

Statistical Tests of Hypothesis Based Color Image Retrieval

K. Seetharaman^{1*}, S. Selvaraj²

¹Department of Computer Science & Engineering, Annamalai University, Annamalai Nagar, India

²Department of Computer Applications, Research Scholar, Manonmaniam Sundaranar University, Tirunelveli, India

Email: ^{*}kseethadde@yahoo.com

Received 25 February 2016; accepted 22 May 2016; published 25 May 2016

Copyright © 2016 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

This paper proposes a novel method based on statistical tests of hypotheses, such as F-ratio and Welch's t-tests. The input query image is examined whether it is a textured or structured. If it is structured, the shapes are segregated into various regions according to its nature; otherwise, it is treated as textured image and considered the entire image as it is for the experiment. The afore-said tests are applied regions-wise. First, the F-ratio test is applied, if the images pass the test, then it is proceeded to test the spectrum of energy, *i.e.* means of the two images. If the images pass both tests, then it is concluded that the two images are the same or similar. Otherwise, they differ. Since the proposed system is distribution-based, it is invariant for rotation and scaling. Also, the system facilitates the user to fix the number of images to be retrieved, because the user can fix the level of significance according to their requirements. These are the main advantages of the proposed system.

Keywords

F-Ratio Test, Welch's Test, Tests of Hypotheses, Mean Average Precision, Target Image, Query Image

1. Introduction

A manifold applications and services of the computer vision and vision communication, viz. transmission of volume of images—medical images, aerial images, heterogeneous data in business, image communication in military, traffic management, cine field, etc., cause a proliferation in the image repository and retrieval for the past two decades. Typical application domains, such as aerial view images in remote sensing, medical images, fingerprints in forensics, museum collections in an art gallery, and registration of trademarks and logos, etc., ef-

*Corresponding author.

fectively utilize the applications of the image mining and image repository. In recent years, the fast evolution of network technology also makes the computer vision and information retrieval more and more efficient and effective. The advancement in computer networks and in communications facilitates the users to search desired images stored in the image repository on a remote server via network.

Nowadays, a number of approaches such as distributional approach [1]-[4], probability-based approach [5] [6], distance measures-based approach [7]-[9], and statistical tests of hypotheses-based approach [4] are available for image retrieval. Though a number of approaches are available, the tests of the hypothesis-based approach play a noteworthy role in the image retrieval. Because the tests of the hypothesis-based approach is distribution-based and it permits the users to fix the level of significance for the test of hypothesis, it enables the users to fix the number of images to be retrieved, only those are required; and it also facilitates the automatic image retrieval system [4].

Most of the techniques have been developed, based on the contents of the images, namely low-level global visual features, viz. color properties, shape, texture, spatial orientation, etc., which are used as a query for the retrieval process [10]-[13]. In the last two decades, the Content-Based Image Retrieval (CBIR) system emerged as great potential for retrieving images, too. At the initial stage, the CBIR system takes the features of the images into account as single representation (global). Subsequently, many researchers suggested that the important features of the individual objects cannot be sufficiently represented by a single signature computed on the entire image, and there is a gap between the visual features and semantic concepts of the images [14]. In sequel, region-based method has been developed [15]-[19], which represents the focus of the users' perceptions on the image contents. This overcomes the drawbacks in [14]. The methods proposed in [3] [13] [20]-[23] classify or segment the entire image into various regions according to the objects or structures present in the image, and then the region-to-region comparison is made to measure the similarity between two images [3] [12] [21] [22]. The method proposed in [24] evaluates the shape descriptors on the shape-based image retrieval; and reports that the region-based descriptors, viz. moment-based and angular radial transform, yield better results than the contour-based Fourier descriptor and curvature scale space. But these techniques demand more computational time complexity, because it adopts Fourier transform and computes the curvature scale space. Normally, the Fourier transform leads to computational complexity; and the computation of the curvature or contour also demand computational time complexity. The method proposed in [25] retrieves only the textured images and gray-scale images. This technique uses orthogonality test, it demands computational complexity.

