Research Article
Retrieval Architecture with Classified Query for Content Based Image Recognition

Rik Das,[1] Sudeep Thepade,[2] Subhajit Bhattacharya,[3] and Saurav Ghosh[4]

[1]Department of Information Technology, Xavier Institute of Social Service, Ranchi, Jharkhand 834001, India
[2]Department of Information Technology, Pimpri Chinchwad College of Engineering, Pune 411057, India
[3]Department of Marketing Management, Xavier Institute of Social Service, Ranchi, Jharkhand 834001, India
[4]A.K. Choudhury School of Information Technology, University of Calcutta, Kolkata 700098, India

Correspondence should be addressed to Rik Das; rikdas78@gmail.com

The consumer behavior has been observed to be largely influenced by image data with increasing familiarity of smart phones and World Wide Web. Traditional technique of browsing through product varieties in the Internet with text keywords has been gradually replaced by the easy accessible image data. The importance of image data has portrayed a steady growth in application orientation for business domain with the advent of different image capturing devices and social media. The paper has described a methodology of feature extraction by image binarization technique for enhancing identification and retrieval of information using content based image recognition. The proposed algorithm was tested on two public datasets, namely, Wang dataset and Oliva and Torralba (OT-Scene) dataset with 3688 images on the whole. It has outclassed the state-of-the-art techniques in performance measure and has shown statistical significance.

1. Introduction

Image data has strong impact on consumer attention that has kindled the buying intention of consumers [1, 2]. Advent of globalization has rapidly influenced the customer preferences and demands [3]. Consumer satisfaction has led to strong and positive behavioral outcomes connected to sustainable purchase [4]. The revenue of digital business process has been adversely affected by predominant dissatisfaction of consumers due to large amount of irrelevant results generation from text based queries [5]. The customers were deprived of taking pleasure in a perishable searching environment. This has led to the popularity of image data which has principally reinstated the text based keyword searching. The progression of multimedia technology has encouraged the use of multimedia in business practice and changed the way we use computers [6, 7]. Image data has revealed increasing importance in the field of contemporary business environment. The authors have proposed a novel methodology for retrieving image data with query classification which has stimulated the performance compared to state-of-the-art techniques for information identification. Statistical measures have been adopted to validate user responses for the fruitfulness of the proposed technique and to establish the significance of the findings.

2. Related Work

Accessibility of diverse online and offline information for products and services has radically altered the customer preferences [8]. Traditional means to locate the product of interest by the customers were based on text queries. However, the method has huge amount of irrelevant results as output. One of the driving factors for inappropriate output has been due to reprehensible selection of keywords as query. Recent approaches of searching have emphasized on the content of the searched object rather than its name as a keyword [9–12]. The content based searching process has been facilitated by the product image which can provide the necessary knowledge for the required product based on its

image contents and has been anticipated to filter out the unwanted results with higher probability. Various methods have been carried out for feature extraction that has applied image binarization as a tool to signify the object of interest and its background, respectively [13–15]. Threshold selection has been vital to enable binarization of image to differentiate the object and its background. It has been stated that uneven illumination and inconsistent gray levels within the image and its backgrounds have adversely affected threshold selection for binarization [16–18]. Threshold selection can be divided into three different categories, namely, mean threshold selection, local threshold selection, and global threshold selection. Feature extraction with mean threshold for binarization has been discussed in [19, 20] and with bit plane slicing in [21]. The problem of uneven illumination in images was efficiently addressed by local threshold techniques [9, 22–26]. The literatures have used measures of dispersion like standard deviation and variance to calculate the threshold. El Alami, 2011 [27], has worked with color and texture based features by 3D color histogram and Gabor filters. Hiremath and Pujari, 2007 [28], have calculated local descriptors of color and texture from the color moments and moments on Gabor filter responses by dividing the image into nonoverlapping blocks. Imagery significant point features were chosen by Banerjee et al., 2009 [29], for retrieval process of images. Jalab, 2011 [30], has fused color layout descriptor and Gabor texture for better detection of images. Shen and Wu, 2013 [31], have extracted signatures from images by dealing with color, texture, and spatial structure descriptors. Irtaza et al., 2014 [32], have explored wavelet packets and eigen values of Gabor filters to generate feature vectors from images. Rahimi and Moghaddam, 2015 [33], have referred to the intraclass and interclass features for effective extraction of image features of interest. The authors have proposed a novel technique for content based product recognition in an Internet based business model and has compared the same with the existing techniques of object recognition discussed in the literature. The performance of the proposed technique has outclassed the existing methods and has shown statistically significant improvement to foster generation of relevant results for a given content based product query.

