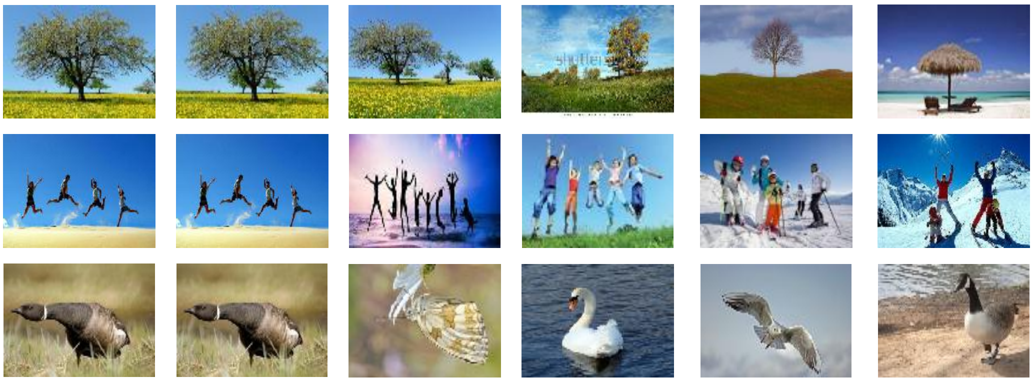

Fig.2. Holiday images. Column 1: input query images; columns 2 – 7: retrieved target images

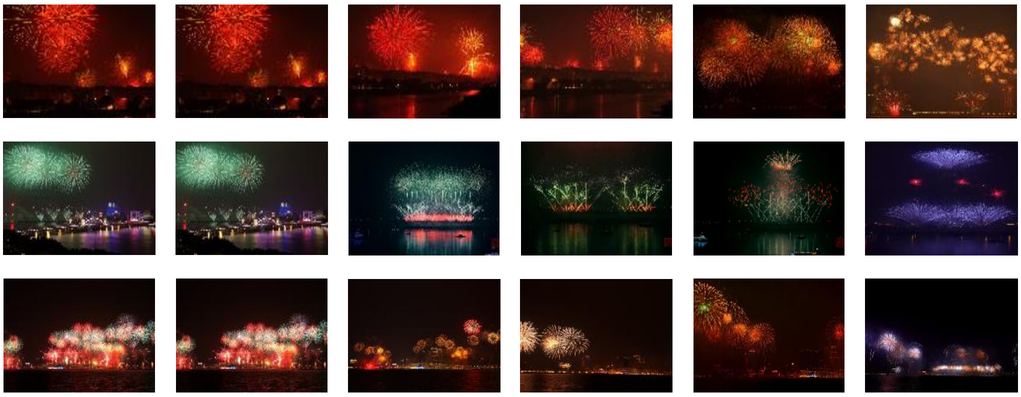

Fig.3. Corel 10K and 20K images – texture or semi-structured images
Column 1: input query image; columns 2-6: retrieved output images

Fig.4. Free downloaded structured images: (a): input query images; (b): retrieved output images

Table.1. Performance of the Proposed Method with Other Existing Methods

Image Database	Proposed system (α is at 0.10 and above)		Kullback-Liebler		Bhattacharyya		Mahalanobis	
	Recall	Precision	Recall	Precision	Recall	Precision	Recall	Precision
CalTech & Holidays	0.8992	0.8651	0.8549	0.7953	0.8096	0.7105	0.7952	0.6315
Corel	0.9098	0.8895	0.9002	0.7663	0.7921	0.7653	0.7965	0.6519
Structure Images	0.9296	0.8952	0.8698	0.8034	0.8928	0.7968	0.7647	0.7018

The experiment is also conducted on Corel images that are collected from the Corel 10K and 20K image databases. The images in the first column of the Fig.3 are inputted to the system, for which the system retrieves the images in columns 2 to 6 at various levels of significance. The images in column 2 are retrieved at the level of significance 0.01; the images in column 3 are retrieved at the level of significance 0.08; at the level of significance 0.12, the images in columns 4 and 5 are retrieved; images in column 6 are retrieved, while the significance level is at 0.18.

Furthermore, to prove the proposed system is invariant for rotation and scaling, a number of structured images are considered that are downloaded from the Internet. The images in columns 6 to 8 are rotated through various angles either clockwise or anti-clockwise, viz. 90^o, 180^o, and 270^o. The images in column 1 of the Fig.4 are fed as input to the system. The system retrieves the images in column 2 at the level of significance 0.05; while the level of significance is fixed at 0.01, the system retrieves the images in column 3; the images in columns 4 and 5 are retrieved when the significance level is at 0.12. The images in columns 6 to 8 are retrieved at the level of significance 0.2.

