



Evaluation, Enhancement, Development and Implementation of Content Based Image Retrieval Algorithms

**A thesis submitted
for award of the degree of**

**DOCTOR OF PHILOSOPHY
in
Electrical Engineering**

Guided by:

Prof. A. I. Trivedi

Submitted by:

Thakore Darshak Gauravbhai



**ELECTRICAL ENGINEERING DEPARTMENT
FACULTY OF TECHNOLOGY & ENGINEERING
THE MAHARAJA SAYAJIRAO UNIVERSITY OF BARODA
VADODARA – 390 001 GUJARAT, INDIA**

2011



Dedicated To

**My Family Members:
Gauravbhai - Pannaben
Dr. Binda - Marmik - Shaivi**



Acknowledgement

Step-advancements in the journey of life are impossible without God's blessings and Guide's direction. The accomplishment of the mammoth task is due to God's grace & inspiration and my guide Prof. A. I. Trivedi Sir's direction, motivation & constant support.

I am thankful to Principal of my parent institute BVM Engineering College - Dr. F. S. Umrigar Sir and Board of Management of the college for allowing me to take up the task, for utilizing resources & for permitting me to present papers in various conferences.

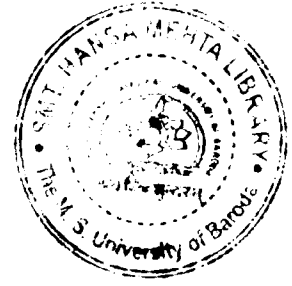
I am also thankful to Prof. A. I. Trivedi Sir, Prof. (Mrs). Kanitkar Madam & Prof. S. K. Shah Sir – Heads of the Department during my period of work and all faculty members at the Faculty of Technology & Engineering, M S University for allowing & extending support to me for the Ph. D. work. The ever-readiness of the supporting staff of the department for providing help is worth taking note of and acknowledged.

I acknowledge Prof. (Dr.) D. P. Vakharia, Mechanical Department, SVNIT, Surat & Prof. S. P. Joshi, Mechanical Department, B V M, V V Nagar, for constantly encouraging me for the work.

I am grateful to my parents, wife and children who sacrificed their needs & expectations to facilitate me working for Ph. D.

Last, but not the least, I am appreciative of all who helped me directly or indirectly during the period of the Ph. D. work.

Certificate



This is to certify that the thesis entitled, "**Evaluation, Enhancement, Development and Implementation of Content Based Image Retrieval Algorithms**" submitted by "**Thakore Darshak Gauravbhai**" in fulfillment of the degree of **DOCTOR OF PHILOSOPHY** in Electrical Engineering Department, Faculty of Technology & Engineering, The M. S. University of Baroda, Vadodara is a bonafide record of investigations carried out by him in the Department of Electrical Engineering, Faculty of Technology & Engineering, M. S. University of Baroda, Vadodara under my guidance and supervision. In my opinion this has attained the standard fulfilling the requirements of the Ph.D. Degree as prescribed in the regulations of the University.

12 May, 2011

Prof. A. I. Trivedi

Department of Electrical Engineering,
Faculty of Technology & Engineering,
The M. S. University of Baroda, Vadodara.

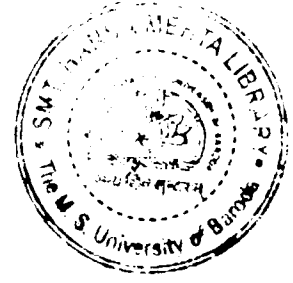
Prof. S. K. Shah,

Head, Department of Electrical Engg.
Faculty of Technology & Engg.,
The M. S. University of Baroda,
Vadodara.

Prof. Dr. K. Babapai

Dean,
Faculty of Technology & Engg.
The M. S. University of Baroda,
Vadodara.

Offg. Dean
Faculty of Tech. & Engg.
M. S. University of Baroda,
Baroda.



Declaration

I, **Thakore Darshak Gauravbhai** hereby declare that the work reported in this thesis entitled, "**Evaluation, Enhancement, Development and Implementation of Content Based Image Retrieval Algorithms**" submitted for the award of the degree of **DOCTOR OF PHILOSOPHY** in Electrical Engineering Department, Faculty of Technology & Engineering, The M. S. University of Baroda, Vadodra is original and was carried out in the Department of Electrical Engineering, Faculty of Technology & Engineering, M. S. University of Baroda, Vadodra. I further declare that this thesis is not substantially the same as one, which has already been submitted in part or in full for the award of any degree or academic qualification of this University or any other Institution or examining body in India or abroad.

Thakore

12th May, 2011

Thakore Darshak Gauravbhai



Abstract

Content Based Image Retrieval (CBIR) – a challenging need of today, aims at retrieving visually similar images from abundantly available / accessible images for a given query image.

The thesis encompasses wide scope covering development of novel algorithms for image segmentation, feature extraction & representation and image retrieval. A challenging task of development, implementation and integration of various novel algorithms to result into GUI based, selectable multimodal processing of single, selectable query image for retrieval of similar images has been achieved successfully. These novel algorithms include:

- Edges and prominent boundaries detection
- Foreground separation
- Image retrieval based on
 - Color codes feature of entire image
 - Foreground color codes feature
 - Foreground shape correlation
 - Combination of foreground color codes feature and shape correlation
 - Extracted face region from images containing complex background for retrieval similar face images

Proposed prominent boundaries detection and foreground separation algorithms emphasize on 'proper' image segmentation, addressing a fundamental problem of computer vision for meeting a prime requirement for a development of a generic CBIR. Prominent boundaries, prominence measure, watershed algorithm with various levels of Haar wavelet decompositions are effectively incorporated together for proper segmentation and feature extraction by enforcing reliable processing of low level cues for avoiding breaks as well as under segmentation by utilizing continuity preserving, well localized visually prominent boundaries for foreground – background separation. The problem of over segmentation is addressed by compositely considering proximity influence and watershed algorithm.

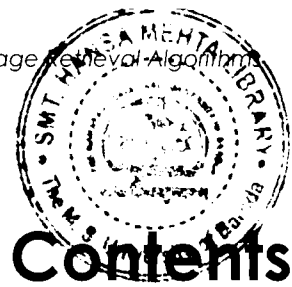
The edge detection method based on candidate boundaries and proximity influence of all four channels detects well localized and delineated perceptually significant edges. The edge detection response of proposed method outperforms edge responses of ACD Photo Editor, MS Photo Editor and Adobe Photoshop. Quantitative analysis & comparisons of edge responses show better performance measure (F-measure) for the proposed method.

The prominent boundaries detection results, segmentation results and foreground extraction results have been qualitatively compared with (i) Human segmented images of standard database BSD3 [Fowlkes, on line] [Martin, 2001] and (ii) with results of JSEG [Deng, on line] [Deng, 2001] for the effect of texture & illumination variations on segmentation and suitability of the algorithms for foreground – background separation. Quantitative analysis yielding high performance measures for extracted foreground with respect to ground truth foreground has been presented to endorse effectiveness of the method.

Proposed image retrieval approaches follow two streams of techniques for achieving a theme - *"Relaxed feature description for better Recall and simultaneous emphasizing of reliable processing of cues leading to precise feature extraction for better Precision"*. The first stream of technique for image comparison is based on color distribution wherein RGB colors are represented by proposed innovative & computationally efficient broad color descriptors called color codes. The technique is suitable for finding near similar images intended for achieving higher Recall. The second technique emphasizes reliable processing of low level cues for precise and well localized prominent boundaries detection, eventually leading to foreground extraction. The extracted foreground is compared on basis of shape correlation and foreground color codes for similar image retrieval. Moreover, a composite approach consisting of foreground shape and foreground color codes provides selectable proportion of weights in composite similarity measures to enable users to match the need based on category & perception for a query image. The exclusion of background and corresponding features supports object based search for image retrieval. The foreground detection based face extraction method and similar face image retrieval from the image containing complex background has been presented as an application specific CBIR, illustrating effectiveness of various proposed algorithms.

Proposed user selectable image retrieval algorithms facilitate user to map needs & choice for bridging semantic gap, to meet subjectivity of perception & challenges offered by image diversities. The query response analysis and performance measures – Precision, Recall & F measures along with P – R curves carried out for various algorithms show effectiveness of the algorithms, indicating higher Precision at higher Recall.

The developed algorithms have been tested on diversified images of various standard image databases, indicative of a step forward on the road-map of continuous & endless technical evolution towards a perfect & versatile CBIR ...



<i>List of Algorithms</i>	<i>XIII</i>
<i>List of Figures</i>	<i>XIV</i>
<i>List of Tables</i>	<i>XIX</i>
<i>List of Abbreviations</i>	<i>XXI</i>
1. Introduction.....	1
1.1 Content Based Image Retrieval (CBIR).....	1
1.2 Application Areas.....	3
1.3 Need & Motivation.....	4
1.4 CBIR System Modules.....	4
1.5 Challenges.....	5
1.6 Image Features.....	6
1.7 Various Approaches	7
1.8 CBIR Systems.....	7
1.9 Scope of the Work	8
1.10 Overview of the Work & Organization of the Thesis	8
1.11 Image Databases.....	12
1.12 Specifications of the Developed CBIR System	13
1.13 Concluding Remark	14
2. CBIR Algorithms & CBIR Systems-Review & Analysis	15
2.1 Introduction.....	15
2.2 Research & Publication Trend Analysis.....	15
2.3 Eminent Contributor's Point of View on CBIR	17
2.4 Query Response Analysis of Some of the Existing CBIR Systems.....	18

2.5	Overview of CBIR Surveys	19
2.6	CBIR and Related Techniques	20
2.7	CBIR Systems.....	28
2.8	Our Observations.....	29
2.9	Our Approaches.....	31
2.10	Discussion.....	32
2.11	Concluding Remark	33
3.	Technical Background.....	34
3.1	Introduction.....	34
3.2	Signal Analysis	34
3.2.1	Types of Mother Wavelets	35
3.2.2	Example - Mother Wavelets & Scaling Functions	35
3.2.3	Continuous Wavelet Transform	38
3.2.4	Discrete Wavelet Transform (DWT) & Stationary Wavelet Transform (SWT)	39
3.3	CBIR Performance Measures	42
3.3.1	Precision.....	42
3.3.2	Recall	42
3.3.3	P - R Curves	43
3.3.4	F - measure	44
3.4	Watershed Regions and Issues.....	44
3.5	Concluding Remark	45
4.	Edges & Prominent Boundaries Detection.....	46
4.1	Introduction.....	46
4.1.1	Key Terminologies	46
4.2	Block Diagram – Edges and Prominent Boundaries Detection.....	47
4.3	Edge Detection	48
4.3.1	The Method	49
4.3.2	Step-wise Results of the Method.....	52
4.3.3	Results – Edge Response Comparisons	55
4.3.4	Discussion	61

4.4	Prominent Boundaries Detection.....	62
4.4.1	The Method	64
4.4.2	Results	66
4.4.3	Discussion	74
4.5	Concluding Remark	75
5.	Foreground Objects Detection & Background Separation	76
5.1	Introduction.....	76
5.2	The Method	77
5.3	Results	79
5.3.1	Step-wise Results of the Method.....	79
5.3.2	Qualitative Comparisons.....	82
5.3.3	Qualitative Result-comparisons: Proposed Method, JSEG & Human Segmented Images of BSDb [Martin, 2001] [Fowlkes, on line]	99
5.4	Discussion.....	106
5.5	Concluding Remark	107
6.	Image Retrieval	108
6.1	Introduction.....	108
6.2	Image Features.....	109
6.3	SIMPLicity Image Database [Wang, 2001] [SIMPLicity, on line] - Classes & Characteristics.....	113
6.4	ALOI Image Database [ALOI, on line] [Geusebroek, 2001]	115
6.5	Proposed Techniques.....	115
6.6	Whole Image Color Codes Based CBIR.....	116
6.6.1	Performance Evaluation	117
6.6.1.1	Query Image Class: Bus	117
6.6.1.2	Query Image Response Examples: Class – Bus	121
6.6.1.3	Query Image Class: Horse	123
6.6.1.4	Query Image Response Examples: Class - Horse	126
6.6.1.5	Query Image Class - Flower	130
6.6.1.6	Query Image Response Example: Class – Flower.....	135

6.6.1.7	Query Image Class – Dinosaur	136
6.6.1.8	Query Image Response Example: Class - Dinosaur	139
6.6.1.9	Query Image Response Examples: Other Classes	140
6.6.2	Discussion	146
6.7	Foreground Color Codes Based CBIR	146
6.7.1	Performance Evaluation	147
6.7.1.1	Query Image Class: Bus	147
6.7.1.2	Query Response Example: Class - Bus	151
6.7.1.3	Query Image Class – Flower	153
6.7.1.4	Query Response Example: Class - Flower	157
6.7.2	Discussion	158
6.8	Foreground Shape Correlation Based CBIR	159
6.8.2	Performance Evaluation	159
6.8.2.1	Query Response Example	162
6.8.3	Discussion	164
6.9	Foreground Color Codes & Foreground Shape Based CBIR	164
6.9.1	Performance Evaluation	165
6.9.1.1	Query Response Examples:	167
6.9.2	Discussion	170
6.10	Comparisons - Query Responses of Various Algorithms	170
6.10.1	Example 1 - SIMPLiCity image database [Wang, 2001] [SIMPLiCity, on line]	170
6.10.2	Example 2 - ALOI image database [ALOI, on line] [Geusebroek, 2001]	173
6.10.3	Discussion	178
6.11	Application Specific CBIR - Similar Face Image Retrieval	179
6.11.1	Face Extraction from Images Containing Complex Background	179
6.11.1.1	The Method	181
6.11.1.2	Results	181
6.11.1.3	Discussion	182
6.11.2	Similar Face Image Retrieval	187
6.11.2.1	Performance Evaluation	188
6.11.2.2	Query Response Example	190
6.11.2.3	Discussion	190
6.12	Performance Comparisons with other CBIR Techniques	190
6.13	Concluding Remark	192

