# **RIO: Rhetorical Structure Theory Based Indexing Technique for Image Objects**

Syed Khaldoon Khurshid and Muhammad Shoaib

Department of Computer Science and Engineering, University of Engineering and Technology, Pakistan

Abstract: Efficient and relevant retrieval of any type of the data, especially multimedia objects, is based on indexing technique. Due to complexity of multimedia data, existing indexing techniques for multimedia objects suffer from irrelevant retrieval. Rhetorical Structure Theory (RST) has already been successfully implemented for indexing text documents, which has reduced irrelevancy. The focus of this research paper is to propose an indexing technique for image objects using RST. An indexing technique is proposed for image objects using image relation framework for RST. To elaborate the functionality of proposed indexing technique, a case study is presented for image objects. Further research can be carried out for indexing technique of other multimedia objects.

Keywords: Indexing techniques, RST, image indexing.

Received September 21, 2011; accepted December 30, 2011; published online August 5, 2012

# 1. Introduction

Since 1970, Image retrieval is an important topic of research and development. The major two disciplines of Computer Science, database management and computer vision contributed a lot, in the relevant retrieval of images. The researchers have studied relevant retrieval of images in great detail. The approach of retrieving an image can be based on text or visual data. The problem with text based image retrieval is that it is limited to manual or automatic annotation provided by people, which resulted in low relevancy [19].

Feature based approach [11], opened up new opportunities in the relevant retrieval of image without using the text affiliated with or within the image objects. The image objects are mapped on the high dimensioned feature vectors; after their automatic extractions of low level features. But with advent of time, the research, to reduce the gap between the high level features of an image object and the understanding of human being's semantic, gained more importance as compared to finding low level features extraction algorithms for image objects [21].

Rhetorical Structure Theory (RST) is framework for analyzing the text data [13]. RST declares two different parts. Nuclei are considered more important parts than satellites which provide support to the nuclei. The relationship among Nucleus and satellites are specifically described by their relations. For example in Evidence relation, the nucleus is the sentence which gives the main idea or focus of the material, and the satellite contains the sentence which provides evidence sentence [14]. For example:

- Sentence 1: Quaid-i-Azam was the first Governor General of Pakistan.
- *Sentence 2:* He took an oath of the office on 14th Aug, 1947.

In this example, first sentence can be considered as nucleus which is providing the main information and second sentence as satellite supporting the first sentence with evidence describing a rhetorical relation. The satellite cannot be understood without nucleus; but on the other hand vice versa is incorrect; text related to nucleus is understandable to a certain limit, even if the satellites have been deleted [13]. RST has successfully reduced the irrelevancy issue but RST is developed only for text documents [19]. A lot of efforts have been undertaken to attain benefits from RST in other fields [3, 4, 5, 6, 7, 12, 17, 22].

A relation is a core concept between nucleus and satellite. The RST and its relations are useful for text data only [21]. To use RST for multimedia objects it was necessary to modify these relations. These relations were modified to accommodate the multimedia objects [10]. A novel indexing technique for image objects using these modified relations has been proposed in this paper. The proposed indexing technique is based on the relevant retrieval of images. The idea is to use low features of an image and convert them according to the semantics of human beings. The feature based approach [23] has been used to describe the content of an image in the proposed framework of modified RST relations. The probability of relevant retrieval of image can be enhanced due to its context as well as its content indexing.

The remainder organization of the paper is as follows: In the next section, related work of image retrieval is presented. In the 3rd section, the proposed indexing technique has been described. In the 4th section, a case study is given to demonstrate the proposed indexing technique. Finally, conclusion and future work is given in the section 5.

## 2. Material and Methods

Text indexing is based on keywords or term frequency. Text indexing focuses on how many times a term have been appeared in the document? Searching is based on the occurrence of a term in a particular document. A term is given importance or weighted and searching is solely based on keywords provided [18]; whereas image cannot be indexed by keywords only [21]. Feature based approach [11] is widely accepted approach. The main idea is to replace keywords with the feature vectors of an image. A number of commercial products have been launched with respect to feature based approach. Content Based Image Retrieval (CBIR) [24] system is the most successful and it retrieved desired images from a huge collection of image database [8]. The retrieval is based on the low level features of an image which can be extracted from the images without human intervention. The problem with CBIR is the gap between the high level semantic features and the low-level features extracted from images. High level semantics means concepts used by humans, and low-level features are like size, shape, texture and color of an image object [23].