To overcome this drawback, this paper introduces content-based image retrieval method, which employs region-based method using statistical tests of hypothesis, such as test for equality of variances of the query and target images; and if both images pass the test then those are included into the test for equality of mean values, *i.e.* spectrum of energy [26] [27]. First, the input query image is segmented into various regions according to its structure. If it is gray-scale, it is considered as it is; otherwise, it is assumed that the image is color and it is transformed to HSV color model. The V represents the intensity values of the pixels, and also it represents the gray-scale. The test statistic expressed in equation (2), *i.e.* F-ratio test, tests the equality of the variances of the query and target images region-wise. If the variances pass the test, then it is proceeded to perform the equality of mean values of both the images region-wise, *i.e.* Welch's test is applied. If the variances and mean values pass the two tests at a particular significant level, namely according to users' suit, it is inferred that the two images are the same or similar; otherwise, it is concluded that the two images are not similar. The proposed technique does not demand computational time since the regions in the image are segregated simply by subtracting the colors. For instance, the red and red oriented region is segregated by subtracting the green and green oriented colors, blue and blue oriented colors from the red oriented colors, because the red oriented color dominates (red pixels' intensity values are larger than the green and blue colors) other colors in the red oriented region. Also, the proposed techniques, *i.e.* F-ratio test and Welch's test, are simple in terms of computational time. Thus, it demands less computational time.

Welch's test is the most significant and powerful when the sample sizes and the variances of the two groups are unequal [26] [27], while compared to that of the student's *t*-test, it is easy to make the correct choice. Generally, the size of the target image is not known in the automatic image retrieval systems; and the size of the target images may differ from the query image; and there is no any guarantee or constraint that the variances of the query and target images are to be equal. Thus, in this paper, the Welch's test is adopted, because it is robust for the sizes and the variances of the query and target images are not the same or unequal.

Overview of the Proposed Method

To verify and validate the proposed methods, a new dataset is constructed which contains a manifold images. The images are collected from various databases. First, the proposed system segments the input query images into various shapes according to its nature as depicted in **Figure 1**. If the segmented query image is color, then it is modelled to HSV color space; otherwise, it is treated as gray-scale image. The intensity values of the gray-scale image are considered for the experiment as it is. The color features are extracted from H and S components; and the texture features are extracted from the V components. Then the test for equality of variances and test for equality of mean values are applied on the color and texture features of the query and target images. If the outcome of these two tests is positively significant, then it is concluded that the query and target images are the same or similar. Otherwise, they belong to different groups.

The rest of the paper is organized as follows. The proposed test statistics are discussed in Section 2; the experimental settings, database design and construction are demonstrated in Section 3. Section 4 presents the measure of performance, while the experimental results are discussed in Section 5. The conclusion is concluded in Section 6.

2. Bases of the Proposed Test Statistics

Generally, when capturing an image through a camera (digital or analog) or scanning through a scanner, there are many possibilities for including noise in the image. The inclusion of noise is a random process, which is independent and identically distributed to Gaussian random process. Let the intensity value f at location (k, l) be a random process, and it is represented as $f(k, l)$. Hence, an image is assumed to be Gaussian Random field [3], [4].

The pixel $f(k, l) \in \mathbb{R}^3$ is a linear combination of three colors, namely red, green, and blue, *i.e.* $f(k, l) = \{r(k, l), g(k, l), b(k, l)\}$. The mean intensity value of each color is represented by μ , and the variation is denoted by σ^2 . The Gaussian density function of the $f(k, l)$ for each color is given by

$$\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad -\infty < x, \mu < \infty, \sigma > 0 \quad (1)$$

The density function in Equation (1) can be denoted as $N(x, \mu/\sigma^2)$ and the distribution law as $N(\mu, \sigma^2)$, where mean μ and variance σ^2 are parameters of the Gaussian distribution. As discussed in section 1, first the F-ratio test is employed to test the variances of the query and target images; if it passes, next the Welch's test is performed.