7. Conclusions & Future Enhancements	193
7.1 Conclusions.....	193
7.2 Limitations.....	196
7.3 Future Enhancements.....	196
7.4 Concluding Remark	197
References	198
Annexure 1 - Publications.....	207
Annexure 2 - GUI & Description	208
A-2.1 Graphical User Interface.....	208
A-2.2 Component Description.....	209
Annexure 3 – Results: Miscellaneous.....	213
A-3.1 Suitability of Proposed Methods for Character Recognition	213
A-3.2 Foreground Extraction - Tiger Images [Fowlkes, on line] [Martin, 2001].....	213
A-3.3 Query Response Examples–Tiger Images [Fowlkes, on line][Martin, 2001].....	213
A.3.4 Face Extraction - Face Image BSDB [Fowlkes, on line] [Martin, 2001]	216
Annexure 4 – Distinctive Issues.....	217
A-4.1 Search Engines - Google Image Search and vSearch.....	217
A-4.2 Quantitative Analysis & Comparisons of Edge Responses	220
A-4.2.1 Discussion.....	224
A-4.3 Quantitative Analysis of Results of Proposed Method for Foreground Extraction w. r. t. Ground Truth.....	224
A-4.3.1 Discussion.....	230
A-4.4 Color Codes	230
A-4.4.1 Results - Color Code Based Segmentation	230
A-4.5 Processing Time Analysis.....	233
A-4.5.1 Discussion.....	236

List of Algorithms

Algorithm 1. Edge detection and thinning 51

Algorithm 2. Prominent boundaries detection..... 64

Algorithm 3. Foreground objects detection & background separation. 79

Algorithm 4. Generic steps for proposed image retrieval methods..... 116

Algorithm 5. Whole image color codes based image retrieval..... 117

Algorithm 6. Foreground color codes based image retrieval. 147

Algorithm 7. Foreground shape correlation based image retrieval. 159

Algorithm 8. Foreground color codes & foreground shape based image retrieval. 165

Algorithm 9. Face extraction from images containing complex background. 181

List of Figures

Figure 1. Image search results for query – rose black a) Google b) Yahoo	1
Figure 2. Block diagram of a CBIR system.....	2
Figure 3. Detailed block diagram of a CBIR system.....	2
Figure 4. Block diagram – Perspective of query image processing for proposed system.	9
Figure 5. Research Interest Trend Based.....	16
Figure 6. Mother Wavelets and Corresponding Scaling Functions.	36
Figure 7. Shift and Scale Operation on Mother Wavelet [Ha, on line].	38
Figure 8. Block diagram for DWT.	39
Figure 9. Wavelet - Multi-level Decomposition.....	40
Figure 10. Decomposition with Discrete Haar Wavelet and Discrete Stationary Haar Wavelet.....	41
Figure 11. Ideal Precision – Recall Curve	43
Figure 12. Watershed Regions – Issues.....	45
Figure 13. Correctly Labeled Regions without using watershed transform	45
Figure 14. Block Diagram – Edge and Prominent Boundaries Detection.	48
Figure 15. Contour Detection ..	52
Figure 16. Processed Contours – Candidate boundaries..	53
Figure 17. Binary Images: Edges from contours.....	54
Figure 18. Edges.....	55
Figure 19. Edge Response Comparison.....	56
Figure 20. Edge Response Comparison.....	59
Figure 21. Edge Response Comparison.....	60
Figure 22. Prominent & non-prominent boundaries.....	65
Figure 23. Comparison of human segmented image results with detected prominent boundaries.....	67
Figure 24. Prominent boundaries detection at different levels of Stationary Haar decompositions.....	69
Figure 25. Prominent boundaries detection results - Different categorical images.....	70
Figure 26. Prominent boundaries detection results - Different categorical images.....	70
Figure 27. Prominent boundaries detection results - Different categorical images.....	71
Figure 28. Comparison of human segmented image results with detected prominent boundaries.....	72
Figure 29. Prominent boundaries detection results - images of PASCAL challenge 2008.....	73
Figure 30. Prominent boundaries detection results - Different categorical images.....	74
Figure 31. Foreground objects detection & background separation - Step-wise results – 1.	80

Figure 32. Foreground objects detection & background separation - Step-wise results – 2..	81
Figure 33. Watershed transform based segmentation without applying proposed method..	84
Figure 34. Revealed fore ground objects & background, incorporating different levels of wavelet decompositions.	85
Figure 35. Revealed fore ground objects & background for images with typical textures..	86
Figure 36. Revealed foreground objects & background, incorporating stationary Haar wavelet decompositions at level 2.....	88
Figure 37. Revealed foreground objects & background, incorporating stationary Haar wavelet decompositions at level 2.....	90
Figure 38. Result-comparison: Human segmented image with revealed foreground objects & background, incorporating different levels of stationary Haar wavelet decompositions.....	92
Figure 39. Result-comparison: Human segmented image with revealed foreground objects & background, incorporating different levels of stationary Haar wavelet decompositions.....	94
Figure 40. Result comparison: Human segmented images with revealed foreground objects & background, incorporating stationary Haar wavelet decompositions at levels 2.	96
Figure 41. Result comparison: Human segmented images with revealed foreground objects & background.	97
Figure 42. Effect of illumination variations on segmentation / foreground extraction on object with fine details.....	100
Figure 43. Effect of illumination variations on segmentation / foreground extraction on non-textured object.....	101
Figure 44. Comparison of segmentation results... ..	102
Figure 45. Comparison of segmentation results.. ..	104
Figure 46. Comparison of segmentation results.. ..	105
Figure 47. Block diagram - Feature extraction.....	110
Figure 48. Color – code based segmentation.....	113
Figure 49. Color – code based segmentation & regions corresponding to two color-codes.	113
Figure 50. ALOI sample images [ALOI, on line] [Geusebroek, 2001].....	115
Figure 51. P- R curves (whole image color codes). Class - Bus.....	120
Figure 52. Average Precision, Average Recall, Average F – measures verses Similarity cut-offs (Whole image color codes). Class – Bus.	120
Figure 53. Query response of a bus image at similarity cut-off 60.	121
Figure 54. Query response of a bus image at similarity cut-off 40.	122
Figure 55. P- R curves (whole image color codes). Class- Horse.	125
Figure 56. Average Precision, Average Recall, Average F – measures verses Similarity cut-offs (Whole image color codes). Class – Horse.	125

Figure 57. Query response of a horse image at similarity cut-off 25.	127
Figure 58. Query response of another horse image at similarity cut-off 70.	130
Figure 59. P- R curves (whole image color codes). Class- Flower.	133
Figure 60. Average Precision, Average Recall, Average F – measures verses Similarity cut-offs (Whole image color codes). Class – Flower.	134
Figure 61. Query response of a flower image at similarity cut-off 40.....	135
Figure 62. P- R curves (whole image color codes). Class- Dinosaur.	138
Figure 63. Average Precision, Average Recall, Average F – measures verses Similarity cut-offs (Whole image color codes). Class – Dinosaur.....	138
Figure 64. Query response of a dinosaur image at similarity cut-off 40.	139
Figure 65. Query response of a sea-shore image at similarity cut-off 60.	141
Figure 66. Query response of another sea-shore image at similarity cut-off 60.....	141
Figure 67. Query response of a sculpture image at similarity cut-off 60.	142
Figure 68. Query response of an elephant image at similarity cut-off 70.	142
Figure 69. Query response of another elephant image at similarity cut-off 60.	143
Figure 70. Query response of served food image at similarity cut-off 60.	143
Figure 71. Query response of an image of a tribal man with color painted on face.	144
Figure 72. P- R curves (whole image color codes). All queries for the method.....	145
Figure 73. Average Precision, Average Recall, Average F – measures verses Similarity cut-offs (Whole image color codes). All queries for the method.	145
Figure 74. P- R curves (Foreground color codes). Class - Bus.	150
Figure 75. P- R curves (Foreground color codes). Class - Bus.	151
Figure 76. Query response of a bus image at similarity cut-off 25. (FGCC)	152
Figure 77. P- R curves (Foreground color codes). Class- Flower.	156
Figure 78. Avg. Precision, Avg. Recall, Avg. F – measures verses Similarity cut-offs (Foreground color codes). Class – Flower.....	156
Figure 79. Query response of a flower image at similarity cut-off 25. (FGCC)	157
Figure 80. Query response of another flower image at similarity cut-off 25. (FGCC)	158
Figure 81. P- R curves (Foreground shape correlation). Class- Flower.	162
Figure 82. Query response of a flower image at similarity cut-off 60. (FG shape correlation)	163
Figure 83. P- R curves for different proportionate weights of Foreground color codes & foreground shape correlation.	166
Figure 84. Query response of a flower image at similarity cut-off 60 with FG CC weight 20. (FGCC & FG shape correlation)	168
Figure 85. Query response of same flower image at similarity cut-off 60 with FG CC weight 10 (FGCC & FG shape correlation).....	168

Figure 86. Query response of same flower image at similarity cut-off 50 with FG CC weight 20 (FGCC & FG shape correlation).....	168
Figure 87. Query response of same flower image at similarity cut-off 50 with FG CC weight 10 (FGCC & FG shape correlation).....	169
Figure 88. Query response of a flower image at similarity cut-off 60. (Whole image color codes).	171
Figure 89. Query response of a flower image at similarity cut-off 60. (Foreground color codes).	171
Figure 90. Query response of a flower image at similarity cut-off 60. (Foreground shape correlation)..	172
Figure 91. Query response of same flower image at similarity cut-off 60 with 30% weight of FGCC. (Foreground color codes & foreground shape correlation).....	173
Figure 92. Query response of an ALOI image at similarity cut-off 70. (Whole image color codes).	174
Figure 93. Query response of same ALOI image at similarity cut-off 70. (Foreground color codes).....	176
Figure 94. Query response of same ALOI image at similarity cut-off 70. (Foreground shape corr.).	177
Figure 95. Query response of same ALOI image at similarity cut-off 70 with 30% weight of FGCC (Foreground color codes & foreground shape correlation).....	178
Figure 96. Various steps of face extraction. Left - stationary Haar wavelet decomposition at level 1...	183
Figure 97. Face extraction in high resolution image reduced to 1/8th of the original size.....	185
Figure 98. Example of unsuccessful face extraction..	185
Figure 99. Face extractions of images with complex background and non-uniform illuminations..	185
Figure 100. P – R curves. Similar face image retrieval.	189
Figure 101. Query Response – similar face-image retrieval.	190
Figure 102. GUI of the CBIR.....	208
Figure 103. Image Selection by Browsing	209
Figure 104. Dialogue box for unselected image.....	211
Figure 105. Dialogue box for unselected check box for image type	211
Figure 106. Dialogue box for confirming all image processing for all attributes.....	212
Figure 107. Dialogue box for wrongly selected target folder for image retrieval	212
Figure 108. Scanned alphabet, Mapped edges, Foreground, Background and Watershed pixels.	213
Figure 109. Tiger images of BSDb [Fowlkes, on line] [Martin, 2001] and extracted foreground.	214
Figure 110. Queries & responses of Tiger images BSDb [Fowlkes, on line] [Martin, 2001].	215
Figure 111. Image of BSDb [Fowlkes, on line] [Martin, 2001] and extracted face	216
Figure 112. Inferred block diagram for Google like search engines	217
Figure 113. Query response for Google – could not tag & retrieving irrelevant images	218
Figure 114. Query response for Google - tagging wrongly & retrieving many irrelevant images.....	219
Figure 115. Query response for vSearch - retrieving many irrelevant (?) images.....	219
Figure 116. Edge Response Comparison & quantitative analysis – example 1	221
Figure 117. Edge response comparison & quantitative analysis – example 2.....	222

Figure 118. Quantitative comparison of edge responses: ACD Photo editor & proposed method 223

Figure 119. Quantitative comparison of edge responses: Adobe Photoshop & proposed method..... 223

Figure 120. Quantitative comparison of edge responses: MS Photo editor & proposed method..... 224

Figure 121. Qualitative & Quantitative Performance Comparisons for foreground extraction..... 226

Figure 122. Qualitative & Quantitative Performance Comparisons for foreground extraction with
respect to different levels of Haar SWT.. 227

Figure 123. Quantitative analysis w.r.t. ground truth for foreground extraction on BSDb Images..... 228

Figure 124. Qualitative & Quantitative Performance Comparisons for foreground extraction on
images with illumination variations..... 228

Figure 125. Qualitative & Quantitative Performance Comparisons for foreground extraction..... 229

Figure 126. Precision – Recall analysis w.r.t. ground truth for foreground extraction on BSDb
images [Fowlkes, on line] 229