In order to compare the proposed method with the existing methods: Kullback-Liebler [15], Bhattacharyya [16], and Mahalanobis methods [17], an empirical study is conducted with manifold images, and the obtained results are tabulated in Table.1. It is observed from the results that the proposed method outperforms the existing methods.

6. CONCLUSION

Since the proposed system uses the global distributional differences of both query and target images, and adopts the Hotelling's T^2 statistic to compare them, and also it uses the eigen vectors, which is computed from the variance-covariance matrices of the query and target images; in the case of structured images, these features are extracted from the shapes in both query and target images, and those are compared shape-wise, it compares the number of shapes between the images. Even for the structured images, the proposed system adopts distributional differences of the corresponding shapes of the query and target images. The proposed system retrieves both textured and structured colour images, and it is robust for scaled and rotated images, since the query and target images are treated as a probability distribution function. Most of the existing methods retrieve a set of similar images, from which the user has to select the required images. But the proposed system facilitates the user to retrieve only the required image by fixing the level of significance at a desired level.

The statistical and the orthogonality techniques pass the test, it is concluded that the two images belong to the same class; otherwise, it is inferred that the images belong to different groups.

REFERENCES

[1] J. Huang, S.R. Kumar, M. Mitra, W.J. Zhu and R. Zabih, "Image Indexing using Colour Correlograms", *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 762-768, 1997.

[2] M.A. Stricker and M. Orengo, "Similarity of Color Images", *Proceedings of the SPIE*, Vol. 2420, pp. 381-392, 1995.

[3] A. Pentland, R.W. Picard and S. Sclaroff, "Photobook: Content-based Manipulation of Image Databases", *International Journal of Computer Vision*, Vol. 18, No. 3, pp. 233-254, 1996.

[4] Chiou-Shann Fuh, Shun-Wen Cho and K. Essig, "Hierarchical Color Image Region Segmentation for Content-based Image Retrieval System", *IEEE Transactions on Image Processing*, Vol. 9, No. 1, pp. 156-162, 2000.

[5] K. Seetharaman and M. Jeyakarthic, "Statistical Distributional Approach for Scale and Rotation Invariant Colour Image Retrieval using Multivariate Parametric Tests and Orthogonality Condition", *Journal of Visual Communication and Image Representation*, Vol. 25, No. 5, pp. 727-739, 2013.

[6] George B. Arfken, "Scalar or Dot Product", in *Mathematical Methods for Physicists*, 3rd edition, Academic Press, pp. 13-18, 1985.

[7] H. Jeffreys, and B. Jeffreys, "Scalar Product", in *Methods of Mathematical Physics*, 3rd edition, Cambridge University Press, pp. 65-67, 1988.

[8] A. Bhattacharyya, "On a Measure of Divergence between Two Statistical Populations Defined by their Probability Distributions", *Bulletin of Calcutta Mathematical Society*, Vol. 35, No. 1, pp. 99-109, 1943.

[9] A. Annamalai Giri and K. Seetharaman, "Color Image Retrieval based on Full Range Gaussian Markov Random Field Model", *International Journal of Innovative Research in Advanced Engineering*, Vol. 1, No. 11, pp. 62-70, 2014.

[10] Richard A. Johnson and Dean W. Wichern, "*Applied Multivariate Statistical Analysis*", 6th Edition, Pearson, 2008.

[11] http://www.corel.com/apac/

[12] http://vision.CalTech.edu/

[13] http://lear.inrialpes.fr/~jegou/data.php

[14] Stephen E. Robertson, Evangelos Kanoulas and Emine Yilmaz, "Extending Average Precision to Graded Relevance Judgments", *Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 603-610, 2010.

[15] I. J. Sumana, G. Lu, G and D. Zhang, "Comparison of Curvelet and Wavelet Texture Features for Content based Image Retrieval", *Proceedings of IEEE International Conference on Multimedia and Expo*, pp. 290-295, 2012.

[16] Y. Yang and S. Newsam, "Geographic Image Retrieval using Local Invariant Features", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 51, No. 2, pp. 878 - 832, 2013.

[17] S.C.H. Hoi, W. Liu and S.-F. Chang, "Semi-Supervised Distance Metric Learning for Collaborative Image Retrieval and Clustering", *ACM Transactions on Multimedia Computing, Communications, and Applications*, Vol. 6, No. 3, pp. 625-637, 2010.