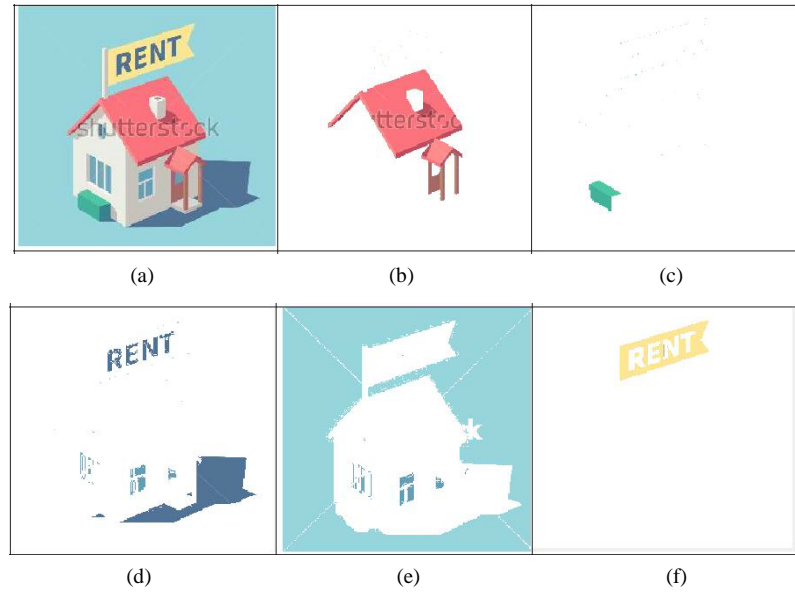


Figure 1. Segmented shapes. (a): actual image; (b): red component; (c): green component; (d): blue component; (e): combination of green and blue; (f): combination of red and green.

2.1. Test Statistics for Equality of Variances

Let the intensity values $F_q = (f_q^1, f_q^2, \dots, f_q^n)$ and $F_t = (f_t^1, f_t^2, \dots, f_t^n)$ of the query and target images be independent and identically distributed samples from two normal populations. In order to test the interactions among the pixels within an image and variations between the images, the F-ratio test is performed. The test statistic is expressed as in Equation (2). The test of the hypothesis is framed for two-tailed test as follows.

Hypotheses:

$H_0 : \sigma_q = \sigma_t$ (Similarity)

$H_a : \sigma_q \neq \sigma_t$ (Non-similarity)

$$F = \frac{\sigma_q^2}{\sigma_t^2} \quad (2)$$

Let

$$\sigma_q^2 = \frac{1}{n_q - 1} \sum_{i=1}^{n_q} (f_{iq} - \bar{f}_q)^2 \quad (3)$$

and

$$\sigma_t^2 = \frac{1}{n_t - 1} \sum_{i=1}^{n_t} (f_{it} - \bar{f}_t)^2 \quad (4)$$

be the sample variances of the query and target images respectively.

Let

$$\bar{f}_q = \frac{1}{n_q} \sum_{i=1}^{n_q} f_{iq} \quad (5)$$

and

$$\bar{f}_t = \frac{1}{n_t} \sum_{i=1}^{n_t} f_{it} \quad (6)$$

be the sample means of the query and target images respectively; n_q and n_t are the number of pixels in the query and target images respectively; f_q and f_t are the standard deviations of the query and target images.

Critical Region: If the F-ratio value is less than the critical value, F_{α, n_q-1, n_t-1} , with degrees of freedom $n_q - 1$ and $n_t - 1$, and the significance level α , it is inferred that the query and target images are same or similar; otherwise, it is concluded that the two images differ. The critical value F_α is referred to the F distribution in the statistical table.

2.2. Test Statistics for Equality of Mean Values

If the test for equality of variances and the means manifest that the query and target images pass the tests such as F-ratio and Welch's, it is concluded that the two images are same or similar; otherwise, it is assumed that they belong to different groups. The Welch's test statistic is expressed as in Equation (7), which compares the mean values of the query and target images. The test of the hypothesis is framed for two-tailed test as follows.

Hypotheses:

$H_0 : \mu_q = \mu_t$ (Similarity)

$H_a : \mu_q \neq \mu_t$ (Non-similarity)

$$WT = \frac{\mu_q - \mu_t}{\sqrt{\frac{s_q^2}{n_q} + \frac{s_t^2}{n_t}}} \quad (7)$$

The μ_q and μ_t are mean values of the query and target images that can be computed using the expression presented in Equations (5) and (6) respectively. The s_q^2 and s_t^2 are the sample variances of the query and tar-

get images respectively those can be computed using the expressions in Equations (3) and (4). The variances of the query and target images are computed separately. The degrees of freedom, ν , is calculated using the formula given in Equation (8), which is used to refer to the statistical table for the significant level α .

$$\nu = \frac{\left(\frac{s_q^2}{n_q} + \frac{s_t^2}{n_t} \right)^2}{\left(\frac{\frac{s_q^2}{n_q}}{\frac{n_q}{n_q-1}} \right)^2 + \left(\frac{\frac{s_t^2}{n_t}}{\frac{n_t}{n_t-1}} \right)^2} \quad (8)$$

Critical Region: If the WT value is less than the critical value, t_α , ν with degrees of freedom, ν , and the significance level α , it is inferred that the query and target images belong to the same class. Otherwise, it is concluded that they belong to different classes of image datasets. The critical value t_α is referred to the t-distribution in the statistical table.