3. Our Approach

The proposed methodology has considered different image categories which have signified diversified nature of product varieties offered by the firms. Uneven illumination of images can be a factor for degradation of product recognition process. It has been addressed by the authors by selection of local threshold for binarization. At the outset, the binarization was carried out using Niblack's method of local threshold selection [9, 18]. Each color component of an image was considered for derivation of pixel-wise threshold values by sliding a rectangular window over the component. The local *mean* $m(i, j)$ and *standard deviation* $\sigma(i, j)$ were calculated primarily with a window size of (25×25). The threshold calculation was done as $T(i, j) = m(i, j) + k \cdot \sigma(i, j)$. The value of k was a constant in between 0 and 1 and was considered to

be 0.6. The quality of binarization was dependent on the size of the sliding window and the value of k.

4. Feature Vector Generation

The pixels were locally divided into two different intensity values, namely, the higher intensity values and the lower intensity values by comparing with the corresponding threshold values. The mean and the standard deviation of the higher intensity value pixels and the lower intensity value pixels were considered to derive the higher intensity feature vector and the lower intensity feature vector, respectively. The process of feature extraction has been graphically illustrated in Figure 1. Initially, the original image displaying the model with a cell phone has been considered for feature extraction as shown in Figure 1(a). The image was divided into red (R), green (G), and blue (B) color component at the beginning in Figures 1(b)–1(d). The process was followed by binarization of each of the color components by Niblack threshold selection method shown in Figures 1(e)–1(g). The binarized image of individual color components has different shades of black and white as clearly visible. Two feature vectors of higher and lower intensity values for each color component were computed from the binarized image and were stored for image recognition as shown in Figure 1(h).

The algorithm has been given in Algorithm 1.

Algorithm 1.

Begin

  (1) Input an image I with three different color components R, G, and B, respectively, of size $m * n$ each.

  (2) Calculate the local threshold value Tx for each pixel in each of the color components R, G, and B using Niblack's method:

$$Tx = \left(\frac{1}{m * n}\right) * \sum_{i=1}^{m} \sum_{j=1}^{n} x(i, j) + k$$

$$* \sigma\left(\left(\left(\frac{1}{m * n}\right) * \left(\frac{1}{m * n}\right) * \sum_{i=1}^{m} \sum_{j=1}^{n} x(i, j)\right)\right. \quad (1)$$

$$\left. - x(i, j)\right)^2,$$
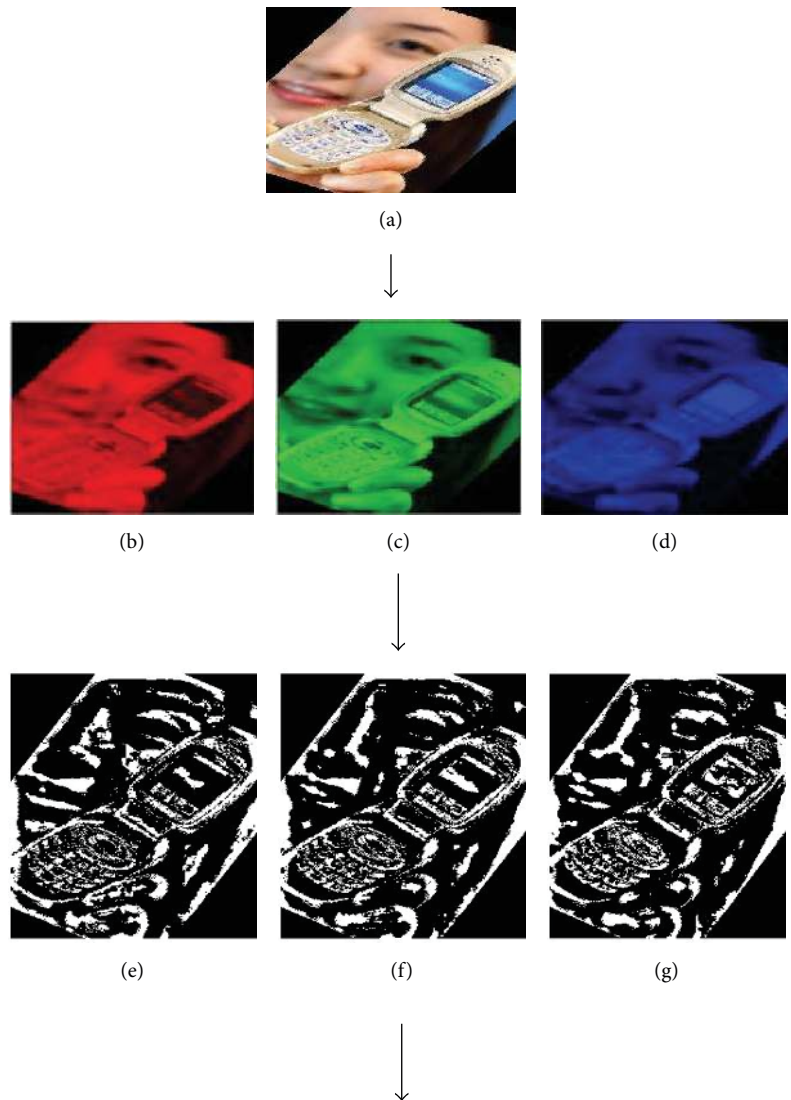
where $k = 0.6$.