To address this problem, researchers realized that the in feature based similarity searching algorithms; they have neither developed efficient nor user friendly algorithms as they had anticipated. An efficient approach to retrieve large amount of images with good precision was required [16]. Retrieval accuracy can be increased by focusing on the research which has been transferred from developing low level feature extraction algorithms to decrease the semantic space between the visual based features and the human semantics [23].

RST is an established theory to understand the semantics of the text due to its content [10]. Hence using RST concepts the semantic gap between the features of an image and human semantics can be decreased.

# 3. The RST-Based Image Indexing Technique

The following case is presented to give the understanding of the proposed technique.

## 3.1. Case for Elaboration

Consider a rectangular box of definite size, shape and color. Searching this box is easy if its features are

known. If it is impossible to find the exact box, it may be possible to find box with similar features. In the next step, divide the box in different parts and marked them with different numerical values, as shown in Figure 1-a and b. Now if a specific segment of a box is required to be found, search only the numerical value of the that particular segment. The chances of finding similar part in another box are greater if it contains the similar numerical value. Finally, consider parts within a box are not only marked but also correlated with one another through pointers. It means that the specific part of the box can be retrieved as well as the parts related to that specific part will also be retrieved. The parts correlated within the box are shown in the Figure 1-c and by searching part 2 of the box, parts 16 and 18 of box are also being retrieved.

			-									
								1	2	3	4	5
							(	6	7	8	9	10
							1	1	12	13	14	15
							1	6	17	18	19	20
a) P	a) Partition of rectangular box. b) Allocating segment numbers.							bers.				
		1		2		3			4	5		
		6	1/	7		8			9	10		
		11	Ι	12		13			14	15		
		16		17		<b>■</b> 18			19	20		
c) Relations among segments.												

Figure 1. Case for elaboration.

Now it is pivotal to understand that anything which is being searched within a box must be converted into the box's format (Divided into parts, allocating of numerical value to each part and finding or giving relations among the parts of the box). Proposed model is based on this concept.

## 3.2. The Model of Architecture

The architecture of indexing technique for image objects using RST is proposed. It consists of four components: image segmentation, image relation framework, image tree and indexing image rhetorically as shown in the Figure 2.



Figure 2. Architecture of model.

Each of these components is explained as follows:

#### 3.2.1. Image Segmentation

Image objects have their own structure and properties which differentiate them from other data. To consider segments within an image is more likely to yield better results than considering a whole image [23]. Process starts with image ID which will be provided as input. Segmentation of images is taken from Berkeley Segmentation Data Set-300 (BSDS-300) [15], to have optimal solution for segmentation problem.

More practical properties of image like 'improved effectiveness of the color coherence Vector' can be used [1], but two simple properties of digital image are given. Pixel Per Inch (PPI) and color depth [2]. Instead of considering multiple images on computer screen; each segment of an image is represented by a single rectangular screen; a hypothetical rectangles, each having specific length and breath. Three values (R, B, G) of color depth of a segment are taken due to less complex calculations. The resolution of an image is related to the number of pixels presented within the digital file, and is measured in PPI. As more pixels are stored in inch of an image; clearer will be the image. Similarly as the density of the colour information will be greater, the details of the image will be clearer and obvious [2]. One screen rectangle has been supposed for one segment of image for the calculation of PPI. Segments of an image are bounded by hypothetical rectangles. Instead of affiliating single satellite with single nucleus as in the case of text [21], Modified RST [10] is affiliate with two satellites (trivial segments). Three segments within an image are taken, for the proposed model. As mentioned in [2], PPI can be calculated as follows:

$$PPI = d_p / d_i \tag{1}$$

where, diagonal resolution in pixels can be calculated by Pythagorean Theorem, as mentioned in [9], as follows:

$$dp = \sqrt{w_p^2 + h_p^2} \tag{2}$$

where in pixels:  $d_p$ =diagonal resolution.  $w_p$ =width resolution.

 $h_p$ =height resolution,

 $\dot{d_i}$ =diagonal size in inches.