3. Experimental Settings and Database Design

Experimental Settings: The input query image is examined whether it is color or gray-scale. If the query image is gray-scale, the shapes in the image are segregated into various regions as depicted in [Figure 1](#). If it is color image, first, the shapes are segregated into various regions, and are transformed to HSV color space. The HSV color space yields better results than the others. The number of shapes in the query and target images is compared. If the shapes in the images match up to 80 percent or above, then the tests of hypotheses such as F-ratio and Welch's tests are performed as discussed in sections 2.1 and 2.2. Otherwise, the tests of hypotheses are dropped. Ultimately, the query and target images pass the tests, viz. the number of shapes in the two images, F-ratio and Welch's test statistics, then it is concluded that the two images are same or similar; otherwise, they differ.

Database Design: In order to examine the efficacy and effectiveness of the above discussed statistical tests of hypotheses, we have constructed our own database. The database is a heterogeneous, since it contains a manifold images those are collected from various image datasets: Brodatz Album, Corel 5K and 10K image data-base, VisTex image database, Holidays image database, CalTech image dataset. The Corel image database consists of the different categories of images such as Elephants, Cracker, Flowers, Beach, Buses, Dinosaurs, African Aborigines, Snow Mountains, and Heritage Buildings. Each category contains more than 100 images. A number of images have been collected from the internet, and some images have been captured through a camera; those have also been incorporated. Totally it contains more than 256, 859 images. Based on this image collection, an image database and their feature vector database are constructed. For a sample, some of them have been presented in this paper.

4. Measure of Performance

In order to validate and verify the performance of the proposed method, the Mean Average Precision score (MAP) [28] is used. The average precision for a single query q is the mean over the precision scores at each relevant item:

$$\text{MAP} = \frac{1}{N} \sum_{q \in n} \text{AvgP}(q) \quad (9)$$

where N is the number of retrieved documents.

$$\text{AvgP}(q) = \frac{\sum_{k=1}^n P(k) \times \text{rel}(k)}{N_r} \quad (10)$$

where $\text{rel}(k)$ is an indicating function, which is equal to 1 if the item at rank k is a relevant document; other-

wise it is equal to 0; the $P(k)$ is the precision at cut-off k in the list.

5. Experiments and Results

An empirical study is conducted on the image database designed and constructed, based on the statistical tests of hypotheses discussed in sections 2.1 and 2.2, which examines the performance of the proposed test statistics. As discussed earlier, the input query image is segmented into various regions according to its nature as presented in **Figure 2**, if structured. Otherwise, it is treated as textured image, 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 F-ratio and Welch's tests. The query and target images are included in the test of the equality of variances, viz.

F-ratio test, and they pass the test; then it is proceeded to the test for equality of spectrum of energy, *i.e.*

Welch's test, the same images also pass the test, then it is concluded that the two images are same. Otherwise, the images are different.

The image in **Figure 2(a)** is submitted as input to the system, and the experiment is conducted at various significant levels. The system returns the images in row 1 of the **Figure 2(b)**, while the level of significance is fixed at 10 percent; while the significance level is at 15 percent, the system retrieves the images in row 3 of the **Figure 2(b)**; and at the level of significance 20 percent, the system retrieves the images in row 2 of the **Figure 2(b)**.

Furthermore, to prove that the system works well for the structured images, the building image presented in **Figure 3(a)** is considered as input. The system retrieves the images in columns 1 to 3 of the **Figure 3(b)** at the level of significance 12 percent; the system returns the images in column 4 of the **Figure 3(b)**, while the significance level is fixed at 20 percent.

The Holidays images are considered for the experiments, the images presented in column 1 of the **Figure 4** are inputted to the system, the system responses with the images in columns 2 to 5, while the significance level is fixed at 15 percent. The images in column 2 are retrieved at the level of significance 1 percent; the system returns the images in column 3 and 4, while the level of significance is fixed at 12 percent; and the images are returned when the significance level is at 15 percent.