  $/^* x = $ R, G and B $^*/$.

  (3) Compute binary image maps for each pixel for the given image:

$$\text{BitMap}_x(i, j) = \begin{cases} 1 & \text{if } x(i, j) > Tx \\ 0 & \text{if } x(i, j) < Tx \end{cases} \quad (2)$$

  $/^* x = $ R, G and B $^*/$.

(a)

(b)　　　　　　　　(c)　　　　　　　　(d)

(e)　　　　　　　　(f)　　　　　　　　(g)

199.2933	125.4946	175.8183	110.2171	164.4336	95.65916	255	255	255	255	255	255.2968
166.1952	85.90133	136.6763	65.23112	132.8543	57.1392	259.1376	278.3944	258.4432	284.6262	258.6312	285.5633
185.4564	76.54298	165.6004	57.77768	151.5858	47.20166	258.062	284.7681	260.363	290.8713	260.9921	292.0651
184.6022	95.15115	168.76	79.95098	159.8879	73.05567	257.7269	274.9995	259.5039	278.455	258.7579	279.4275
179.0961	55.9985	185.2628	61.49882	170.9321	45.0085	258.8851	283.4065	258.0038	281.5065	260.6284	291.729
188.7057	99.45191	156.032	68.90949	144.5566	60.47847	258.5856	274.3959	258.0165	286.0434	259.5283	286.4907
152.673	64.57613	127.1092	42.14967	122.9239	41.58239	259.6199	279.401	260.8346	291.6732	261.2346	290.1892
177.853	76.25525	146.8589	59.53834	143.2831	52.54985	259.0512	278.4283	262.1247	287.933	262.3707	288.3263
154.1838	73.8058	126.5994	53.309	119.9515	46.17131	256.2727	277.9949	257.3016	289.7538	259.0722	289.9174
154.6726	71.11406	148.2096	63.2554	139.1838	54.1262	258.8584	282.6307	259.476	288.4395	260.5965	285.7649
194.1738	112.0663	183.369	81.19479	173.9577	67.24584	257.2266	274.082	258.557	281.3253	260.8187	282.6252
175.4419	96.69125	170.5811	81.63748	150.5545	60.74239	257.6588	279.844	257.9719	277.4522	259.3805	285.7406
167.2141	74.73959	150.5169	50.92828	138.716	41.63226	258.9075	283.4199	266.2363	290.7316	266.0484	290.4801
157.5847	66.1645	147.1577	53.06438	133.0852	42.26251	257.2553	282.1624	258.8852	290.9833	260.5877	291.6093
154.8712	97.29923	173.2174	103.8346	132.2863	77.03318	255.3473	281.5237	255.2684	280.2173	258.166	285.8375
181.083	90.68569	169.4036	65.24256	151.2251	51.15155	256.3134	284.3142	260.1679	291.4813	260.8914	290.2542

(h)

FIGURE 1: (a) Original image. (b) Red component. (c) Green component. (d) Blue component. (e) Binarization of red component. (f) Binarization of green component. (g) Binarization of blue component. (h) Extracted feature vectors.

(4) Generate image features for the given image for each color component:

$$X_{\text{upmean}} = \frac{1}{\sum_{i=1}^{m} \sum_{j=1}^{n} \text{BitMap}_x (i, j)}$$

$$* \sum_{i=1}^{m} \sum_{j=1}^{n} \text{BitMap}_x (i, j) * X (i, j)$$

$$X_{\text{lomean}} = \frac{1}{m * n - \sum_{i=1}^{m} \sum_{j=1}^{n} \text{BitMap}_x (i, j)}$$

$$* \sum_{i=1}^{m} \sum_{j=1}^{n} \left(1 - \text{BitMap}_x (i, j)\right) * X (i, j)$$

$$/^* x = \text{R, G and B} ^*/.$$

$$(3)$$

5. Complexity Analysis

The proposed technique has parted the images into three color components, namely, red, green, and blue. If the total number of gray levels for each component was assumed to be G then linear time $O(G)$ was consumed for selection of threshold for all the gray values in each component. Hence, the number of iterations for three color components was $O(3N)$. Consequently, it was inferred that the time complexity for the feature selection process was linear. Conventional feature extraction techniques have same feature dimension to that of the image from which the feature was extracted. Hence, for an image of size $S * S$ the feature size was S^2. The proposed method has radically reduced the feature size to 12 irrespective of the image dimension. Thus, the space complexity was efficiently addressed with small feature size.