Digital colored pictures color depth of 24bit or three (23) 8bit color representation for each pixel of the segment, is considered [1], for example: color depth for Grey=(84, 84, 84).

#### **3.2.2.** Construction of Image Relation Framework

PPI and Color depth values of segments of an image are calculated for nucleus and satellites. One segment is marked as nucleus and other two segments as its satellites. The nucleus of the image will be decided on the basis of the desired PPI values and color depth of the segment. These segments of an image are represented in three tuple image relation framework. Within the framework, segment dictionary will be constructed. Segment dictionary contains segment IDs which are tagged to different segments within an image, with respect to given categories, as shown in the Table 1, in section 4. Segment tables are constructed by human subjects on the basis of most commonly searched categories in the image database.

Following two equations 3 and 4 in image relation frame are proposed, which contains multimedia nucleus and satellites:

$$N_m = (PPI, Color depth, Segment ID)$$
 (3)

$$S_{mn} = (PPI, Color depth, Segment ID)$$
 (4)

where, 'm' is a constant value; and it identifies the multimedia Nucleus along with its multimedia satellites and 'n' is a variable value which relates satellite with its Nucleus.

## **3.2.3.** Construction of Image Tree

Based on information provided by Image Relation Framework, Image tree will be constructed to identify the Multimedia Relations;  $R_{mn}$  as follows:

$$N_m \longrightarrow S_{mn}$$
 (5)

In image tree, nodes will represent the multimedia nucleus i.e., salient segment,  $(N_m)$  and trivial segments i.e., multimedia satellites  $(S_{mn})$  & edges will represent the multimedia relations  $(R_{mn})$ , as shown in the Figure 3.



Figure 3. General image tree.

Nucleus is the located at the root node and satellites become its children. Multimedia relations are simply pointers from salient segment to trivial segments. Retrieval of nucleus node along with the satellite nodes due to these established relations.

#### 3.2.4. Construction of Proposed Rhetorical Based Indexers

Proposed indexer is composed of following parameters: Image ID which will be allocated before segmentation, segment ID of salient segment is only considered for salient indexer and it is being allocated by segment dictionary, after segmentation. The values of PPI and color depth of the segment within an image is calculated during construction of image relation framework. Segment phrase is attained during image tree construction, Segment phrase contains segment "title/ caption/ words/ sentences" within an image, for context based retrieval and finally weight parameter is included which indicates the frequency of searches for the particular salient segment of an image within a database. Weight will be incremented as the number of searches for the particular segment increases. The general view of indexer is represented by Figure 4, mentioning image depended part and segment depended parts.

Image ID	Segment	Pixel	Color	Segment	Weight
	ID	Per Inch	Depth	Phrase	
Numeric	Alphanume	Numeric	Three	"Title/Caption	Numer
value	ric value	value	numeric	/ words/	ic
	(Salient		values	sentences"	value
	Segment)				

Segments depended parts

Mage depended part

Figure 4. General view of proposed indexer.

The final output is the population of RST based indexer containing the values extracted from the salient segment of an image. The content based retrieval is possible due to the index terms; segment ID, Weight PPI and color depth values of a segment. Context based retrieval is possible due to the segment phrases only, as shown in the following Figure 5.

Content based indexing

	$\sim$				
Image	Segment	Pixel Per	Color	Weight	Segment
ID	ID	Inch	Depth		Phrase
Numer ic value	Alphanume ric value (Salient Segment)	Numeric value	Three numeric values	Numeric value	"Title/Caption / words/ sentences"

Context based indexing

Figure 5. Content and Context based indexing.

Similarly the related trivial indexer was created after comparing the image ID value with the salient indexer. The information related to trivial segments having same image ID has been saved through relation pointers as shown in Figure 6.

Ima ID	age Segment		Pixel Per Inch		Color S Depth P		Segm	Segment Phrase		nt	
Nur valı	Numeric Alphanumeric value value (Salient Segment)		Nur valu	Numeric value		e eric es	"Title/Caption/ words/ sentences"		Numer value	ric	
	Imag ID Num value	eric	Segment ID Alphanum value (Sal Segment)	eric ient	Pixel Per I Num value	l nch leric e	Colo Dept Thre num value	er e eric es	Segment Phrase "Title/Cap words/ sentences'	otion/	
	Imag ID Num	eric	Segment ID Alphanum	ieric	Pixel Per I Num	l nch leric	Colo Dept Thre	or th e	Segment Phrase "Title/Cap	otion/	
	value Va Se		Segment)	Segment)		value		es	sentences'	,	

Figure 6. Relations among salient and trivial segments.