Furthermore to emphasis the proposed system is robust for rotation and scaling, the image in **Figure 5(a)** is inputted, and the system responses with the images in **Figure 5(b)**. The images in column 1 are returned, while the level of significance is fixed at 1 percent; at the level of significance 5 percent, the images in columns 2 and 3 are retrieved; while the significance level is at 12 percent, the system returns the images in columns 4 and 5.

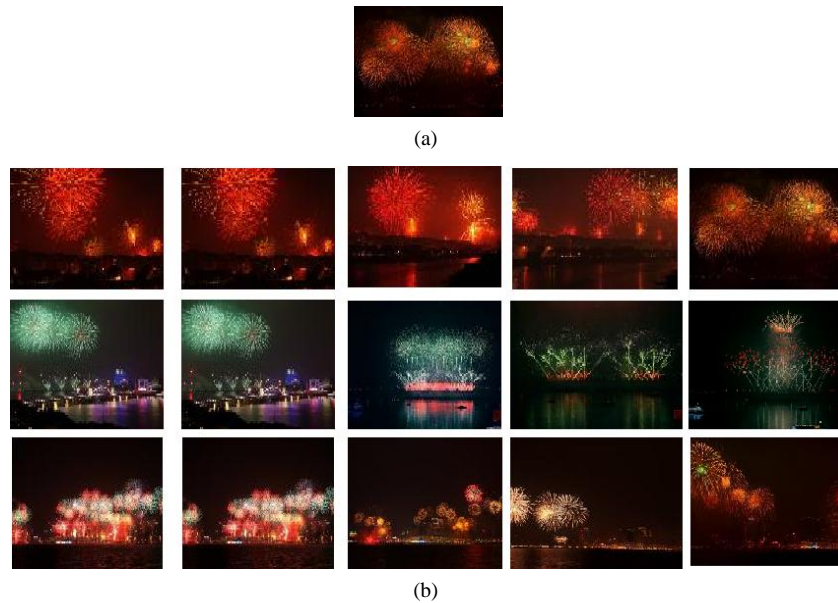


Figure 2. Corel images: texture or semi-structured images. (a): input query image; (b): columns 1 - 5: retrieved images.



Figure 3. Corel images: building images. (a): input query images; (b): retrieved output images.



Figure 4. Holidays images. Column 1: input query images; columns 2 - 5: retrieved target images.

To explore the attributes of the system, MAP score is computed using the expressions given in Equations (9) and (10) for the results obtained at various significance levels; and those are graphically represented as in **Figure 6**. It is observed from the experiments that the proposed system yields better results for structure images compared to those of the textured or semi-textured, because the structured images are segmented into various regions according to its nature, the image data become homogeneous and so the mean and variances computed for the shapes are more precise than the textured or semi-structured images. Also, the semi-textured images are mixed data types than the fine texture. In the case of semi-textured images, the images are not segmented so the data are mixed (either non-homogeneous or heterogeneous). The graph, also, reveals that while the significance level increases, the MAP score decreases monotonically. The increase of the significance level causes the system to re-turn more images. Thus, it facilitates the user to fix the level of significance according to their suit, so as to the user can get a number of images as they desire.

6. Conclusion

In this paper, statistical approach-based tests of hypotheses are employed, and the obtained results reveal that the proposed system outperforms the existing systems. Since the proposed system is distribution oriented, so it



Figure 5. Wall Street Bull-downloaded from internet. (a): Input query image; (b): Retrieved output images: row 1: scaled down image of size 75×100 ; row 2: actual image with size 96×128 ; row 3: images in row 2 are rotated clockwise by 90 degrees; row 4: images in row 2 are rotated clockwise by 180 degrees.