6. Retrieval Architecture

The process of retrieval was carried out by means of classified query as in Figure 2. Conventional retrieval process comprised searching the entire dataset with a generic user query. On the contrary, retrieval with classified query initially classifies the query image into the nearest category of images. The classification process was followed by retrieval of images only from the class of interest. The rest of the image categories were pruned down for a classified query as they do not belong to the native class of the query. Thus, the process has evidently improved the recognition performance compared to state-of-the-art techniques. However, a wrong query classification resulting from bleak feature extraction process would result in zero image retrieval of relevant class as all images will be retrieved from the misclassified category. In case of generic query the scenario is different and in most of the cases a minimum number of image retrieval can be expected from the class of interest. Higher degree of misclassification of retrieval query would have adverse effect on retrieval performance which may be considered as a disadvantage of the proposed technique with respect to existing methods. Nevertheless, it can be avoided by designing effective techniques for robust feature extraction.
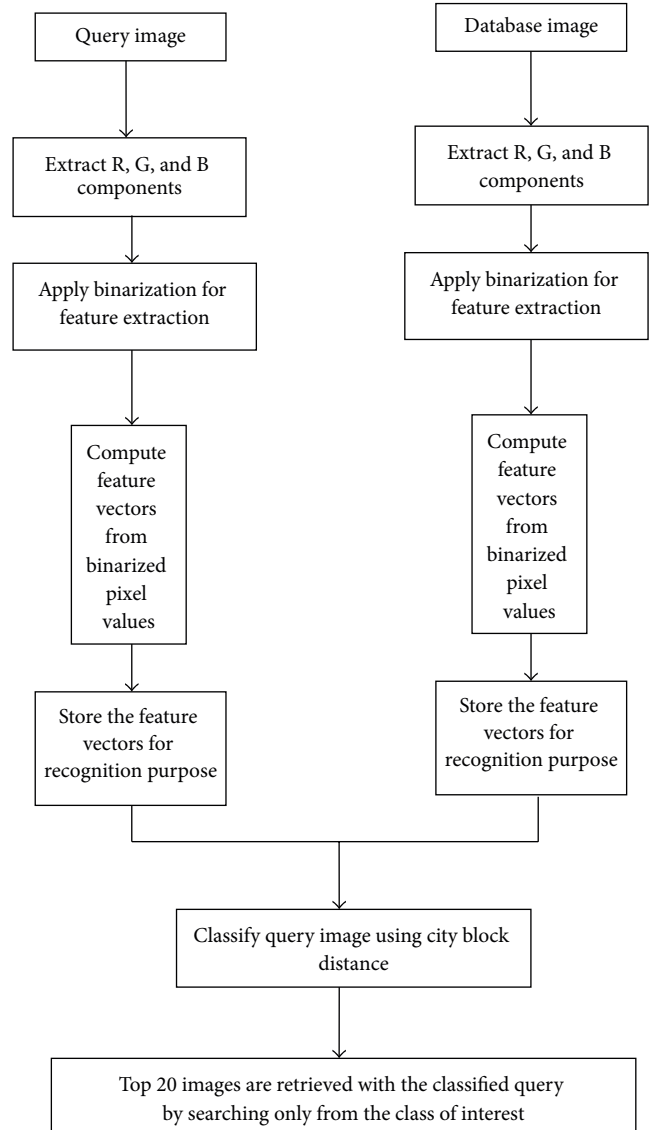

Figure 2: Block diagram of feature extraction and recognition process.

7. Experimental Verification

Two different datasets, namely, Wang dataset (10 categories, 1000 images) and Oliva and Torralba (OT-Scene) dataset (8 categories, 2688 images) were considered for the evaluation purpose. Figures 3 and 4 have illustrated the sample of the datasets used for the experimentation process. The validation process was carried out with 10-fold cross validation in which 9 subsets were considered as training set and 1 subset was considered as the testing set. 10 trials were conducted for the performance assessment of the classifiers. The final decision was made by combining the 10 results thus obtained after evaluating 10-fold cross validation. The evaluation process was performed using three different classifiers, namely, K Nearest Neighbor (KNN), Support Vector Machine (SVM), and Artificial Neural Network (ANN) [36]. KNN was considered as an instance based classifier which has acted based

FIGURE 3: Sample image for Wang dataset.

FIGURE 4: Sample image for OT-Scene dataset.

TABLE 1: Precision and recall for two datasets under different classifier environment.