By attaching trivial segments with the salient segment of a particular image not only data is saved from being wasted but the probability of relevant retrieval of image have also increased as more information related to an image is being stored.

## 4. Results and Discussions

For the narration of proposed indexing technique as a case study, an image Figure 7 is taken from Berkeley BSDS-300, produced by user #1109 [15]. This particular image is being selected because it gives three different segment categories which cover the wide scope of the proposed solution as well as makes it easy to explain each step of process clearly.

*Case Study:* Image ID 01=A small image of brown eagle sitting on the branch of tree with blue sky in the background.



Figure 7. Case study image from (BSDS-300) [15].

The segmentation [15] of the case study, image is shown in Figure 8 as follows:



Figure 8. Segmentation process by user# 1109 of BSDS-300.

Twenty two segments are being produced by user #1109 [15]. For proposed indexing technique, only three segments are chosen from these segments. Hypothetical rectangles are being drawn which are marked over an image to differentiate the chosen three segments as shown in the following Figure 9.



Figure 9. Identifying hypothetical rectangles.

Segment IDs are tagged to segments from segment dictionary, constructed in Table 1, is shown in Figure 10.

Table 1. Segment dictionary for image objects.



Figure 10. Tagging segments IDs.

Three segments of an image are acquired and marked as *Seg2*, *Seg3* and *Seg5*. *Seg2* is decided as Multimedia Nucleus ( $N_1$ ) based on the identification of 'Eagle' segment after segmentation and desired color density of brown as (150, 75, 0). Figure 11 represents Seg2 as follows: diagonal size in inches  $d_i$ =6, width resolution in pixels;  $w_p$ =4 and height resolution in pixels  $h_p$ =7.



Figure 11. Calculations of  $d_p$  and PPI values.

Using equation 2,  $d_p = \sqrt{w_{p+}^2 h_p^2}$  [9], is used to calculate diagonal resolution in pixels of Seg2 and by rounding off the value:

$$=\sqrt{(4)^2 + (7)^2} = 80.62 \approx 81 \tag{6}$$

Using equation 1,  $PPI=d_p/d_i$  [2], is used to calculate Pixel Per Inch for Seg2 and by rounding off the value:

$$= d_p / d_i = 81/6 = 13.5 \approx 14 \tag{7}$$

Similarly segments *Seg5* and *Seg3* are taken as, "sky"  $(S_{10})$  and "branch of a tree"  $(S_{11})$  and their PPI values are calculated as 16 and 28 from Figures 12 and 13, respectively, after repeating the same above calculations.



Figure 12. Seg5 (sky) PPI value is calculated.



Figure 13. Seg3 (the branch of tree) PPI value is calculated.

Taking color densities of sky as blue (176, 226, 255) and branch of tree as wood brown color (92, 64, 51) and putting all values in equations 3 and 4 for  $N_I$ ,  $S_{I0}$  and  $S_{II}$  to get equations 8, 9 and 10.

$$N_l = (14, (150, 75, 0), Seg2)$$
 (8)

$$S_{10} = (16, (92, 64, 51), Seg3)$$
 (9)

$$S_{II} = (28, (176, 226, 255), Seg5)$$
 (10)

Using equation 5,  $R_{10}$  and  $R_{11}$  will be represented as follows in equations 11 and 12:

$$R_{10} = N_1 \longrightarrow S_{10} \tag{11}$$

$$R_{II} = N_I \longrightarrow S_{II} \tag{12}$$

Now image relation tree is constructed as shown in Figure 14, based on equations 11 and 12.



Figure 14. Image tree of example.

Suppose the number of searches is 54 for segment "Eagle". The values are taken from equation 8 to populate root node  $(N_l)$  in the proposed RST based indexer as is shown in Table 2.

Table 2. Indexer values.