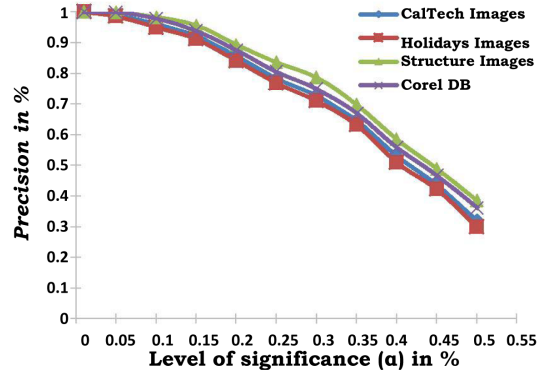


Figure 6. Line graph: threshold (t) vs. mean average precision scores.

facilitates the user to fix the level of significance according to their convenience, so as to the user can fix the number images to be retrieved; and it is invariant for rotation and scale. Besides the proposed test for equality of mean facilitates the user to retrieve the images which are unequal with sizes and even if they have unequal variances. Because the test for equality of means (Welch's test) is robust for unequal variances and unequal sizes of the images. Thus, the system retrieves the images despite either the query image or the target image is scaled.

Acknowledgements

First author would like to thank the University Grants Commission, New Delhi, India for providing financial

support for this research work under the Major Research Project Scheme, File No. 42-145/2013(SR).

References

- [1] Zhou, X.S. and Huang, T.S. (2001) Small Sample Learning during Multimedia Retrieval Using Biasmap. *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 1, 11-17.
- [2] Fukunaga, K. (1991) Introduction to Statistical Pattern Recognition. 2nd Edition, Academic Press, Cambridge.
- [3] Seetharaman, K. (2015) Image retrieval Based on Micro-Level Spatial Structure Features and content Analysis Using Full Range Gaussian Markov Random Field Model. *Engineering Applications of Artificial Intelligence*, **40**, 103-116. <http://dx.doi.org/10.1016/j.engappai.2015.01.008>
- [4] Seetharaman, K. and Jeyakarthic, M. (2013) Statistical Distributional Approach for Scale and Rotation Invariant Color Image Retrieval Using Multivariate Parametric Tests and Orthogonality Condition. *Journal of Visual Communication and Image Representation*, **25**, 727-739. <http://dx.doi.org/10.1016/j.jvcir.2014.01.004>
- [5] Kherfi, M.L. and Ziou, D. (2006) Relevance Feedback for CBIR: A New Approach Based on Probabilistic Feature Weighting with Positive and Negative Examples. *IEEE Transactions on Image Processing*, **15**, 1017-1030. <http://dx.doi.org/10.1109/TIP.2005.863969>
- [6] Kinga, I. and Jina, Z., (2003) Integrated Probability Function and Its Application to Content-Based Image Retrieval by Relevance Feedback. *Pattern Recognition*, **36**, 2177-2186. [http://dx.doi.org/10.1016/S0031-3203\(03\)00043-8](http://dx.doi.org/10.1016/S0031-3203(03)00043-8)
- [7] Hoi, S.C.H., Liu, W., Lyu, M.R. and Ma, W.-Y. (2006) Learning Distance Metrics with Contextual Constraints for Image Retrieval. *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 2, 2072-2078.
- [8] Yang, Y. and Newsam, S. (2013) Geographic Image Retrieval Using Local Invariant Features. *IEEE Transactions on Geoscience and Remote Sensing*, **51**, 818-832. <http://dx.doi.org/10.1109/TGRS.2012.2205158>
- [9] Hoi, S.C.H., Liu, W. and Chang, S.F. (2010) Semi-Supervised Distance Metric Learning for Collaborative Image Retrieval and Clustering. *ACM Transactions on Multimedia Computing, Communications, and Applications*, **6**, Article No. 18. <http://dx.doi.org/10.1145/1823746.1823752>
- [10] Huang, J., Kunar, S.R., Mitra, M., Zhu, W.J. and Zabih, R. (1997) Image Indexing Using Color Correlogram. *Proceeding of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Juan, 17-19 June 1997, 762-768. <http://dx.doi.org/10.1109/cvpr.1997.609412>
- [11] Stricker, M. and Orengo, M. (1995) Similarity of Color Images. *Storage and Retrieval for Image and Video Databases*, No. 2420, 381-392. <http://dx.doi.org/10.1117/12.205308>
- [12] Pentland, A., Picard, R. and Sclaroff, S. (1994) Photobook: Content-Based Manipulation of Image Databases. *SPIE Storage and Retrieval for Image and Video Databases II*, No. 2185, 34-47. <http://dx.doi.org/10.1117/12.171786>
- [13] Fuh, C.-S., Cho, S.-W. and Essig, K. (2000) Hierarchical Color Image Region Segmentation for Content-Based Image Retrieval System. *IEEE Transactions on Image Processing*, **9**, 156-162. <http://dx.doi.org/10.1109/83.817608>
- [14] Jing, F., Li, M., Zhang, H.J. and Zhang, B. (2004) An Efficient and Effective Region-Based Image Retrieval Framework. *IEEE Transactions on Image Processing*, **13**, 699-709. <http://dx.doi.org/10.1109/TIP.2004.826125>
- [15] Hsieh, J.-W., Eric, W. and Grimson, L. (2003) Spatial Template Extraction for Image Retrieval by Region Matching. *IEEE Transactions on Image Processing*, **12**, 1404-1415. <http://dx.doi.org/10.1109/TIP.2003.816013>
- [16] Belongie, S., Carson, C., Greenspan, H. and Malik, J. (2002) Recognition of Images in Large Databases Using Color and Texture. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **24**, 1026-1038.
- [17] Jing, F., Zhang, B., Lin, F.Z., Ma, Y.-W. and Zhang, H.-J. (2001) A Novel Region-Based Image Retrieval Method Using Relevance Feedback. *Proceedings of the 2001 ACM Workshops on Multimedia: Multimedia Information Retrieval*, Ottawa, 30 September-5 October 2001, 28-31.
- [18] Deng, Y. and Manjunath, B.S. (1999) An Efficient Low-Dimensional Color Indexing Scheme for Region-Based Image Retrieval. *IEEE International Conference on Acoustics, Speech, and Signal Processing*, **6**, 3017-3020.
- [19] Hsieh, I.-S. and Fan, H.C. (2001) Multiple Classifiers for Color Flag and Trademark Image Retrieval. *IEEE Transactions on Image Processing*, **10**, 938-950. <http://dx.doi.org/10.1109/83.923290>
- [20] Smith, J.R. and Li, C.S. (1999) Image Classification and Querying Using Composite Region Templates. *Journal of Computer Vision and Image Understanding*, **75**, 165-174. <http://dx.doi.org/10.1006/cviu.1999.0771>
- [21] Wang, J.Z. and Du, Y.P. (2001) Scalable Integrated Region-Based Image Retrieval Using IRM and Statistical Clustering. *Proceedings of the 1st ACM/IEEE-CS Joint Conference on Digital Libraries*, Roanoke, 24-28 June 2001, 268-277. <http://dx.doi.org/10.1145/379437.379679>
- [22] Kimura, F. and Shridhar, M. (1991) Handwritten Numerical Recognition Based on Multiple Algorithm. *Pattern Rec-*