	Wang dataset			OT-Scene dataset		
	KNN	SVM	ANN	KNN	SVM	ANN
Precision	0.776	0.807	0.838	0.63	0.753	0.665
Recall	0.763	0.804	0.837	0.619	0.754	0.663

	Precision	Recall
■	83.8	83.7
■	65.1	64.8
■	64.7	64.9
■	64.2	64
■	63	63.3
■	62.9	62.4
■	56.7	56.8
■	52.8	51.9

■ Proposed
■ Feature extraction from Image Bit Planes with mean threshold selection (Kekre et al., 2012)
■ Feature extraction from even + odd image using mean threshold (Thepade et al., 2013)
■ Feature extraction by binarization with Bernsen's local threshold technique (Yanli and Zhenxing, 2012)
■ Feature extraction by binarization with Sauvola's local threshold technique (Hamza et al., 2005)
■ Feature extraction by ternary threshold selection (Thepade et al., 2013)
■ Feature extraction by binarization with Niblack's local threshold technique (Valizadeh et al., 2009)
■ Feature extraction by binarization with Otsu's global threshold technique (Shaikh et al., 2013)

FIGURE 5: Graphical illustration for precision and recall comparison.

on similarity functions of two different instances. SVM has followed the learning process of Self-Organizing Map (SOM) which has presumed that only nearby nodes have affected the behavior of each other. Finally, ANN was based on a feed-forward architecture known as multilayer perceptron (MLP). The retrieval process was conducted using city block distance as a measure to match the query image with the database image. The equation for city block distance has been given in

$$D = \sum_{i-1}^{n} \left| Q_i - T_i \right|, \tag{4}$$

where D is distance, Q_i is query image, and T_i is database image.

Two different metrics were considered for evaluation purpose, namely, precision and recall. Precision was defined as the probability that an object is classified correctly as per the actual value and recall was considered as the probability of a classifier to produce true positive results.

Table 1 has given the precision and recall rate for two different datasets, namely, Wang dataset and OT-Scene dataset under three different classifier environments. It was observed that in case of Wang dataset the highest precision was of 0.838 and the highest recall was of 0.837 using ANN classifier. The highest precision rate for OT-Scene dataset was noted to be of 0.753 with a recall rate of 0.754.

The precision and recall rate for classification with proposed technique of feature extraction were further compared with state-of-the-art techniques as shown in Figure 5.

It was clearly observed that classification with proposed technique of feature extraction has outperformed the existing techniques.

Hypothesis 1. Classification with the proposed method of feature extraction has outclassed the existing techniques.

TABLE 2: Paired *t*-test for significance testing of precision results for classification.

Comparison	*t*-calc	*p* value	Significance of difference in value
Feature extraction by binarization using bit plane slicing with Niblack's local threshold method (Thepade et al., 2014 [9])	3.1626	0.0133	Significant
Feature extraction by binarization with multilevel mean threshold (Kekre et al., 2013 [19])	3.1626	0.0133	Significant
Feature extraction by binarization using Bit Plane Slicing with mean threshold (Kekre et al., 2013 [19])	2.9059	0.0197	Significant
Feature extraction by binarization of original + even image with mean threshold (Thepade et al., 2013 [20])	3.4116	0.0092	Significant
Traditional feature extraction by binarization with Bernsen's local threshold method (Yanli and Zhenxing, 2012 [26])	3.2593	0.0115	Significant
Traditional feature extraction by binarization with Sauvola's local threshold method (Ramírez-Ortegón and Rojas, 2010 [23])	3.0157	0.0167	Significant
Traditional feature extraction by binarization with Niblack's local threshold method (Liu, 2013 [22])	3.9038	0.0045	Significant
Traditional feature extraction by binarization with Otsu's global threshold method (Shaikh et al., 2013 [24])	3.661	0.0064	Significant

Table 2 has shown the significance of *p* values obtained from precision comparison and hence the null hypothesis of equal precision rate for the proposed algorithm and existing algorithms was rejected. Hence it was inferred that the proposed method has been capable of improving the classification performance radically. Further, the proposed technique of feature extraction was tested for retrieval performance. The performance was assessed by comparing retrieval with classified query and retrieval with generic query. Table 3 has shown the category wise comparison of precision for retrieval for both the retrieval techniques. The comparison in Table 3 has revealed that retrieval with classified query has improved the identification for each category of images in the Wang dataset. Hence the overall performance of retrieval has increased considerably. The average precision for retrieval with classified query has also outperformed the conventional retrieval technique with generic query.

Hypothesis 2. Retrieval with classified query cannot give higher precision results compared to retrieval with generic query.

p value in Table 4 has shown significance and hence Hypothesis 2 was rejected. Henceforth, the precision results of proposed technique of feature extraction for retrieval with classified query were compared to state-of-the-art techniques of retrieval as in Table 5.

Hypothesis 3. Retrieval with classified query with proposed feature extraction technique has higher precision compared to the existing techniques for product recognition for diversified product categories.

TABLE 3: Category wise comparison for retrieval with classified query and generic query.

Categories	Retrieval with classified query	Retrieval with generic query
Tribals	80	54
Sea beach	60	48
Gothic structure	60	39
Bus	80	61
Dinosaur	100	100
Elephant	80	65
Roses	80	70
Horses	100	100
Mountains	80	56
Food	60	54
Average precision	**78**	**64.7**

TABLE 4: Paired *t*-test for significance of retrieval results with classified query.