Figure 127. Results: Color codes based segmentation. 232

List of Tables

Table 1. Types of Wavelets.	35
Table 2. Test Images & Their Performance Challenging Salient Characteristics.....	56
Table 3. Categorical Representative Test Images & Characteristics.....	66
Table 4. Categorical Representative Test Images & Their Performance Challenging Salient Characteristics.....	82
Table 5. Extracted features.	111
Table 6. SIMPLicity Image Database [Wang, 2001] [SIMPLicity, on line] - Classes and Characteristics.....	114
Table 7. Precision, Recall & F –measure at different similarity cut-offs. (Whole image color codes) Class - Bus.	118
Table 8. Average Recall, Average Precision & Average F -measure. (Whole image color codes) Class – Bus.....	120
Table 9. Precision, Recall & F –measure at different similarity cut-offs. (Whole image color codes) Class - Horse.	123
Table 10. Average Recall, Average Precision & Average F - measure. (Whole image color codes) Class – Horse.	124
Table 11. Precision, Recall & F –measure at different similarity cut-offs. (Whole image color codes) Class - Flower.	131
Table 12. Average Recall, Average Precision & Average F - measure. (Whole image color codes) Class – Flower.....	133
Table 13. Precision, Recall & F –measure at different similarity cut-offs. (Whole image color codes) Class – Dinosaur.	136
Table 14. Average Recall, Average Precision & Average F - measure. (Whole image color codes) Class – Dinosaur.	137
Table 15. Average Recall, Average Precision & Average F - measure. (Whole image color codes) All queries for the method.....	144
Table 16. Precision, Recall & F –measure at different similarity cut-offs. (Foreground color codes) Class - Bus.	148
Table 17. Average Recall, Average Precision & Average F -measure. (Foreground color codes) Class – Bus.....	150

Table 18. Precision, Recall & F –measure at different similarity cut-offs. (Foreground color codes)
Class - Flower..... 154

Table 19. Average Recall, Average Precision & Average F - measure. (Foreground color codes)
Class – Flower..... 155

Table 20. Precision, Recall & F–measure at different similarity cut-offs.(Foreground shape correlation). Class - Flower..... 160

Table 21. Average Recall, Average Precision & Average F - measure. (Foreground Shape correlation). Class – Flower. 161

Table 22. Precision, Recall & F –measure for different proportionate weights at different similarity cut-offs. (Foreground Color codes & foreground shape correlation)..... 165

Table 23. Precision, Recall & F –measure for two different proportionate weights at different similarity cut-offs. (Foreground Color codes & foreground shape correlation). 167

Table 24. Performance evaluation of face extraction method for various images [Caltech, on line] [Fei-Fei, 2004]. 187

Table 25. Precision, Recall & F – measure at different similarity cut-offs. Similar face-image retrieval. 188

Table 26. Average Recall, average Precision & average F – measure at different similarity cut-off. (Similar face image retrieval)..... 189

Table 27. Formulation of Color Codes231

Table 28. Processing time analysis for SIMPLicity images.....234

Table 29. Processing time analysis for BSDb images.....235

List of Abbreviations

CBIR - Content Based Image Retrieval

CC - Color codes

FG - Foreground

FGCC - Foreground color codes

p - Precision, image retrieval

P_e - Precision, edge response

P_{fg} - Precision, foreground extraction

P - R curves - Precision - Recall curves

r - Recall, image retrieval

R_e - Recall, edge response

R_{fg} - Recall, foreground extraction

rr - Retrieved relevant images

total - Total retrieved images

Total - Total relevant images in the database



1. Introduction

1.1 Content Based Image Retrieval (CBIR)

The content based image retrieval (CBIR) aims at retrieving visually similar images from image database for a given query image (or sketch). Retrieval of *required-query-similar* images from abundantly available / accessible digital images is a challenging need of today. The image retrieval techniques based on visual image content has been in-focus for more than a decade. Many search-engines, including state of the art web-search-engines retrieve similar images by searching and matching textual metadata associated with digital images. For quicker response time, association of metadata is carried out as an off-line process known as image-annotation. The image search results, appearing on the first page for fired text query *rose black*, are shown in Figure 1 for leading web search engines – Google & yahoo.

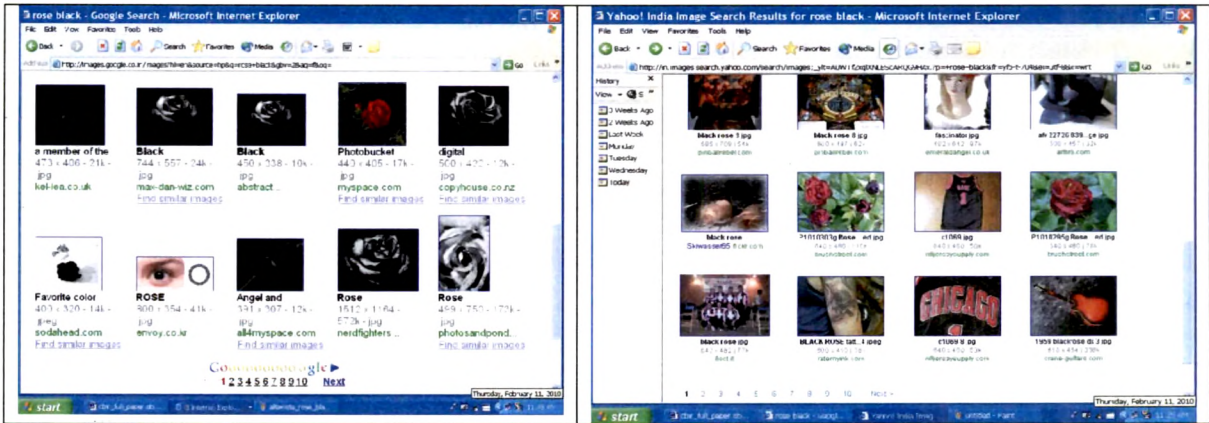


Figure 1. Image search results for query – rose black. a) Google b) Yahoo

As can be seen, many resultant images of Figure 1 lack semantic matching with the meaning of the text-query, showing vast scope of research leading to improvements in the state-of-art-techniques for image retrieval.

The need for image retrieval evolved two solutions – image annotation (or image tagging) and content based image retrieval. The annotation of images is in the form of textual description, carried out either automatically or manually e.g. Google image labeler (<http://images.google.com/imagelabeler/>). The automatic image annotation / tagging analyses the image content for producing and associating textual description with images under considerations. The textual query is then matched with annotations for image retrieval. The content based image retrieval techniques aim to respond to a query image (or sketch) with query-similar resultant images obtained from the image database, as shown in Figure 2.

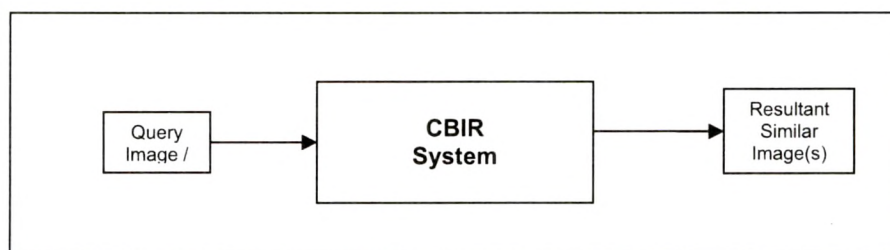


Figure 2. Block diagram of a CBIR system.

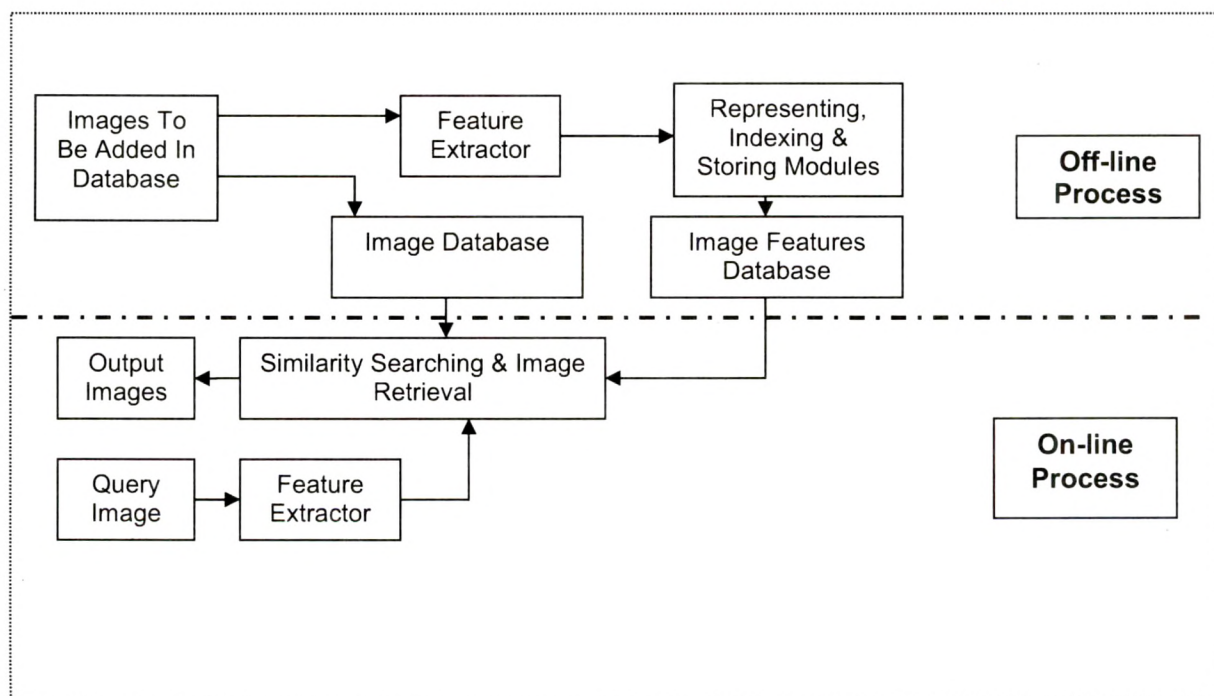


Figure 3. Detailed block diagram of a CBIR system.

The detailed block diagram of a CBIR system available in the literature is shown in the Figure 3. The system consists of two databases – image database and feature

database. The image feature database is a collection of the image features on which image search is to be carried out. While adding a new image in the image database, image features are extracted and stored in the feature database. The features are represented in a convenient format before storing them appropriately for faster search. Features of the query image are extracted and compared with the features of the images, available in feature database. The similarity comparison and search are carried out with the image feature database for finding similar-featured-images. The corresponding images are retrieved and displayed as a result on the basis of similarity measures.

1.2 Application Areas

Typical areas involving content based image retrieval [Gudivada, 1995] are

- Art galleries and museum management
- Architectural and engineering design
- Interior design
- Remote sensing and management of earth resources
- Geographic information systems
- Scientific database management
- Weather forecasting
- Retailing
- Fabric and fashion design
- Trademark and copyright database management
- Law enforcement and criminal investigation
- Picture archiving and communication systems.

Other applications reported in the literature are

- Web image searching
- Education and training
- Home applications like digital photograph cataloging and retrieval
- Medical image database maintenance and medical diagnosis
- Military applications
- Pornography detection & elimination

1.3 Need & Motivation

Advances in processor speed & digital sensor technology, expansion & availability of internet have tremendously increased volume & accessibility of digital images. This has caused a need to have a system that returns images based on the given query considering visual similarity among them to facilitate image searching / cataloging. The conventional image search techniques involves search on metadata comprises of textual annotation associated with images, having severe drawbacks and limitations as illustrated in Figure 1. The CBIR system aims to overcome these limitations. The vast application areas ([Section 1.2](#)) and lack of availability of general purpose CBIR system imply huge research potential on the subject.

1.4 CBIR System Modules

Various modules of a CBIR system shown in Figure 3 can be described as under.

- **Input / output interface:**

This module provides Graphical User Interface to user to give or select input query either in a form of an image within or outside of the image database. It may also provide GUI for giving sketch as an input query. The output interface displays images based on the similarity measures. It may also provide a mechanism to give positive or negative relevance feed back to refine search results.

- **Image database and its interface:**

The image database is an unorganized collection of images. The corresponding interface provides accessibility to image database.

- **Image features database and its interface:**

The image feature database is an organized collection of image features of corresponding images of image-database. The corresponding interface provides accessibility to image feature database.

- **Feature extractor:**

It is the most important module of any CBIR system. The relevant-image-retrieval performance of the CBIR system is directly proportionate to the performance of the feature extractor module. The module extracts required image features listed in [Section 1.6](#) by meeting challenging issues specified in [Section 1.5](#). These image features are compared to determine similarity.

- **Representing, indexing & storing modules:**

The extracted features are represented appropriately to facilitate storing and indexing.

- **Similarity searching & image retrieval:**

The extracted features of the query image are compared with the stored image features of the feature database by applying similarity criteria for retrieving content-similar-images.

1.5 Challenges

Following issues make development of versatile techniques for image feature extraction and hence retrieval difficult and challenging.

- **Semantic gap:**

It is a most crucial factor affecting the relevant-image-retrieval performance of the image retrieval. The semantic gap - as defined in [Smeulders, 2000], is the lack of coincidence between the information that one can extract from visual data and the interpretation that the same data have for a user in a given situation. It is also described in the literature as a gap between human perception for the image content description and its feature representations.

- **Subjectivity:**

The subjectivity of human being for the content analysis and description, which is characterized by human psychology, emotions and imaginations, is a second most crucial factor affecting the relevant-image-retrieval performance of the system.

- **Inter tuning of various phases:**

The feature extractor module generally consists of a series of operations / phases whose proper tuning is important for better over-all performance of the system. Inter tuning of phases plays very important role for example in CBIR systems incorporating relevance feedback or hierarchical frame work for feature extraction & representation.

- **Variety of image categories & characteristics:**

Different image categories & varieties of image characteristics add to the difficulty levels for development of versatile image feature extraction algorithms. Few of them are summarized here.

- Image resolutions and resizing of images
- Image categories
- Intra-image illumination variations
- Non- homogeneity of intra-region and inter-region textures
- Multiple and occluded objects
- Affine transforms of objects
- **Parameter tuning and threshold value selection:**

A large variety in the image characteristics and image categories require parameter tuning and threshold value selection for meeting required scale of feature extraction.
- **Time performance related issues:**

The optimization of feature processing time and query response time may become crucial for a large image database. The issue of feature dimensionality reduction is equally important for a large image database.
- **Application domain specific issues:**

The selection of algorithms, parameters, scale of segmentation etc. are many a times application specific, e.g. the segmentation algorithm used for natural images may not be suitable for X-ray or histopathological images.