- ognition, **24**, 969-983. [http://dx.doi.org/10.1016/0031-3203\(91\)90094-L](http://dx.doi.org/10.1016/0031-3203(91)90094-L)
- [23] Ho, T.K., Hull, J.J. and Srihari, S.N. (1994) Decision Combination in Multiple Classifier Systems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **16**, 66-75. <http://dx.doi.org/10.1109/34.273716>
- [24] Amanatiadis, A., Kaburlasos, V.G., Gasteratos, A. and Papadakis, S.E. (2011) Evaluation of Shape Descriptors for Shape-Based Image Retrieval. *IET Image Processing*, **5**, 493-499. <http://dx.doi.org/10.1049/iet-ipr.2009.0246>
- [25] Krishnamoorthy, R. and Sathya devi, S. (2013) A Multiresolution Approach for Rotation Invariant Texture Image Retrieval with Orthogonal Polynomial Model. *Journal of Visual Communication and Image Representation*, **23**, 18-30. <http://dx.doi.org/10.1016/j.jvcir.2011.07.011>
- [26] Wilcox, R. (2012) Introduction to Robust Estimation Hypothesis Testing. 3rd Edition, Academic Press, New York.
- [27] Ruxton, G.D. (2006) The Unequal Variance *t*-test Is an Underused Alternative to Student's *t*-test and the Mann-Whitney *U* Test. *Behavioral Ecology*, **17**, 688-690. <http://dx.doi.org/10.1093/beheco/ark016>
- [28] Robertson, S.E., Kanoulas, E. and Yilmaz, E. (2010) Extending Average 1086 Precision to Graded Relevance Judgments. *Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, Geneva, July 19-23, 603-610.