Comparison	*t*-calc	*p* value	Significance of difference in value
Retrieval with generic query and retrieval with classified query	4.4888	0.0015	Significant

It was evidently revealed in Table 5 that the proposed method has higher performance compared to the existing techniques and has outclassed the category wise precision

TABLE 5: Category wise comparison of precision results of the proposed technique with the state-of-the art techniques.

Categories	Proposed	Subrahmanyam et al., 2012 [34]	Jalab, 2011 [30]	Walia and Pal, 2014 [25]	Hiremath and Pujari, 2007 [28]	Yue et al., 2011 [35]	Rahimi and Moghaddam, 2015 [33]
Tribals	80	54.1	32.3	54	48	58.75	41.8
Sea beach	60	48.3	61.2	51	34	41.19	44.6
Gothic structure	60	40.2	39.2	38	36	42.35	49.2
Bus	80	89.5	39.5	46	61	71.69	30.5
Dinosaur	100	97.7	99.6	100	95	74.53	55.4
Elephant	80	39.1	55.7	63	48	65.08	69.9
Roses	80	85.9	89.3	90	61	83.24	61.1
Horses	100	81.4	65.2	93	74	69.3	57.2
Mountains	80	32	56.8	48	42	44.86	43.7
Food	60	41	44.1	39	50	44.54	43.5
Average precision	78	60.9	58.29	62.2	54.9	59.55	49.69

TABLE 6: Paired t-test for significance of precision results for retrieval.

Comparison	t-calc	p value	Significance of difference in value
Subrahmanyam et al., 2012 [34]	2.9118	0.0173	Significant
Jalab, 2011 [30]	3.3348	0.0087	Significant
Walia and Pal, 2014 [25]	3.5295	0.0064	Significant
Hiremath and Pujari, 2007 [28]	7.1826	0.0001	Significant
Yue et al., 2011 [35]	5.3223	0.0005	Significant
Rahimi and Moghaddam, 2015 [33]	5.8345	0.0002	Significant

results for each of the existing techniques. The statistical significance of retrieval with the proposed technique was established by a paired t-test and the result has been given in Table 6. The test was carried out to validate that the difference in precision results was not generated from a population with zero mean.

Table 6 has shown significant p values for the comparison of precision for retrieval with proposed feature extraction technique with respect to the existing techniques. Hence Hypothesis 3 was accepted and the supremacy of the proposed method for content based image retrieval was established.

Hypothesis 4. Image data based query and text data based query in digital marketplace have similar impact on consumer satisfaction.

Analysis in Table 7 has shown significant association in between consumer satisfaction and image data based query (likelihood ratio = 79.270; Phi = 0.734; Cramer's V = 0.424; and $p < 0.01$) related to product recognition in diversified product categories compared to text data based query (likelihood ratio = 18.577; Phi = 0.296; Cramer's V = 0.210; and $p < 0.05$). It was observed that image data based query was having greater association with consumer satisfaction related to product recognition with respect to text data based query. Henceforth, the relation between consumer satisfaction and product query based on image data has been analyzed in Table 8.

From the above results shown in Table 8, significant correlation was observed between consumer satisfaction and query based on image data related to product recognition in diversified product categories. On the other hand it was detected that significant negative correlation lies in between consumer satisfaction and query based on text data related to product recognition in diversified product categories.

8. Conclusion

The model of digital marketplace has enriched the firms with competitive strategy to gain technical payback from their competitors. The authors have investigated the application orientation of innovation information identification by image data. The study has addressed the challenge of relevant result generation for customer query and has efficiently suggested a content based search method for relevant output generation. The novel technique of product identification based on query with image data has envisioned a new direction for innovation and value addition for using information technology in organizational development. The technique has outclassed all the state-of-the-art techniques and has increased customer satisfaction in product recognition from

TABLE 7: Chi-square test for consumer satisfaction analysis related to product recognition in diversified product categories in digital environment.

| | Chi-square tests | | | |
| | Raw variable → consumer satisfaction related to product recognition in diversified product categories | | | |
Column variables ↓	Likelihood ratio	Sig. (2-sided)	Phi	Cramer's V
Image data based query	79.270	.000	.734	.424
Text data based query	18.577	.039	.296	.210

Number of samples 185.

TABLE 8: Nonparametric correlation test for consumer satisfaction analysis related to product recognition in diversified product categories in online environment.