Image ID	Segment ID	PPI	Color Depth	Segment Phrase	Weight
01	Seg2	14	(150, 75, 0)	" Eagle"	54

Similarly if 20 sample images are taken from User #1109; BSDS-300 [15], and then each image is converted into the similar framework format as

presented in the above described case study. Proposed RST based Indexer for salient segment  $(N_I)$  can be populated as shown in the Table 3.

Image ID	Segment ID	PPI	Color Depth	Segment Phrase	Weight
01	Seg2	14	(150, 75, 0)	"Eagle"	54
02	Seg6	26	(255, 255, 240)	"Boat"	47
03	Seg6	28	(139, 139, 131)	"opening bridge"	42
04	Seg5	32	(190,190,190)	"Mountains"	41
05	Seg1	12	(255, 255, 240)	"Skiing Man"	36
06	Seg6	24	(107, 142, 35)	"House with lake"	33
07	Seg6	31	(245, 255, 250)	"Boats in sea"	31
08	Seg5	34	(142, 107, 35)	"Bears with river"	29
09	Seg4	24	(210, 180, 140)	"Two Lions"	28
10	Seg2	20	(139, 139, 131)	"Penguin"	28
11	Seg4	22	(139, 58, 58)	"Horses"	25
12	Seg5	34	(255, 255, 255)	"Mountains in sea"	24
13	Seg5	24	(217, 217, 243)	"Rushing water"	23
14	Seg2	28	(0,0,0)	"Bird"	21
15	Seg6	28	(84, 84, 84)	"Old bridge"	20
16	Seg1	24	(255, 255, 240)	"Girl Standing"	17
17	Seg1	22	(250, 250, 210)	"Boy praying"	17
18	Seg1	28	(54, 54, 54)	"Two men"	16
19	Seg3	30	(255, 193, 37)	"Bonsai tree"	13
20	Seg1	38	(255 255 240)	"Cricketers"	12

Table 3. Salient indexer  $(N_I)$ .

Table 4. First trivial indexer  $(S_{10})$ .

Image ID	Segment ID	PPI	Color Depth	Segment Phrase
01	Seg5	28	(176, 226, 255)	" Sky"
02	Seg6	10	(0, 0, 0)	"Sky"
03	Seg6	22	(34, 139, 34)	"Trees"
04	Seg5	16	(255, 228, 196)	"Sky"
05	Seg5	20	(255, 250, 250)	"Snow"
06	Seg5	32	(34, 139, 34)	"Trees"
07	Seg5	24	(176, 226, 255)	"Sky"
08	Seg5	24	(209, 209, 209)	"Rocks"
09	Seg3	26	(205, 149, 12)	"Grass"
10	Seg5	20	(238, 233, 233)	"Snow"
11	Seg5	32	(173, 255, 47)	"Grass"
12	Seg3	16	(34, 139, 34)	"Trees"
13	Seg1	18	(255, 255, 240)	"People"
14	Seg3	18	(205, 170, 125)	"Tree"
15	Seg5	24	(255, 255, 255)	"Waterfall"
16	Seg3	28	(92, 64, 51)	"Old tree"
17	Seg5	28	(245, 245, 220)	"Rocks"
18	Seg6	38	(205, 170, 125)	"Old walls"
19	Seg5	36	(139, 115, 85)	"Soil"
20	Seg3	34	(124, 252, 0)	"Ground"

Weight will be assigned on the basis of number of searches for a particular segment of an image within the database. The most searched segments of an image will be allocated in the top slots of proposed RST based indexer while the least searched image segments will be placed at the bottom slots of the proposed indexer.

Similarly Tables 4 and 5 are being generated, which represent the population of trivial segments (Satellites;  $S_{10}$  and  $S_{11}$ ) of the images in the similar manner. With the inclusion of trivial segments more information is being saved. Like, in case study salient image segment "Eagle" is processed and stored and the trivial image segments, the "Branch of a tree" and "Sky" are also being processed and saved in two separate indexers. Simply by establishing pointers through image ID

increases the probability of relevant retrieval of an image.

All calculations are being performed with the focus that the proposed indexing technique executes at the segment level to retrieve an image. Indexing technique is not only based on its context value but also based on its content value. Various design metrics for RST based IR systems are available to measure the retrieval relevancy and to improve performance of the information retrieval systems. There is also validation of the metrics by taking various case studies [19] but these designs and validations are only for text data.