1.6 Image Features

Various image features and variants of there of; found in the literature for content description of the images are as under.

- Histograms
 - Local , global and cumulative
- Colors, Color layouts and color distributions
- Edges
- Contours
- Boundaries & regions
- Textures
- Shapes

1.7 Various Approaches

The CBIR algorithm issues are addressed by various approaches in the literature, viz.

- Iterative relevance feed back from user
- Fuzzy, evolutionary and neural network
- Hierarchical approaches
- Focusing on improvements on processing of low level cues so as to precisely extract features
- Semantic domain based image retrieval systems, comparing *meaningful* concepts

1.8 CBIR Systems

Various CBIR systems reported in [Thakore, 2010, 1] are summarized below:

- **QBIC:** Query By Image Content system [QBIC, on line] [Flickner, 1995], developed by IBM, a pioneer commercial product
- **Ultimedia Manager Product:** developed by IBM [Barber, 1994], based on QBIC technology.
- **VisualSEEk:** Developed at Columbia University [Smith, 1996] [VisualSEEk, on line].
- **Photobook:** Developed at Media Laboratory, Massachusetts Institute of Technology – MIT [Photobook, on line], incorporating a unique feature of interactive learning agent, named *FourEyes* for selecting & combining feature-based models
- **MARS:** Multimedia Analysis and Retrieval Systems [MARS, on line]
- **FIRE:** Flexible Image Retrieval Engine [Fire, on line]
- **PicSOM:** (Picture & Self-organizing Map) [Laaksonen, 1999] implemented using tree structured SOM
- **NeTra** [Ma, 1997]

The elaborative survey on CBIR system can be found in [Veltkamp, on line].

1.9 Scope of the Work

The work carried out includes:

1. Literature survey – CBIR algorithms and CBIR systems [Thakore, 2010, 1].
2. Development and implementation of a novel algorithm for edge detection and edge thinning & qualitative comparison of results with those of Adobe Photoshop, ACD photo editor, MS photo editor for color images [Thakore, 2010, 2].
3. Development and implementation of a novel algorithm for prominent boundaries detection and qualitative comparison of results with those of standard databases of color images [Thakore, 2010, 3].
4. Development and implementation of a novel algorithm for foreground object revealing by separating background for prominent boundaries detected color images [Thakore, 2010, 4].
5. Development and implementation of novel algorithms for image retrieval, based on:
 - color code attributes of whole image
 - color code attributes of separated foreground
 - foreground shape correlation coefficients
 - combination of foreground color codes and shape correlation
- 5.1 Development and implementation of prominent boundaries & foreground separation based algorithm for extracting human-face and retrieving similar face images containing complex backgrounds.
- 5.2 Performance evaluation and analysis of results of above image retrieval algorithms at various selectable similarity-cut-offs.
6. Integration of various modules and development of Graphical User Interface for the CBIR system comprising of above algorithms.

1.10 Overview of the Work & Organization of the Thesis

CBIR being the most demanding and challenging need of the recent years, the work emphasizes on development and implementation of algorithms for achieving content based image retrieval. A complete GUI based CBIR system having selectable multimodal image retrieval with selectable cut-offs of (dis)similarity for selectable input-query image have been developed, implementing novel algorithms. The perspective

of query image processing for feature extraction & image retrieval given as the block diagram in Figure 4 gives the overview of the work and the relationship / dependency of various phases of the developed CBIR system / algorithms.

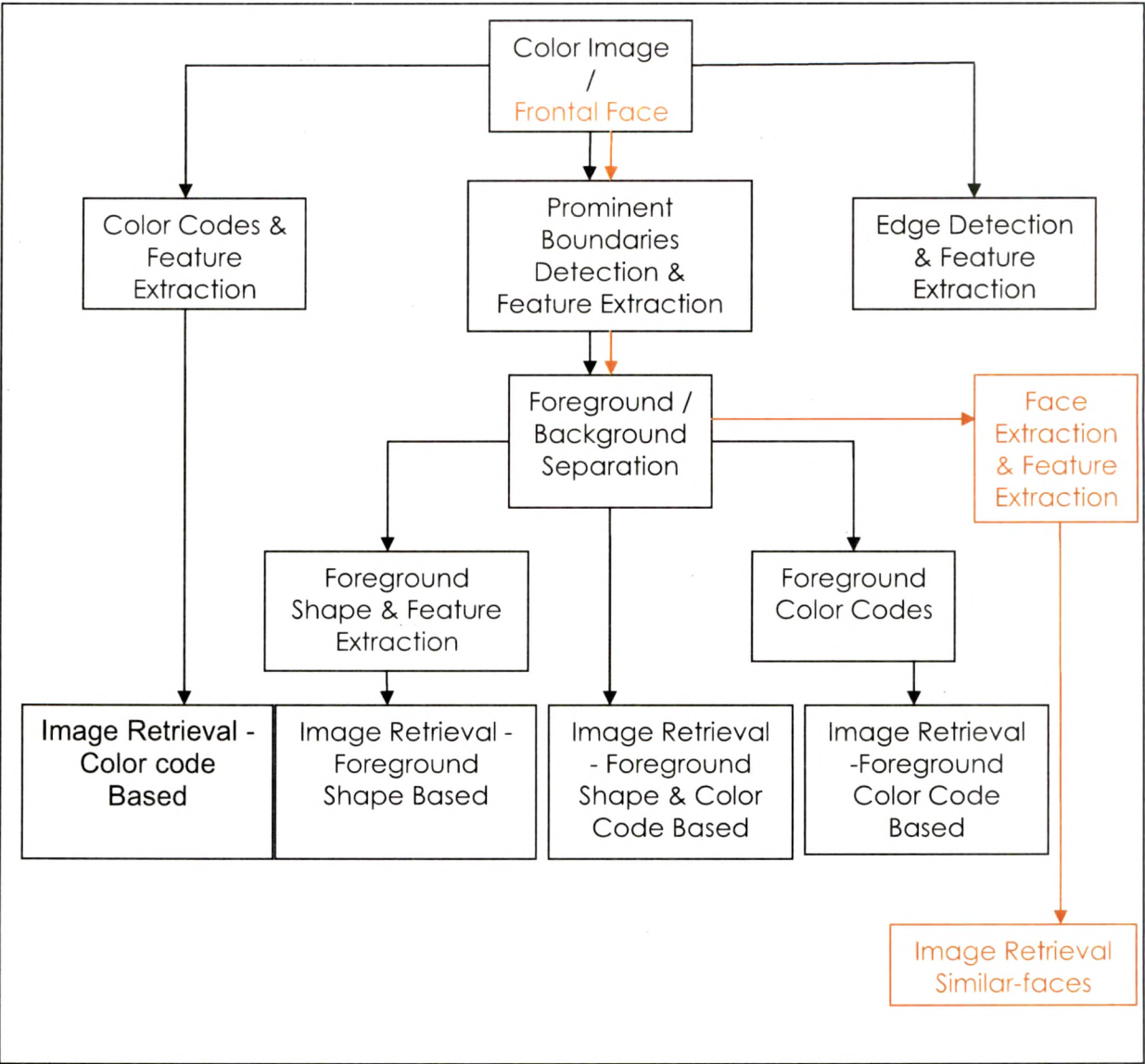


Figure 4. Block diagram – Perspective of query image processing for proposed system.

Our approach is based on two streams of features – prominent boundaries based features and Color code based features. These features and combination of them have been used for image retrieval. The color codes are broadest descriptors of colors whereas prominent boundaries are well localized boundaries detected due to reliable processing of low level cues. One of the most fundamental requirements for better performance of CBIR is the 'proper' image segmentation needed for precise and

reliable image-feature extraction. A novel approach based on detection of prominent boundaries for image segmentation & image feature extraction has been developed. The versatility and suitability of prominent boundary based image segmentation has been proved by testing various categories of images of various standard databases listed in [Section 1.1.1](#). The qualitative comparison of results with human segmented images available at standard image database BSDb [Fowlkes, on line] [Martin, 2001] shows effectiveness of the method.

The second technique for image feature extraction is based on *color codes* - a technique to map a set of colors to a specific color code leading to image segmentation based on broad color descriptors. Though the method may sometimes under-segment the image, the computationally inexpensiveness and broad color description of regions are two great advantages of the method.

The prominent boundaries and watershed algorithm have been utilized for region feature extraction and foreground / background separation. The incorporation of Stationary Haar wavelet decomposition at various levels makes the method suitable for hierarchical approach. The results of the proposed method are also qualitatively compared with those of JSEG [Deng, on line] [Deng, 2001] for the effect of texture and illumination variations and suitability of the algorithms for foreground – background separation.

The developed algorithms extract image features like contours, edges, boundaries & thin boundaries, regions & region attributes, color & color distribution, derived features of foreground & background regions, shape of foreground region, color & color distribution of foreground object. The edge-features are not incorporated for image retrieval purpose. It may be utilized for further reduction in number of over-segmented regions – *a suggested future enhancement of the work*. The edge detection results are compared qualitatively in **[Thakore, 2010, 2]** with edge detection performance of leading DTP and image processing software packages – Adobe Photoshop, ACD photo-editor and MS photo-editor. The results of the proposed method outperform others for i) detection of significant perceptual edges ii) elimination of insignificant edges corresponding background and foreground textures iii) better preservation of continuity.

Image retrieval methods based on (i) color codes of entire image (ii) Foreground color codes (iii) Foreground shape correlation (iv) Combination of foreground color codes and shape correlation have been developed and implemented. The image query responses for various images at different similarity cut-off along with precision and recall for various images of standard database SIMPLcity [SIMPLcity, on line] and ALOI [ALOI, on line] [Geusebroek, 2001] have been presented for aforesaid retrieval techniques.

The prominent boundaries and foreground detection based method for extracting face and face region attributes from images containing complex background and illumination variations have been developed. The application specific CBIR for retrieving similar-face-images has been implemented. The method measures similarity of normalized face regions based on correlation coefficients. The effectiveness of prominent boundaries and foreground detection techniques for face extraction has been shown for 115 images of standard database Caltech [Caltech, on line] [Fei-Fei, 2004].

The algorithms are tested on vast set of images inclusive of standard image databases listed in Section 1.11, consisting of various categories of images. The implementation has been carried out on Intel® dual core T 2050, 1.6 GHz processor with 1.5 GB of RAM using Matlab R14.

The thesis is organized as follows. Chapter 2 includes review and evaluation of various CBIR algorithms and CBIR systems. The relevant technical background is covered in the chapter 3. The chapter 4 deals with the developed novel techniques for edges & prominent boundaries detection. It also covers analysis and qualitative comparison of the results of prominent boundaries with human segmented images of standard database BSDb [Fowlkes, on line] [Martin, 2001]. The edge detection results are qualitatively compared with edge detection response of Adobe Photoshop, ACD photo-editor and MS photo-editor in the same chapter. The next chapter 5 covers prominent boundaries based novel technique for foreground object detection and background separation, results and qualitative comparison of them. It also includes qualitative comparison of results of the proposed method of segmentation / foreground detection with that of segmentation algorithm JSEG [Deng, on line] [Deng, 2001]. The chapter 6 includes the details of the image features, proposed novel method for image retrieval based on (i) color codes of entire image (ii) Foreground color codes (iii)

Foreground shape correlation (iv) Combination of foreground color codes and shape correlation. It also covers results and performance evaluation and analysis with precision-recall & precision recall curves. The method and results of face extraction and similar-face-image retrieval are also covered in the chapter. And finally, conclusions & future enhancements are covered in the last chapter 7.

The Annexure 1 contains the list of the publications made so far. The Annexure 2 describes various components of Graphical User Interface (GUI) for the developed CBIR application. The Annexure 3 contains typical miscellaneous results.

1.11 Image Databases

Proposed novel algorithms have been tested on various images including following standard image databases.

- **BSDb** [Fowlkes, on line] [Martin, 2001] - The Berkeley Segmentation Dataset and Benchmark (BSDb).

Available: <http://www.cs.berkeley.edu/projects/cs/vision/grouping/segbench/>

The standard data set consists of variety and vast range of images, segmented images and human segmented images of medium size and resolutions.

- **SIMPLcity** [SIMPLcity, on line] [Wang, 2001]- The SIMPLcity web site:

Available: http://wang14.ist.psu.edu/cgi-bin/zwang/regionsearch_show.cgi

A data set consists of images of 10 different image categories and 100 images per category (Table 6). Images are of small size and medium resolution.

- **Amsterdam Library of Object Images (ALOI)** [ALOI, on line] [Geusebroek, 2001]-

Available: <http://staff.science.uva.nl/~aloi/>

A collection of images of small objects with dark background. Multiple images of objects captured with varied and controlled illumination conditions viewing angle, illumination angle, and illumination color.

- **Caltech** [Caltech, on line] [Fei-Fei, 2004] - Caltech 101 face dataset. Available:

http://www.vision.caltech.edu/Image_Datasets/Caltech101/Caltech101.html

A data set consisting of 101 objects, of which 437 face images of 27 different people with various indoor-outdoor complex background and different illumination conditions. Images are of medium size and medium resolution. The same face images of larger size and higher resolution are available as Caltech frontal face images 1999.