Kendall's tau: nonparametric correlation coefficient		Consumer satisfaction related to product recognition in diversified product categories in online environment
Image data based query	Correlation coefficient	.498
	Sig. (2-tailed)	.000
Text data based query	Correlation coefficient	−0.32
	Sig. (2-tailed)	0.643

multiple product categories. It has efficiently pruned the inappropriate outcomes for customer queries by replacing text based search process with image query formulation. The work can be extended towards unstructured data analysis and e-commerce and for other analytical activities related to consumer orientated revenue generation for modern business process.
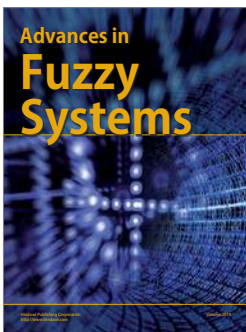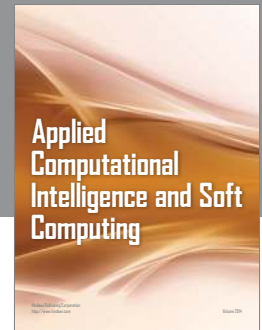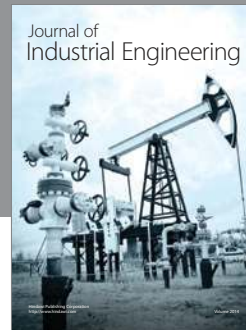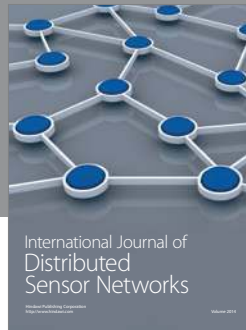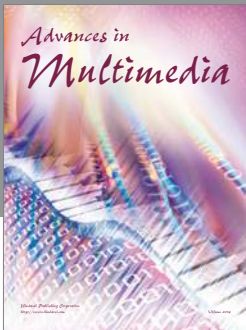
Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] Y. Hu, H. Yin, D. Han, and F. Yu, "The application of similar image retrieval in electronic commerce," *The Scientific World Journal*, vol. 2014, Article ID 579401, 7 pages, 2014.

[2] J. Zhai, L. Shen, Y. Liang, and J. Jiang, "Application of fuzzy ontology to information retrieval for electronic commerce," in *Proceedings of the International Symposium on Electronic Commerce and Security (ISECS '08)*, pp. 221–225, Guangzhou, China, August 2008.

[3] N. Gupta, "Globalization does lead to change in consumer behavior: an empirical evidence of impact of globalization on changing materialistic values in Indian consumers and its after-effect," *Asia Pacific Journal of Marketing and Logistics*, vol. 23, no. 3, pp. 251–269, 2011.

[4] J. G. Maxham III, "Service recovery's influence on consumer satisfaction, positive word-of-mouth, and purchase intentions," *Journal of Business Research*, vol. 54, no. 1, pp. 11–24, 2001.

[5] D. Su and X. Huang, "Research on online shopping intention of undergraduate consumer in China—based on the theory of planned behavior," *International Business Research*, vol. 4, no. 1, pp. 86–92, 2011.

[6] D. Chaffey, *E-Business and E-Commerce Management—Strtaegy, Implementation and Practice*, Prentice Hall, 2011.

[7] R. Das, S. Thepade, and S. Ghosh, "Framework for content-based image identification with standardized multiview features," *ETRI Journal*, vol. 38, no. 1, pp. 174–184, 2016.

[8] R. Datta, D. Joshi, J. Li, and J. Z. Wang, "Image retrieval: ideas, influences, and trends of the new age," *ACM Computing Surveys*, vol. 40, no. 2, article 5, 2008.

[9] S. Thepade, R. Das, and S. Ghosh, "A novel feature extraction technique using binarization of bit planes for content based image classification," *Journal of Engineering*, vol. 2014, Article ID 439218, 13 pages, 2014.

[10] R. Das, S. Thepade, and S. Ghosh, "Multi technique amalgamation for enhanced information identification with content based image data," *SpringerPlus*, vol. 4, article 749, 2015.

[11] R. Das, S. Thepade, and S. Ghosh, "Content based image recognition by information fusion with multiview features," *International Journal of Information Technology and Computer Science*, vol. 7, no. 10, pp. 61–73, 2015.

[12] R. Das, S. Thepade, and S. Ghosh, "Novel technique in block truncation coding based feature extraction for content based image identification," in *Transactions on Computational Science XXV*, vol. 9030 of *Lecture Notes in Computer Science*, pp. 55–76, Springer, Berlin, Germany, 2015.

[13] H. Hamza, E. Smigiel, and A. Belaid, "Neural based binarization techniques," in *Proceedings of the 8th International Conference on Document Analysis and Recognition (ICDAR '05)*, vol. 1, pp. 317–321, IEEE, September 2005.

[14] S. Thepade, R. Das, and S. Ghosh, "A novel feature extraction technique with binarization of significant bit information," *International Journal of Imaging and Robotic*, vol. 15, no. 3, pp. 164–178, 2015.

[15] D. Thepade, R. Das, and S. Ghosh, "Content based image classification with thepade's static and dynamic ternary block truncation coding," *International Journal of Engineering Research*, vol. 4, no. 1, pp. 13–17, 2015.

[16] Y.-F. Chang, Y.-T. Pai, and S.-J. Ruan, "An efficient thresholding algorithm for degraded document images based on intelligent block detection," in *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics (SMC '08)*, pp. 667–672, Singapore, October 2008.