Image ID	Segment ID	PPI	Color Depth	Segment Phrase
01	Seg3	16	(92, 64, 51)	"Branch of tree"
02	Seg6	10	(0, 0, 0)	"Bridge"
03	Seg5	24	(255,182,193)	"Flowers"
04	Seg5	20	(0, 0, 0)	"Trees"
05	Seg6	36	(222, 184, 135)	"Houses"
06	Seg5	38	(0, 250, 154)	"Sea"
07	Seg5	28	(34, 139, 34)	"Mountains"
08	Seg5	22	(50, 153, 204)	"River"
09	Seg4	24	(184, 134, 11)	"Lions"
10	Seg5	22	(120, 120, 120)	"Rocks"
11	Seg5	24	(153, 204, 50)	"Flowers"
12	Seg5	18	(0, 34, 102)	"Sea"
13	Seg6	14	(250, 0, 0)	"Boat"
14	Seg1	16	(222, 184, 135)	"Baby birds"
15	Seg3	18	(34, 139, 34)	"Bushes"
16	Seg5	22	(173, 216, 230)	"River"
17	Seg6	34	(92, 64, 51)	"Old walls"
18	Seg6	32	(255, 255, 255)	"Drums"
19	Seg6	34	(238, 238, 224)	"Floor"
20	Seg1	38	(255, 255, 240)	"Cricketers"

Table 5. Second trivial indexer  $(S_{11})$ .

## 5. Conclusions and Future Work

In this research paper only image objects are considered as multimedia objects. Future work can be considered for other multimedia objects. In this proposed indexing technique, human interaction is required at different levels, for the optimal solution. In future, this interaction can be finished or minimized by developing computerized automatic indexing technique. In future research work, implemented RST based indexing for image object can be produced based on proposed indexing technique and can be compared with the existing multimedia indexing techniques.

### References

- [1] Al-Hamami A. and Al-Rashdan H., "Improving the Effectiveness of the Color Coherence Vector," *The International Arab Journal of Information Technology*, vol. 7, no. 3, pp. 324-332, 2010.
- [2] Anderson S., Mike P., Mick E., Tony A., Andrew W., and Malcolm P., "Digital Images Archiving Study," *Technical Report*, Arts and Humanities Data Service, 2006.
- [3] Bateman J., Thomas K., Jörg K., and Klaus R., "Towards Constructive Text, Diagram, and Layout Generation for Information Presentation,"

Computational Linguistics, vol. 27, no. 3, pp. 409-449, 2001.

- [4] Cesare R., Oliviero S., Zancanaro M., Kruppa M., and Krüger A., "The Museum Visit: Generating Seamless Personalized Presentations on Multiple Devices," in Proceedings of the 9<sup>th</sup> International Conference on IUI, Intelligent user Interfaces, Portugal, pp. 316-318, 2004.
- [5] Georg G., Hernault H., Cavazza M., Prendinger H., and Ishizuka M., "From Rhetorical Structures to Document Structure, Shallow Pragmatic Analysis for Document Engineering," in Proceedings of the 9<sup>th</sup> ACM Symposium on Document Engineering, Germany, pp. 185-192, 2009.
- [6] Iraky K., Zakareya F., and Abdelfatah F., "Arabic Discourse Segmentation Based on Rhetorical Methods," *International Journal of Electric & Computer Sciences IJECS-IJENS*, vol. 11, no. 01, pp. 10-16, 2011.
- [7] Iria C., Juan M., and Torres M., "Automatic Discourse Segmentation: Review and Perspectives," in Proceedings of International Workshop on African Human Language Technologies, pp.17-20, 2010.
- [8] Jorma L., Markus K., Sami L., and Erkki O., "Picsom, Content-Based Image Retrieval with Self-Organizing Maps," *Pattern Recognition Letters*, vol. 2, no. 1, pp. 1199-1207, 2000.
- [9] Kazuhiko K., Seiji K., and Shin-chi Y., "Accurate Root Length Measurement by Image Analysis," *Plant and Soil, Kluwer Academic Publishers*, vol. 216, no. 1-2, pp. 117-127, 1999.
- [10] Khaldoon K. and Shoaib M., "A Paradigm for Rhetorical Structure Theory (RST) Based Relations to Accommodate Multimedia Objects," *IADIS International Conference e-society*, Spain, pp. 297-299, 2009.
- [11] Manjunath B., Chellappa R., and Von-der C., "A Feature Based Approach to Face Recognition," *in Proceedings of CVPR, Computer Vision and Pattern Recognition*, USA, pp. 373-378, 1992.
- [12] Martin M. and Victor F., "A Presentation Model for Multimedia Summaries of Behavior Proceeding," in Proceedings of the 13<sup>th</sup> International Conference on IUI, Intelligent user Interfaces ACM, Spain, pp. 369-372, 2008.
- [13] Maite T. and William C., "Rhetorical Structure Theory: Looking Back and Moving Ahead," *Discourse Studies*, vol. 8, no. 3, pp. 423-459, 2006.
- [14] Maite T. and Mann W., "Applications of Rhetorical Structure Theory," *Discourse Studies*, vol. 8, vol. 4, pp. 567-588, 2006.
- [15] Martin F., Martin D., Fowlkes C., Tal D., and Malik J., "A Database of Human Segmented Natural Images and its Application to Evaluating Segmentation Algorithms and Measuring