- **PASCAL challenge 2008 image database** [Everingham, on line] - Available:
<http://pascallin.ecs.soton.ac.uk/challenges/VOC/voc2008/workshop/index.html>
A data set of image segmentation competition – consisting of various challenging images.
- University of Washington, Image database [University of Washington, on line]. Available:
<http://www.cs.washington.edu/research/imagedatabase/demo/seg/>
A collection of images used for segmentation.
- **MedPics** [MedPics, on line] – An Image Library for Medical Education, UCSD – School of Medicine.
Available: <http://medpics.ucsd.edu/>
A collection of medical images.

The images of PASCAL challenge 2008 image database [Everingham, on line], University of Washington, Image database [University of Washington, on line] and MedPics [MedPics, on line] have been used only for edge detection, prominent boundaries detection & foreground extraction. Rest all have been used for testing of all proposed algorithms.

1.12 Specifications of the Developed CBIR System

- **Query**
 - By example - single image
 - Selectable from image database
 - GUI based
- **Results**
 - Retrieved similar images in order of decreasing similarity
 - Presented in multiple windows & 16 thumbnails (images) / window
- **Supported Image types**
 - jpeg
 - png

- **Image Features**

- Color Codes
- Histograms
- Regions & region attributes
- Foreground & background regions and region attributes
- Shape correlation coefficients

- **Image Retrieval**

- Based on
 - Color codes of entire image
 - Foreground color codes
 - Foreground shape correlation
 - Combination of foreground color codes and shape correlation
 - With selectable percentage proportion of weight of foreground color codes and foreground shape correlation for composite similarity measure
- Similar face – images containing complex background

- **Selectable similarity cut-offs**

The feature of the CBIR system allows user to specify the cut-off level for similarity measurement, helpful to specify the maximum permitted inexactness in the retrieved images

The GUI of the CBIR system and its description are shown in Annexure 2.

1.13 Concluding Remark

A successful endeavor of development of CBIR algorithms and the CBIR system...

2. CBIR Algorithms & CBIR Systems- Review & Analysis

2.1 Introduction

The revolutionary internet and digital technologies have imposed a need to have a system to organize abundantly available digital images for easy categorization and retrieval. The need to have a versatile and general purpose content based image retrieval (CBIR) system for a very large image database has attracted focus of many researchers of information-technology-giants and leading academic institutions for development of CBIR techniques. These techniques encompass diversified areas viz. image segmentation, image feature extraction, representation, mapping of features to semantics, storage & indexing, image similarity-distance measurement and retrieval - making CBIR system development a challenging task.

The chapter addresses and analyses challenges & issues of CBIR techniques/systems, evolved during recent years covering various methods for segmentation; edge, boundary, region, color, texture & shape based feature extraction; object detection & identification and image retrieval. The chapter also covers overview of published surveys on CBIR techniques / system. An eminent contributor's point of view on CBIR, query response analysis of some existing CBIR systems, our observations on CBIR issues and finally our approaches for developed CBIR Algorithms and system are included in the chapter.

2.2 Research & Publication Trend Analysis

The publication trend for a period of 1995 – 2005 has been shown for 'image retrieval' in [Datta, 2008]. Extending the analysis up to year 2010 for 'image retrieval' article indexed by Google Scholar as on 17-01-2011 shows tremendous increase in the research interest and hence in the number of publications post year 2005, as shown in

Figure 5, left. The publication trend for last five years has been shown in Figure 5, right for 'image retrieval' as indexed by Google Scholar as on 17-01-2011. A slight fall in number of publications for year 2010 may be due to non-visit of Google-web crawler to publishing server till date i.e. early part of year 2011.

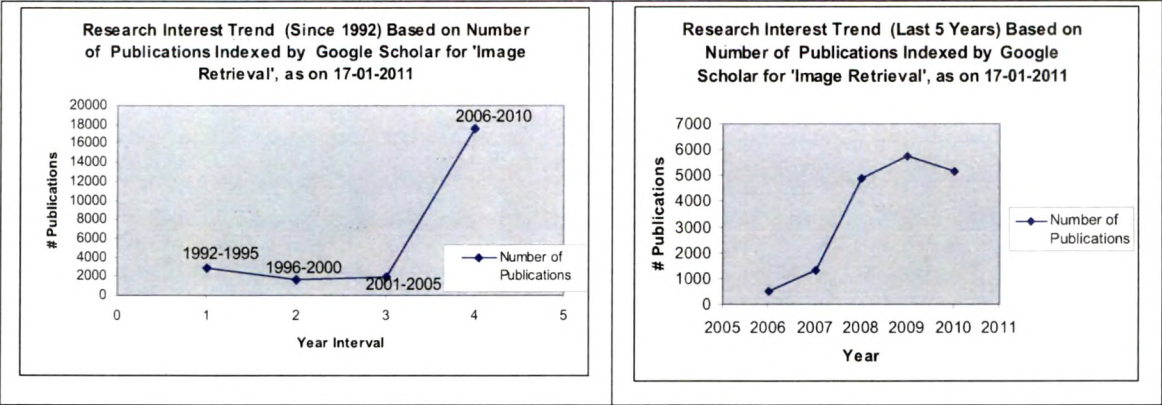


Figure 5. Research Interest Trend Based on Publications Indexed by Google Scholar for Image Retrieval. Left: Since 1992. Right: Last 5 Years.

The biggest issue for CBIR system is to incorporate versatile techniques so as to process images of diversified characteristics and categories. Many techniques for processing of low level cues are distinguished by the characteristics of domain-images. The performance of these techniques is challenged by various factors like image resolution, intra-image illumination variations, non-homogeneity of intra-region and inter-region textures, multiple and occluded objects etc. The other major difficulty, described as semantic-gap in the literature, is a gap between inferred understanding / semantics by pixel domain processing using low level cues and human perceptions of visual cues of given image. In other words, there exists a gap between mapping of extracted features and human perceived semantics. The dimensionality of the difficulty becomes adverse because of subjectivity in the visually perceived semantics, making image content description a subjective phenomenon of human perception, characterized by human psychology, emotions, and imaginations. The image retrieval system comprises of multiple inter-dependent tasks performed by various phases. Inter-tuning of all these phases of the retrieval system is inevitable for over all good results. The diversities in the images and semantic-gaps generally enforce parameter tuning & threshold-value specification suiting to the requirements. For development of a real time

CBIR system, feature processing time and query response time should be optimized. A better performance can be achieved if feature-dimensionality and space complexity of the algorithms are optimized. Specific issues, pertaining to application domains are to be addressed for meeting application-specific requirements. Choice of techniques, parameters and threshold-values are many a times application domain specific e.g. a set of techniques and parameters producing good results on an image database of natural images may not produce equally good results for medical or microbiological images.

2.3 Eminent Contributor's Point of View on CBIR

The issues presented in invited plenary talk at the *19th International Conference on Pattern Recognition*, held at Tampa, Florida during year 2008 by Theo Pavlidis - "Limitations of Content-based Image Retrieval" [Pavlidis, on line] are summarized below:

- Answer to the question to title of editorial in special issue of IEEE Proceedings [Hanjalic, 2008]: "*The Holy Grail of Multimedia Information Retrieval: So Close or Yet So Far Away?*" was
 - **So close** if published results are taken at face values
 - **Yet so far away**, based on
 - closer look at the published results
 - the results of many test sites like GazoPa [GazoPa, on line] by Hitachi, TILTOMO [tiltomo, on line], ALIPR [aliper, on line] & many others
- The CBIR systems are designed to respond to broader queries than the specific.
- The methods and trainer / classifier are tested on too small number of images.
- Whether methods search for similar 2-D images or similar objects (3-D) of the image.
- Illumination / pose / viewpoint changes offer great challenges to simple feature based CBIR systems.
- The segmentation and salient point matching fail to deal pose / viewpoint variations of many objects.
- General similarity measures are needed for real-world CBIR.

- Computationally close images may have different perceptual meaning and computationally slightly differing images may have vast perceptual differences.
- “Accept that practically significant results for *real time general CBIR* cannot be obtained unless there are major breakthroughs both in image analysis and in computer architecture.
 - **As long as we do not have general segmentation methods that can identify objects on an image, it is unwise to pursue general CBIR.”**

The current state of art and issues for automatic image annotation has also been presented by Theo Pavlidis [Pavlidis, on line] concluding direct human labeling is mandatory for proper image-tagging as automatic image tagging is very hard.

2.4 Query Response Analysis of Some of the Existing CBIR Systems

To illustrate the difficulties and complexity of CBIR, let us consider the query response of on-line demo of SIMPLIcity, developed at Penn State University, is available at http://wang14.ist.psu.edu/cgi-bin/zwang/regionsearch_show.cgi . The huge image database was consisting of 59895 images (as on 13-01-2011). The query of a deer image with ID 35600 gave no deer image in the response of displayed 32 images. Similarly, a horse-face, a close up image with id 35031 also resulted into query response of no horse image in displayed 32 images. The other on-line demo is available at <http://amazon.ece.utexas.edu/~qasim/cires.htm> having 50 image categories and around 30-50 images per category, developed at Computer and Vision Research centre, University of Texas, Austin. The image query for a wolf image named nat_anm_mam_wol_art_aw00001.jpg gave response with precision 3/20 (0.15). The enhanced version of CIRES provides browsing based on image-tags. The image browsing for tag provides 458 images (As on 14-01-2011) with precision 268/458 (0.58) which is available at <http://cires.matthewriley.com/browse/view/7117>. The precision in the enhanced version is better, but the result includes many images having human faces, perhaps due to significant weight to color in similarity matching. Given query response examples are only for the purpose of illustrating the challenges and difficulties involved in any CBIR algorithms / system and not for criticizing individual(s) or shortcomings of CBIR systems.

2.5 Overview of CBIR Surveys

The paragraph summarizes the points presented in the exhaustive survey by Datta et al. [Datta, 2008] carried out in 2008 on image retrieval covering about 300 theoretical and practical contributors. CBIR is cited in it as an emerging technology that in principal helps to organize digital pictures by their visual contents. The user perspective based on the clarity of user intention (browsing, surfing or searching), data scope (personal collection, domain specific, enterprise, archives or web) and the query form has been classified in the survey. The alternate view from system perspective involves mode of query processing & output visualization for variety of data scopes has been proposed along with it. The output visualization has been characterized as i) Relevance order e.g. Google and Yahoo image search engines ii) time order e.g. Google's Picasa – providing an option for organizing personal collection of images in time order iii) clustered iv) hierarchical and v) composite. Different image similarity comparison measures viz. Euclidian distance, weighted Euclidian, Hausdorff, Mallow, Integrated Region Matching - IRM [Li, 2000] and K –L divergence have been tabulated in the survey.

A. W. Smeulders et al. [Smeulders, 2000] reviewed 200 references in CBIR in year 2000 covering patterns of use, types of pictures, the role of semantics, the sensory gap, color, texture, local geometrical features, accumulative and global features, object and shape features, types and means of feedback and other related issues.

More than 100 articles have been reviewed in a survey on content based multimedia information retrieval by Lew et al. [Lew, 2006] in year 2006. The survey includes review of various face detection techniques for concept detection in simple and complex background. Some of the observations presented in the survey [Lew, 2006] are:

- o "One of the most important challenges and perhaps the most difficult problem in semantic understanding of media is visual concept detection in the presence of complex backgrounds."
- o "Another limiting case is where researchers have examined the problem of detecting visual concepts in laboratory conditions where the background is simple and therefore can be easily segmented."

Müller et al. presented analysis on performance evaluation in CBIR and proposed few new measures in [Müller, 2001]. They have summarized performance

measures like Precision, Recall, Precision - Recall curves, Precision – Number of Retrievals, Recall – Number of Retrievals, measures specified by TREC (Text REtrieval Conference, co-sponsored by NIST, US) and variational performance measures used by different research-contributors. They have proposed performance evaluation measures similar to those specified by TREC are summarized below:

- o Rank₁ – Rank at which first relevant image is retrieved
- o $\overline{\text{Rank}}$ – Normalized average rank of relevant images
- o PR curves
- o P (20), P (50) and P (N_g) – Precision after 20, 50 and N_g relevant images are retrieved, where N_g represents total number of relevant images.
- o R_p (0.5) – recall when precision falls below 0.5
- o R(100) – Recall after 100 images are retrieved

A comprehensive survey of methods for Colour Image Indexing and Retrieval in Image Databases has been presented by Schettini et al. [Schettini, 2001] covering various color discretization methods, color indexing, issues of histogram comparisons due to color shifts, color spatial indices, illuminant invariant color image indexing techniques and related issues along with methods combining other features with color feature.

2.6 CBIR and Related Techniques

Various techniques for extraction and representation of image features like histograms – local (corresponding to regions or sub-image) or global , color layouts, gradients, edges, contours, boundaries & regions, textures and shapes have been reported in the literature.

Histogram is one of the simplest and computationally inexpensive image features. Despite being invariant to translation and rotation about viewing axis, lack of inclusion of spatial information is its major draw back. Many totally dissimilar images may have similar histograms as spatial information of pixels is not reflected in the histograms. Consequently, many histogram refinement techniques have been reported in the literature. The simplest form of histogram is having fixed number of bins. A bin corresponds to a fixed range of intensity values. The range of intensity values defines the width of a bin. Generally, all bins are of equal width. The intensity (color) changes in the same image may shift the bin-membership of pixels altering the intensity distribution.

The nature and amount of change in distribution depends on number of bins and value of change of intensity. The increase in bin-width will reduce the effect of changes in the distribution due to intensity (color) changes. Increase in bin-width results into loss of intra-range distribution information for a larger range. Hence, deciding the bin-width plays important role in the similarity measure methods. Given fixed width of bins may not be suitable to all categories of images. The cumulative histogram gives the cumulative distribution of the intensity values.