[17] B. Gatos, I. Pratikakis, and S. J. Perantonis, "Efficient binarization of historical and degraded document images," in *Proceedings of the 8th IAPR International Workshop on Document Analysis Systems (DAS '08)*, pp. 447–454, IEEE, Nara, Japan, September 2008.

[18] M. Valizadeh, N. Armanfard, M. Komeili, and E. Kabir, "A novel hybrid algorithm for binarization of badly illuminated document images," in *Proceedings of the 14th International CSI Computer Conference (CSICC '09)*, pp. 121–126, IEEE, Tehran, Iran, October 2009.

[19] H. B. Kekre, S. Thepade, R. K. Kumar Das, and S. Ghosh, "Multilevel Block Truncation Coding with diverse color spaces for image classification," in *Proceedings of the International Conference on Advances in Technology and Engineering (ICATE '13)*, pp. 1–7, Mumbai, India, January 2013.

[20] S. Thepade, R. Das, and S. Ghosh, "Performance comparison of feature vector extraction techniques in RGB color space using block truncation coding for content based image classification with discrete classifiers," in *Proceedings of the 10th Annual Conference of the IEEE India Council (INDICON '13)*, Mumbai, India, December 2013.

[21] H. B. Kekre, S. Thepade, R. Das, and S. Ghosh, "Performance boost of block truncation coding based image classification using bit plane slicing," *International Journal of Computer Applications*, vol. 47, no. 15, pp. 45–48, 2012.

[22] C. Liu, "A new finger vein feature extraction algorithm," in *Proceedings of the 6th International Congress on Image and Signal Processing (CISP '13)*, pp. 395–399, IEEE, Hangzhou, China, December 2013.

[23] M. A. Ramírez-Ortegón and R. Rojas, "Unsupervised evaluation methods based on local gray-intensity variances for binarization of historical documents," in *Proceedings of the 20th International Conference on Pattern Recognition (ICPR '10)*, pp. 2029–2032, IEEE, Istanbul, Turkey, August 2010.

[24] S. H. Shaikh, A. K. Maiti, and N. Chaki, "A new image binarization method using iterative partitioning," *Machine Vision and Applications*, vol. 24, no. 2, pp. 337–350, 2013.

[25] E. Walia and A. Pal, "Fusion framework for effective color image retrieval," *Journal of Visual Communication and Image Representation*, vol. 25, no. 6, pp. 1335–1348, 2014.

[26] Y. Yanli and Z. Zhenxing, "A novel local threshold binarization method for QR image," in *Proceedings of the IET International Conference on Automatic Control and Artificial Intelligence (ACAI '12)*, pp. 224–227, Xiamen, China, March 2012.

[27] M. E. El Alami, "A novel image retrieval model based on the most relevant features," *Knowledge-Based Systems*, vol. 24, no. 1, pp. 23–32, 2011.

[28] P. S. Hiremath and J. Pujari, "Content based image retrieval using color, texture and shape features," in *Proceedings of the 15th International Conference on Advanced Computing and Communication (ADCOM '07)*, pp. 780–784, Guwahati, India, December 2007.

[29] M. Banerjee, M. K. Kundu, and P. Maji, "Content-based image retrieval using visually significant point features," *Fuzzy Sets and Systems*, vol. 160, no. 23, pp. 3323–3341, 2009.

[30] H. A. Jalab, "Image retrieval system based on color layout descriptor and Gabor filters," in *Proceedings of the IEEE Conference on Open Systems (ICOS '11)*, pp. 32–36, IEEE, Langkawi, Malaysia, September 2011.

[31] G. L. Shen and X. J. Wu, "Content based image retrieval by combining color, texture and CENTRIST," in *Proceedings of the IEEE International Workshop on Signal Processing*, vol. 1, pp. 1–4, London, UK, January 2013.

[32] A. Irtaza, M. A. Jaffar, E. Aleisa, and T.-S. Choi, "Embedding neural networks for semantic association in content based image retrieval," *Multimedia Tools and Applications*, vol. 72, no. 2, pp. 1911–1931, 2014.

[33] M. Rahimi and M. E. Moghaddam, "A content-based image retrieval system based on Color Ton Distribution descriptors," *Signal, Image and Video Processing*, vol. 9, no. 3, pp. 691–704, 2015.

[34] M. Subrahmanyam, R. P. Maheshwari, and R. Balasubramanian, "Expert system design using wavelet and color vocabulary trees for image retrieval," *Expert Systems with Applications*, vol. 39, no. 5, pp. 5104–5114, 2012.

[35] J. Yue, Z. Li, L. Liu, and Z. Fu, "Content-based image retrieval using color and texture fused features," *Mathematical and Computer Modelling*, vol. 54, no. 3-4, pp. 1121–1127, 2011.

[36] M. H. Dunham, *Data Mining Introductory and Advanced Topics*, Pearson Education, 2009.