Ecological Statistics," in Proceedings of the  $\delta^{th}$ IEEE International Conference. Computer Vision, Canada, vol. 2, pp. 416-423, 2001.

- [16] Michael S., Nicu S., Chabane D., and Ramesh J., "Content-Based Multimedia Information Retrieval: State of the Art and Challenges," *in* ACM Transactions on Multimedia Computing, Communications, and Applications, vol. 2, no. 1, pp. 1-19, 2006.
- [17] Nakasone A., Prendinger H., and Ishizuka M. "Web Presentation System using RST Events," in Proceedings of the 5<sup>th</sup> International Conference on Autonomous Agents and Multi-Agent Systems, Japan, pp. 995-997, 2006.
- [18] Shoaib M. and Abad A., "A New Indexing Technique for Information Retrieval Systems Using RST," *Journal of Computer Science*, vol. 2, no. 3, pp. 224-228, 2006.
- [19] Shoaib M., Arshad S., Jabeen S., and Tariq M., "Software Quality Metrics for Rhetorical Structure Theory Based Information Retrieval Systems," *Pakistan Journal of Science*, vol. 62, no. 3, pp. 177-184, 2010.
- [20] Sumathi T., Lakshmi C., and Hemalatha M., "An Overview of Automated Image Annotation Approaches," *International Journal of Research and Reviews in Information Sciences*, vol. 1, no. 1, pp. 1-5, 2011.
- [21] William C. and Sandra A., "Rhetorical Structure Theory: A Framework for the Analysis of the Text," *IPM Papers in Pragmatics*, vol. 1, no.1, pp. 79-105, 1987.
- [22] Yijuan L., Nicu S., Ross H., and Qi T., "Personalization in Multimedia Retrieval: A Survey," *Multimedia Tools and Applications*, *ACM*, vol. 51, no. 1, pp. 247-277, 2011.
- [23] Ying L., Dengsheng Z., Guojun L., Wei-Ying M., "A Survey of Content-Based Image Retrieval with High-Level Semantics," *Pattern Recognition, Elsevier Science*, vol. 40, no. 1, pp. 262-282, 2006.
- [24] Zhang L., Lin F., and Zhang B., "A CBIR Method Based on Color-Spatial Feature," *in Proceedings of the IEEE Region 10 Conference TENCON*, South Korea, pp. 166-169, 1999.



Khaldoon Syed Khurshid is working an assistant professor in Computer Science and Engineering Department at University of Engineering and Technology Lahore, Pakistan. He has completed his MSc in computer science from

International Islamic University Islamabad, Pakistan in 2004. Currently, he is PhD scholar in University of Engineering and Technology, Pakistan. His current research interests is in content based indexing and information retrieval systems.



Muhammad Shoaib is a professor at Computer Science and Engineering Department at the University of Engineering and Technology Lahore, Pakistan. He received his MSc in computer science from Islamia University,

Pakistan. He has completed his PhD from the University of Engineering and Technology, Pakistan in 2006. His Post Doc. is from Florida Atlantic University, USA, in 2009. His current research interests include information retrieval systems, information systems, software engineering and semantic web.