The multi-resolution histogram and its use for recognition for image and video retrieval have been proposed in [Hadjidemetriou, 2004]. As defined in it, the multi-resolution histogram is the set of intensity histograms of an image at multiple image resolutions. Like plain histograms, multi-resolution histograms are fast to compute, space efficient, invariant to rigid motions, and robust to noise. In addition, spatial information is directly encoded with multi-resolution histograms. A novel matching algorithm based on the multi-resolution histogram that uses the differences between histograms of consecutive image has been proposed in it [Hadjidemetriou, 2004] along with the effect of shape parameters on the multi-resolution histograms.

Histogram intersection based method for comparing model and image histograms was proposed in [Swain, 1991] for object identification. Histogram refinement based on color coherence vectors was proposed in [Pass, 1996]. The technique considers spatial information and classifies pixels of histogram buckets as coherent if they belong to a small region and incoherent otherwise. Though being computationally expensive, the technique improves performance of histogram based matching. Color correlogram feature for images was proposed in [Huang, 1997] which takes into account local color spatial correlation as well as global distribution of this spatial correlation. The correlogram gives the change of spatial correlation of pairs of colors with distance and hence performs well over classical histogram based techniques. A modified histogram based technique to incorporate spatial layout information of each color with annular, angular and hybrid histograms has been proposed in [Rao, 1999]. In [Stricker, 1995], cumulative histogram and respective distances for image similarity measures, overcoming quantization problem of the histogram bins was proposed. The representation of color distribution features for each color channel based on average, variance and skewness, described as moments, for image similarity was also presented.

Various segmentation techniques based on edge detection, contour detection and region formation have been reported in the literature. These techniques, in general, process low level cues for deriving image features by following bottom-up approach. Automatic image segmentation is a very crucial phase as the overall performance of retrieval results significantly depends on the precision of the segmentation. The most difficult task for any automatic image segmentation algorithm is to avoid under and over segmentation of images, possessing diversified characteristics. Hence, for required scale of segmentation, parameter tuning or threshold adjustment becomes unavoidable for versatile image segmentation algorithms.

Directional changes in color and texture have been identified in [Ma, 2000], using predictive color model to detect boundaries by iteratively propagating edge flow. This iterative method is computationally expensive because of processing of low level cues at all pixels for given scale.

A novel hierarchical classification frame work based approach for boundary extraction with Ultrametric Contour Maps UCM - representing geometric structure of an image has been proposed in [Arbel'aez, 2006]. A generic grouping algorithm based on Oriented Watershed Transform and UCM [Arbel'aez, 2006] has been proposed in [Arbel'aez, 2009] to form a hierarchical region tree, finally leading to segmentation. The method enforces bounding contour closures, avoiding leaks - a root cause of under segmentation. Exhaustive precision-recall evaluation of OWT-UCM technique for different scales also has been presented. The precise low level processing is very crucial for feature extraction. J. Malik et al. emphasis on perfect boundary detection leading to segmentation and / or object shape description. Local and global cue based contour and junction detection has been proposed in [Maire, 2008].

Image segmentation has been treated as a graph partitioning problem using Normalized cuts in [Shi, 2000]. Two powerful segmentation strategies—mean shift clustering and normalized cuts based accurate and rapid object initialization scheme—weighted mean shift normalized cuts for geodesic active contour model for segmentation of Histopathology images have been presented in [Xu, 2010].

Region based image retrieval, incorporating graphs, multiple low level labels and their propagation, multilevel semantic representation and support vector machine has been proposed by Li et al., implying effectiveness of the method by showing various precision measures only. The recall measure analysis for the incorporated large image

database of Corel consisting of 10000 images would have been helpful throwing lights on effect of various types of feedbacks on region matching. It should be noted that the paper [Li, 2008] uses JSEG [Deng, on line] [Deng, 2001] algorithm for image segmentation.

The brightness and texture gradient based probability of boundary [Martin, 2004] has been used to generate edge-map, which is scale-invariant representation of image from the bottom up, using a piecewise linear approximation of contours and constrained Delaunay triangulation for completing gaps. The curvilinear grouping on top of this graphical/geometric structure using a conditional random field to capture the statistics of continuity and different junction types has been proposed in [Ren, 2008] for contour completion in natural images. A new concept of Boolean derivatives as a fusion of partial derivatives of Boolean function for edge detection algorithms for binary and gray scale image has been presented and results have been compared with traditional edge detection algorithms in [Agaian, 2010]. Color gradient detection based technique for automatic image segmentation has been proposed in [Ugariza, 2009]. A method for unsupervised determination of hysteresis thresholds for edge detection by combining advantages and disadvantages of thresholding methods by finding best edge map, a subset and an overset of the unknown edge point set has been proposed in [Medina-Carnicer, 2010]. Combined top-down and bottom-up approach for image segmentation has been proposed in [Borenstein, 2008]. Top-down & bottom-up cue based probabilistic method for image segmentation overcoming the limitations of traditional conditional random field (CRF) based approach has been proposed in [Pawan Kumar, 2010]. YCbCr color model based automatic seeded region growing algorithm for image segmentation has been proposed in [Shih, 2005].

YCbCr color space based face detection algorithm for varying lightning conditions and complex background has been proposed in [Hsu, 2002] that incorporates light compensation technique and non-linear color transform. Major face detection techniques have been also listed in [Hsu, 2002]. Analysis and comparison of color representation, color quantization and classification algorithms for skin segmentation have been reported in [Phung, 2005]. Exhaustive survey of face detection issues and techniques has been found in [Yang, 2002]. The foreground objects revealing by separating background in the images of standard data set has been proposed in [Thakore, 2010, 4].

Various techniques based on generalized Hough transform and Fourier descriptors have been reported in the literature for shape and object boundary detection. A review of methods for shape comparison has been reported in [Veltkamp, 2000]. Active contour model – snake has been used in [Kass, 1988] for interactive interpretation, where user-imposed constraint forces guide the snake to feature of interest. Many variations based on active contour methods have been found in literature. The boundary detection precision of active contour based methods is generally sensitive to seed-points or seed-contours; if not provided properly, snakes may not converge to true object boundaries. $L^* u^* v^*$ color space recursive mean shift procedure based analysis of multimodal feature space and delineation of arbitrarily shaped cluster can be found in [Comaniciu, 2002]. Scale invariant local shape features with chains of k -connected roughly straight family of contour segments has been used for object class detection in [Ferrari, 2008].

The boundary structures and global shape feature based approach for segmentation and object detection has been proposed in [Toshev, 2010]. Image segmentation and object detection using iterated Graph Cuts, based on local texture features of wavelet coefficient has been reported in [Fukuda, 2008]. The application of watershed algorithm for contour detection leading to segmentation was proposed in [Beucher, 1979].

Many relevance feed back techniques have been proposed in literature to bridge the semantic gap by specifying positive and negative feed backs given by the user for refinement of results. A relevance feedback based interactive image retrieval approach to address issues of semantic-gap and subjectivity of human perception of visual contents was introduced in [Rui, 1998], which showed significant improvement in the results. In [Tao, 2008], orthogonal complement component based relevance feed back technique is proposed that does not treat positive and negative feed backs equivalently, as the former share homogenous concepts whereas latter do not. Generalized Bayesian learning framework with target query and a user conception based user-model has been proposed in [Hsu, 2005], where target distribution, target query and matching criteria have been updated at every feed back step.

A fuzzy approach based CBIR, named FIRST - Fuzzy Image Retrieval SysTem, has been proposed in [Krishnapuram, 2004] to handle the vagueness in the user queries and inherent uncertainty in image representation, similarity measure and relevance

feedback incorporating fuzzy attributed relational graph comparisons for similarity measures. Contour and texture cues have been exploited simultaneously in [Malik, 2001] using intervening contour framework and textons for image segmentation with spectral graph theoretic framework of normalized cuts. As stated in [Malik, 2001] contour based image segmentation approaches have edge detection as the first stage followed by edge linking stage to exploit curvilinear continuity. Perceptual grouping of block based visual patterns using modified Hough transform for object search technique in heterogeneous cluster-oriented CBIR with load balancing implementation has been reported in [Cheng, 2007]. Two new texture features - Block difference of inverse probabilities (BDIP), measuring local brightness variations & block variation of local correlation coefficients (BVLC), and measuring local texture smoothness have been used in [Chun, 2003] and the combination of BDIP and BVLC moments for image retrieval improves performance compared to wavelet moments. Evolutionary group algorithm to optimize the quantization thresholds of the wavelet-correlogram has been reported in [Saadatmand-Tarzjan, 2007].

A color image edge detection algorithm was proposed in [Dutta, 2009], taking up average maximum color difference value was used to predict the optimum threshold value for a color image and thinning technique was applied to extract proper edges producing comparable results with other edge detection algorithms. The presented method results [Dutta, 2009] were of images containing color patches with no or minimum textures. The performance of the algorithm could have been tested for textured and natural images.

Prasad et al. [Prasad, 2004] proposed image retrieval using integrated color-shape-location index has been proposed based on grouping RGB color space into 25 perceptual color categories, dominant region eccentricity for 8 shape categories and grid cell of image for location identification and indexing with performance measure on database consisting of various national flags and vegetables-fruits images.

The use of CIE Lab color space based color descriptors for CBIR and comparisons for different quantization methods, histograms calculated using color-only and/or spatial-color information with different similarity measures have been presented in [Gavrielides, 2006] covering retrieval results for images with different transformations like scaling, rotation, cropping, jpeg compression with different quality factors, blurring, illumination changes, contrast adjustments and various adaptive noise attacks.

Morphology-based approaches for CBIR by making use of granulometries independently computed for each sub-quantized color and employing the principle of multi-resolution histograms for describing color, using respectively morphological leveling and watersheds has been proposed in [Aptoula, 2009] for LSH color space.

The relevance feed back based biased discriminative Euclidean embedding (BDEE) was proposed in [Bian, 2010] which parameterizes samples in the original high-dimensional ambient space to discover the intrinsic coordinate of image low-level visual features showing precise modeling of both - the intra-class geometry and interclass discrimination evaluated on Corel image database presenting query response examples.

Comparison of the mean average precision of three content based image retrieval methodologies have been presented in [Vasconcelos, 2007], indicating improvements in the performance over last few years. Performance comparison of query by visual example and query by semantic example has been reported in [Vasconcelos, 2007], demonstrating superior performance of the latter. As reported, the content based image retrieval methodologies have evolved from modeling visual appearance, to learning semantic models and finally to making inferences using semantic spaces. Performance comparison of minimum probability of error retrieval frame work based query by visual example and query by semantic example has been reported in [Rasiwasia, 2007], concluding semantic representations of images have an intrinsic benefit for image retrieval. Elaborative study of query by semantic example addressing structure of semantic space and effect of low level visual features & high level semantic features on over all performance of CBIR system has been reported in [Rasiwasia, 2008].

Paitakes et al. [Pratikakis, 2006] proposed a novel unsupervised method for image retrieval based on hierarchical watershed algorithm applied on $L^*a^*b^*$ color or texture feature space for achieving meaningful segmentation & automatic meaningful region extraction leading to construction of region adjacency graph – RAG incorporating scale based weights in the multi-scale hierarchical frame work for Earth mover's distance (EMD) computation as region-similarity-comparison. The segmentation results produced with their proposed method have been compared with those of JSEG [Deng, on line] [Deng, 2001], E-M algorithm (Blobworld) [Carson, 2002] and graph-based segmentation [Felzenszwalb, 2004]. The image retrieval results have

been compared with the regions generated with their proposed method, JSEG [Deng, on line] [Deng, 2001], E-M algorithm (Blobworld) [Carson, 2002], and graph-based segmentation [Felzenszwalb, 2004] for segmentation by applying their method of region similarity. The mean precision-recall have been measured in [Pratikakis, 2006] for 10 queries per image-class of image database consisting of total 1000 images of 10 different classes with 100 images per class, which reads (approximate values) for all categories of images, highest mean precision of 0.7 at mean recall of 0.07 and highest mean recall of 0.425 with precision of 0.41. Further, P – R curves corresponding to all image categories indicate that it is not possible to retrieve images with precision as 1 at any cost of recall, i.e. for no case, only relevant images (may be very few in number) gets retrieved. As no examples of query responses have been presented in [Pratikakis, 2006], analysis and inference about ordering and ranking of the retrieved resultant images cannot be carried out.

The slope magnitude method along with Sobel, Prewitt, Robert and Canny gradient operators have been used for forming shape image on which block truncation coding (BTC) is applied in [Kekre, 2010, 1] for performing image retrieval on 1000 images of 11 different classes of SIMPLiCity image database [Wang, 2001] [SIMPLiCity, on line]. The performance analysis of 55 queries in a form of precision & recall plotted for number of images retrieved with different combination of methods. Simple morphological edge detection, top-hat morphological edge detection, bottom-hat morphological edge detection methods are combined with BTC for image retrieval in [Kekre, 2010, 2]. The precision is not exceeding 0.8 even for 2 retrieved images. The maximum recall obtained in all seven methods proposed is about 0.35 for 100 retrieved images (relevant + irrelevant) with maximum precision of about 0.35. Here also in both publications [Kekre, 2010, 1] & [Kekre, 2010, 2], no examples of results of query response has been presented to enable analysis of results for order and ranking of retrieved images.

Basak et al. [Basak, 2006] presented Multiple Exemplar-Based Facial Image Retrieval Using Independent Component Analysis as a specific CBIR application for three different image databases, of which one is Caltech [Caltech, on line] [Fei-Fei, 2004] for which, as can be seen from the illustrative query response examples, non-face portion of the image constituting background have been cropped. It is to be noted that our proposed method for similar-face-image retrieval uses Caltech images, where the complex-background, which contributes to major portion of the image has been

excluded based on prominent boundaries and foreground detection based techniques to extract the face region before comparing them for similarity measures.

2.7 CBIR Systems

A brief summary of some of the CBIR systems has been presented in this section. QBIC - Query By Image Content system, developed by IBM, makes visual content similarity comparisons of images based on properties such as color percentages, color layout, and textures occurring in the images. The query can either be example images, user-constructed sketches and drawings or selected color and texture patterns [QBIC, on line] [Flickner, 1995]. The IBM developed QBIC technology based Ultimedia Manager Product for retrieval of visually similar images [Barber, 1994]. Virage [Virage, on line] and Excalibur are other developers of commercial CBIR systems.

VisualSEEk - a joint spatial-feature image search engine developed at Columbia university performs image similarity comparison by matching salient color regions for their colors, sizes and absolute & relative spatial locations [Smith, 1996] [VisualSEEk, on line]. Photobook developed at Media Laboratory, Massachusetts Institute of Technology – MIT for image retrieval based on image contents where in color, shape and texture features are matched for Euclidean, mahalanobis, divergence, vector space angle, histogram, Fourier peak, and wavelet tree distances. The incorporation of interactive learning agent, named FourEyes for selecting & combining feature-based models has been a unique feature of Photobook [Photobook, on line]. MARS - Multimedia Analysis and Retrieval Systems [MARS, on line] and FIRE- Flexible Image Retrieval Engine [Fire, on line] incorporate relevance feed back from the user for subsequent result refinements. Similar images are retrieved based on color features, Gabor filter bank based texture features, Fourier descriptor based shape features and spatial location information of segmented image regions in NeTra [Ma, 1997]. For efficient indexing, color features of image regions has been represented as subsets of color code book containing total of 256 colors. The frame work proposed in [Ma, 2000] has been incorporated for image segmentation in NeTra. PicSOM (Picture & Self-organizing Map) was implemented using tree structured SOM, where SOM was used for image similarity scoring method [Laaksonen, 1999]. Visual content descriptors of MPEG-7 (Moving Pictures Expert Group Multimedia Content Description Interface) were used in PicSOM [Laaksonen, 1999] for CBIR techniques and performance comparison with Vector Quantization based system

was proposed in [Laaksonen, 2002]. Incorporation of relevance feedback in it caused improvements in the precision of results of PicSOM. SIMPLicity – Semantic-sensitive Integrated Matching for Picture Libraries incorporates integrated region matching methodology for overcoming issues related to improper image segmentation. The segmented images are represented as sets of regions. These regions, roughly corresponding to objects are characterized by their colors, shapes, textures and locations. The image search is narrowed-down by applying image-semantic-sensitive categorization for better retrieval performance [Wang, 2001]. The online demo of SIMPLicity is available at [SIMPLicity, on line].

A comparative survey of various 42 CBIR systems developed by year 2001 has been reported in [Veltkamp, on line]. The survey compares those systems for mode of querying, features & method of similarity comparisons and indexing techniques. The sample query responses along with URL for demo have been provided (Though very few demo URL pages are accessible as on date, the survey is worth noting for its contents & the query response examples).

Many technology-giants now have research focus on multimedia / video retrieval techniques. The multimedia information retrieval issues and recent research publications covered in the IEEE special issue on Advances in Multimedia Information Retrieval (April 2008) have been summarized in the editorial article [Hanjalic, 2008]. Similarly, the editorial article of special issue on Recent Advances in Image and Video Retrieval, in proceedings of IEE, 2005, summarizes the papers of the issue [O'connor, 2006].

The video-frame based image analysis as the application for video abstraction / video summary has been overviewed in [Li, 2001] along with list of major players in the field.

2.8 Our Observations

The observations learnt and derived for CBIR are as under.

- 'Proper' segmentation is a mandatory requirement for feature extraction.
 - The word 'Proper' is subjective and depends on image characteristics and categories of images. Improper segmentation leads to under-segmentation / over-segmentation of the image under consideration.

- Minimization of over segmentation sometimes leads to under segmentation.
- Inter-region and intra-region texture & illumination variations are root causes of over-segmentation.
- The segmentation algorithm should not be very sensitive to such variations.
- The automatic, non-parametric generic segmentation algorithm should be multi-scale / hierarchical to accommodate image characteristic variations.
- The performance of region matching algorithms greatly depends on the quality of segmentation.
- Region matching based techniques are characterized by
 - Image comparisons by parts – may not be suitable to some categories of images.
 - Finding best matched region by performing one to many region-attribute comparison faces the challenges of resolving clash to determine best matched region and accommodating color / illumination / shape / spatial location changes in the region.
- Proper identification of regions constituting object(s) is the second most important requirement. The process of region-identification may be characterized by region merging / region separating / region propagating / region eliminating operations.
- The user interaction (intervention) in a form of relevance feed-back generally improves the performance by accommodating variations in images and by overcoming limitations of algorithms.
- Stringent image-feature descriptors or strict similarity constraints end up into poor recall with higher precision, assuming other conditions favorable and constant. Broad image-feature descriptors or relaxed similarity measures improve recall by sacrificing precision. (Examples of broad image-descriptors are – histograms, color quantization based features, concept descriptors etc. A Gabor filter applied at various orientations for texture description is an example of stringent image feature.)
- Recall improves if a technique is capable of taking care of variations in color / illumination / poses / shapes in similar images.

- Precision improves if a technique is precise in feature extraction and similarity measures.
- One of the objectives for a design / implementation of a good CBIR system would be 'To retain maximum precision for higher recall in a large image database consisting of variety of images'.
- And, exhaustive testing of methods on variety of image categories is necessary for proving applicability and suitability.

2.9 Our Approaches

The theme of our approaches is *"Relaxed feature description for better Recall and simultaneous emphasizing of reliable processing of cues leading to precise feature extraction for better Precision"*.

Our approaches follow two streams of techniques. The first one is based on broad color feature descriptors called color codes, which is a simple and computationally efficient technique, suitable for image comparison on a broader scale, on the basis of color comparison without considering shapes. The approach is suitable for finding near-similar images having nearly similar color distributions. The second approach emphasizes reliable processing of low level cues for precise and well localized prominent boundaries detection eventually leading to foreground extraction. The extracted foreground is compared on basis of shape - correlation and foreground color codes. The composite approach consisting of foreground shape and foreground color codes provides selectable proportion of weights in composite similarity measures enables users to match the need based on category of query image. The exclusion of background and corresponding features enables object based search for image retrieval. The foreground detection based face extraction method for similar-face-image retrieval from the image containing complex-background has been presented as an application-specific CBIR, illustrating effectiveness of various proposed algorithms.

Our approaches address some of the issues observed by Theo Pavlidis, described in [Section 2.3](#):

- The proposed segmentation method yields good results for wide categories of images enabling foreground (objects) extraction by separating background and facilitating foreground comparison instead of comparison of whole

images. (Addressing of segmentation and object based search issues cited in [Section 2.3](#))

- The foreground extraction method meets the challenges offered by Illumination / pose / viewpoint changes.
- The other approach enables whole image comparison for image retrieval based on broad image-feature descriptors (Color Codes) giving user a selection for foreground or whole image based searches depending upon user's intentions.

Additional characteristics of the approaches are,

- Broader color descriptors – color codes are less sensitive to illumination and color variations up to certain extent, a very helpful characteristic to improve recall.
- The reliable processing yielding precise prominent boundaries & foreground shape combined with broader color descriptors of foreground form a good proposition for object based image retrieval intended for improvements in precision and recall respectively.
- Encouraging results for exhaustive testing of methods on various images.

2.10 Discussion

The road map of development of CBIR techniques began with simple primitive features based indexing methodologies that later got enhanced with combinational features. Two major issues, semantic-gap and subjectivity of semantics are addressed by the state of the art techniques. Many state of the art techniques incorporate iterative relevance feed back from user for refinement of results. Semantic gap bridging approaches based on fuzzy, evolutionary and neural network have also been reported. Hierarchical approaches for feature extraction and representations achieve hierarchical abstraction; help matching semantics of visual perception of human beings. Several modern techniques focus on improvements on processing of low level cues so as to precisely extract features. Many state of the art techniques suggest that semantic domain based image retrieval systems, comparing meaningful concepts improve quality of retrieved image set. Effective learning and inferring of meaningful concepts may get proved critical for such systems.

The proposed combination of reliable processing leading to precise feature extraction and broader color descriptors applied to foreground shape leads to encouraging results for foreground based image retrieval for better precision and recall measures. The results of proposed independent approaches, based on – whole image color codes, foreground shape and foreground color codes have shown applicability and suitability of the methods for image retrieval. Various proposed methods put together for development of application specific CBIR - similar-face-image retrieval for images containing complex background, produces results endorsing the effectiveness of methods.

2.11 Concluding Remark

The state of the art image retrieval techniques have a vast scope of under-going significant technical evolution...

3. Technical Background

3.1 Introduction

The chapter briefly covers wavelets, continuous wavelet transform, discrete wavelet transform and stationary wavelet transform. It also includes image retrieval performance measures - Precision and Recall along with Precision-Recall curves presented as performance evaluation measures of various algorithms. The watershed region related issues are covered lastly in the chapter. The prominent boundaries detection and prominent boundaries detection based algorithms incorporates stationary Haar wavelet decompositions at different levels. The watershed transforms and prominent boundary has been utilized in foreground revealing method.

3.2 Signal Analysis

Fourier transform is a time domain to frequency domain transformation that represents a signal under consideration into harmonics of sinusoidal having different frequencies, amplitudes and phases. As the signal is being transformed into frequency domain, the powerful transformation has a serious drawback of not retaining time information. Hence, time localization of an instance of an event cannot be detected and analyzed with Fourier transform. To overcome this shortcoming, Gabor proposed Short-time Fourier analysis using *windowing* a signal under consideration. The windowed signal is analyzed using Fourier analysis giving frequency domain representation of a windowed signal by retaining time information. The accuracy of detection of the instance of an event depends on the size of the selected window and the frequency content of the signal. For all practical non-stationary signals, a fixed sized window is not suitable and adaptively adjusting the window size is difficult. Hence, short-time Fourier analysis is not suitable for many applications.

The solution to the question of detecting time-instance of an event is Wavelet analysis. As the name suggests, Wavelets are small time limited waves having zero average value. Different types of available wavelets are shown in Table 1 and mother

wavelets and corresponding scaling functions in Figure 6. These wavelets are the basis function for wavelet analysis. The wavelet analysis represents signal with scaled and shifted versions of mother wavelets.

3.2.1 Types of Mother Wavelets

Different wavelet families and corresponding mother wavelets are tabulated below. ** following wavelets in the last column of the table indicate a wavelet being a part of an infinite family of wavelets.

Table 1. Types of Wavelets.

Sr. No	Mother wavelet family names	Abbreviations	Wavelets
1	Haar	haar	
2	Daubechies	db	db1 db2 db3 db4 db5 db6 db7 db8 db9 db10 db**
3	Symlets	sym	sym2 sym3 sym4 sym5 sym6 sym7 sym8 sym**
4	Coiflets	coif	coif1 coif2 coif3 coif4 coif5
5	BiorSplines	bior	bior1.1 bior1.3 bior1.5 bior2.2 bior2.4 bior2.6 bior2.8 bior3.1 bior3.3 bior3.5 bior3.7 bior3.9 bior4.4 bior5.5 bior6.8
6	ReverseBior	rbio	rbio1.1 rbio1.3 rbio1.5 rbio2.2 rbio2.4 rbio2.6 rbio2.8 rbio3.1 rbio3.3 rbio3.5 rbio3.7 rbio3.9 rbio4.4 rbio5.5 rbio6.8
7	Meyer	meyr	
8	DMeyer	dmey	
9	Gaussian	gaus	gaus1 gaus2 gaus3 gaus4 gaus5 gaus6 gaus7 gaus8 gaus**
10	Mexican_hat	mexh	
11	Morlet	morl	
12	Complex Gaussian	gaus	cgau1 cgau2 cgau3 cgau4 cgau5 cgau**
13	Shannon	shan	shan1-1.5 shan1-1 shan1-0.5 shan1-0.1 shan2-3 shan**
14	Frequency B-Spline	fbsp	fbsp1-1-1.5 fbsp1-1-1 fbsp1-1-0.5 fbsp2-1-1 fbsp2-1-0.5 fbsp2-1-0.1 fbsp**
15	Complex Morlet	cmor	cmor1-1.5 cmor1-1 cmor1-0.5 cmor1-0.1 cmor**

3.2.2 Example - Mother Wavelets & Scaling Functions

The mother wavelets and corresponding scaling functions are plotted with the help of Matlab functions and shown below in Figure 6.

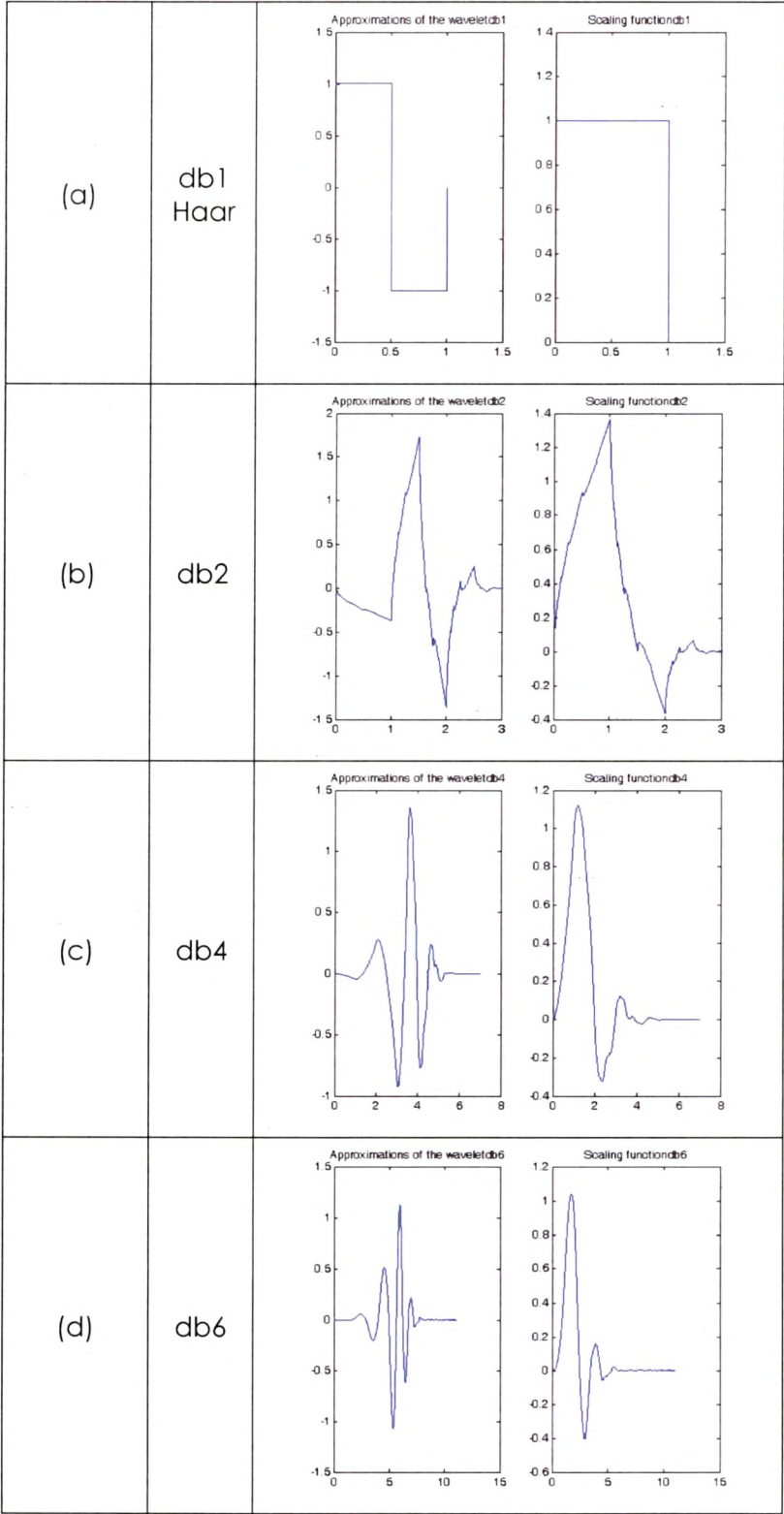


Figure 6. Mother Wavelets and Corresponding Scaling Functions.

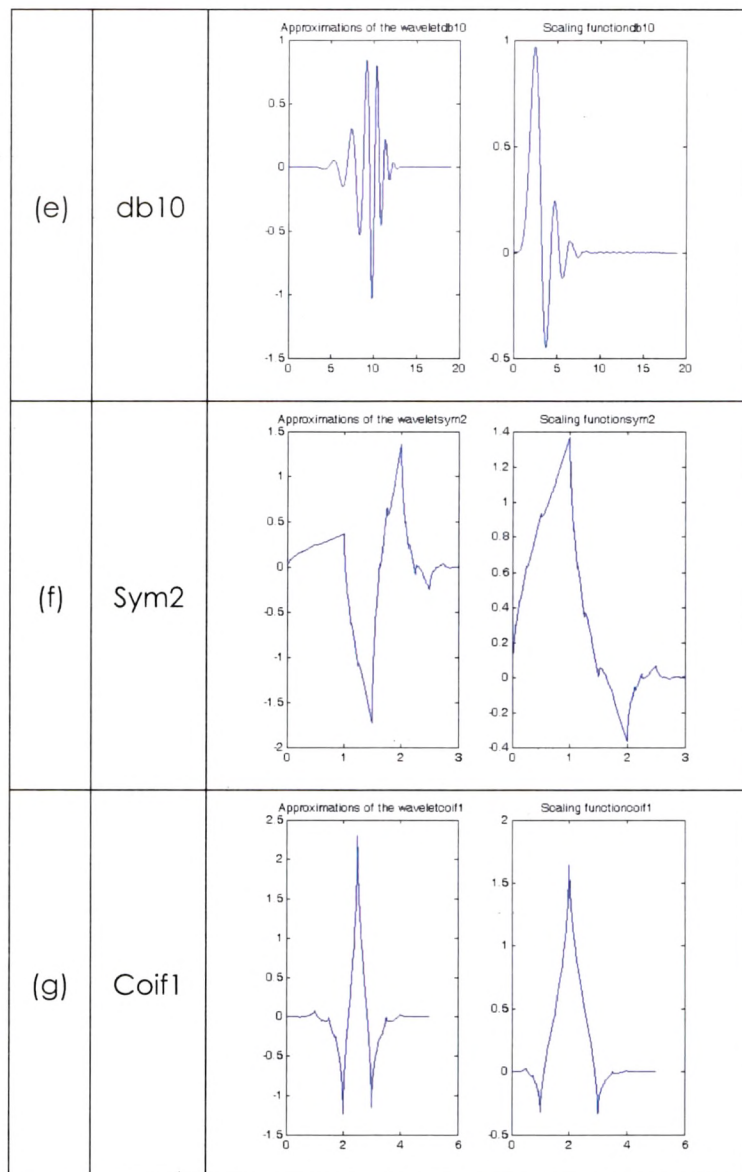


Figure 6 (Contd). Mother Wavelets and Corresponding Scaling Functions.

The Haar mother wavelet function $\psi(x)$, also known as db1 mother wavelet, can be described as

$$\begin{aligned}
 \psi(x) &= 1, & \text{if } x \in [0, 1/2[, \\
 \psi(x) &= -1, & \text{if } x \in [1/2, 1[, \\
 \psi(x) &= 0, & \text{otherwise.}
 \end{aligned}
 \tag{3.1}$$

and its scaling function $\phi(x)$ can be given as

$$\begin{aligned}
 \phi(x) &= 1, & \text{if } x \in [0, 1[, \\
 \phi(x) &= 0, & \text{otherwise.}
 \end{aligned}
 \tag{3.2}$$

Haar mother wavelet is the simplest and the first wavelet. The step like wavelet and simple scaling functions are shown in Figure 6 (a).

3.2.3 Continuous Wavelet Transform

Continuous Wavelet transform (CWT) is defined as

$$C(a, b) = \int_{-\infty}^{\infty} \phi(t) \Psi(a, b, t) dt \quad (3.3)$$

Where,

C is the wavelet coefficients characterized by scale a , and shift b .

$\phi(t)$ is a function, whose wavelet transform is sought.

$\Psi(a, b, t)$ is a mother wavelet that is scaled and shifted for coefficient computation.

Thus wavelet coefficient at given scale and shift is computed by multiplying $\phi(t)$ with scaled and shifted version of mother wavelets and summing up the results in other words, $\phi(t)$ is convolved with scaled and shifted version of $\Psi(a, b, t)$ for a given scale a and shift b . Figure 7 [Ha, on line] shows scale and shift operations on mother wavelet.

The scale and shifts are continuous in the CWT resulting into production of a large number of wavelet coefficients. These computations are not only computationally-expensive, but also more than enough for majority of practical applications.

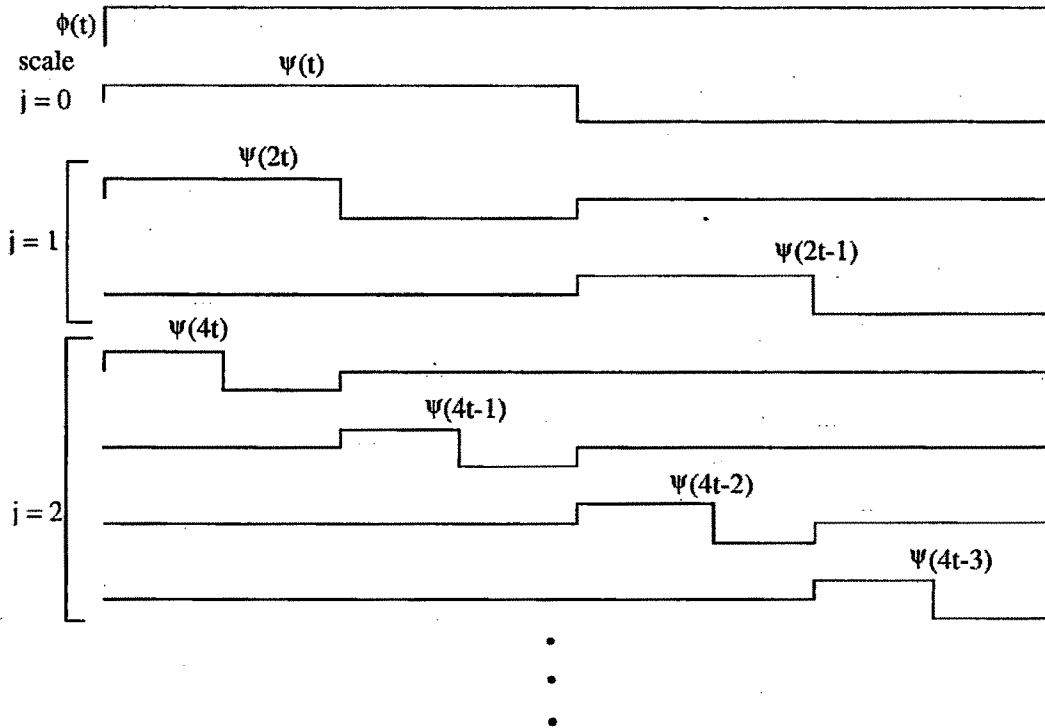


Figure 7. Shift and Scale Operation on Mother Wavelet [Ha, on line].

3.2.4 Discrete Wavelet Transform (DWT) & Stationary Wavelet Transform (SWT)

The Discrete Wavelet transform selects scale and shifts based on power of 2 to have faster and accurate enough transform with reduced number of wavelet coefficients. Figure 8 illustrates DWT operation. The classical Discrete Wavelet Transform (DWT) at each level convolves the signal with low and high pass filters and then performs decimation operation to generally discard odd coefficients and preserve even ones. The DWT decomposes sampled signal into approximate and detailed components. The filter pair is designed corresponding to the mother wavelet under consideration. These coefficients are down sampled by a factor of 2 by generally discarding every odd coefficient. This process is known as decimation operation. The approximate and detail coefficients have half the length totaling same number of coefficients as original number of samples. The approximate coefficients are low frequency components whereas detail coefficients are high pass counter parts. The mathematical model for computation and interpretation of multi-resolution signal decomposition as wavelet representation and the extension of orthogonal wavelet representation for images was proposed in [Mallat, 1989].

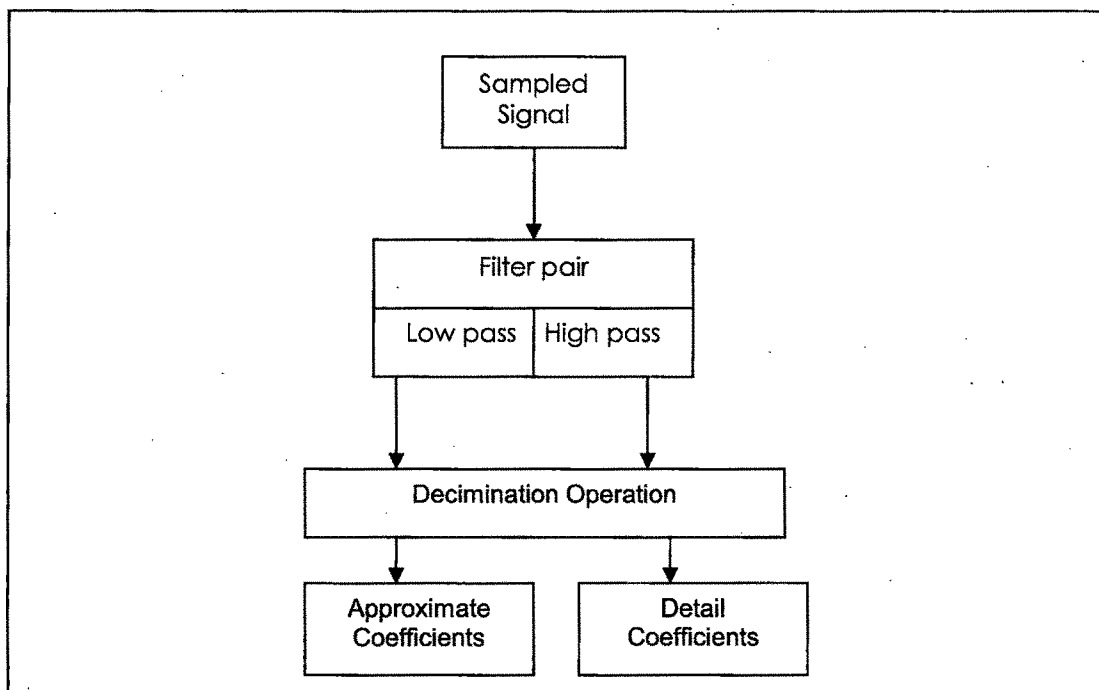


Figure 8. Block diagram for DWT.

The approximate coefficients can be further decomposed into approximate and detailed coefficients of next level, producing multi-level DWT as shown in Figure 9. The resultant tree known as wavelet decomposition tree that is useful for hierarchical analysis. The wavelet synthesis reconstructs the signal using wavelet coefficients, up-sampling and complementary filters.

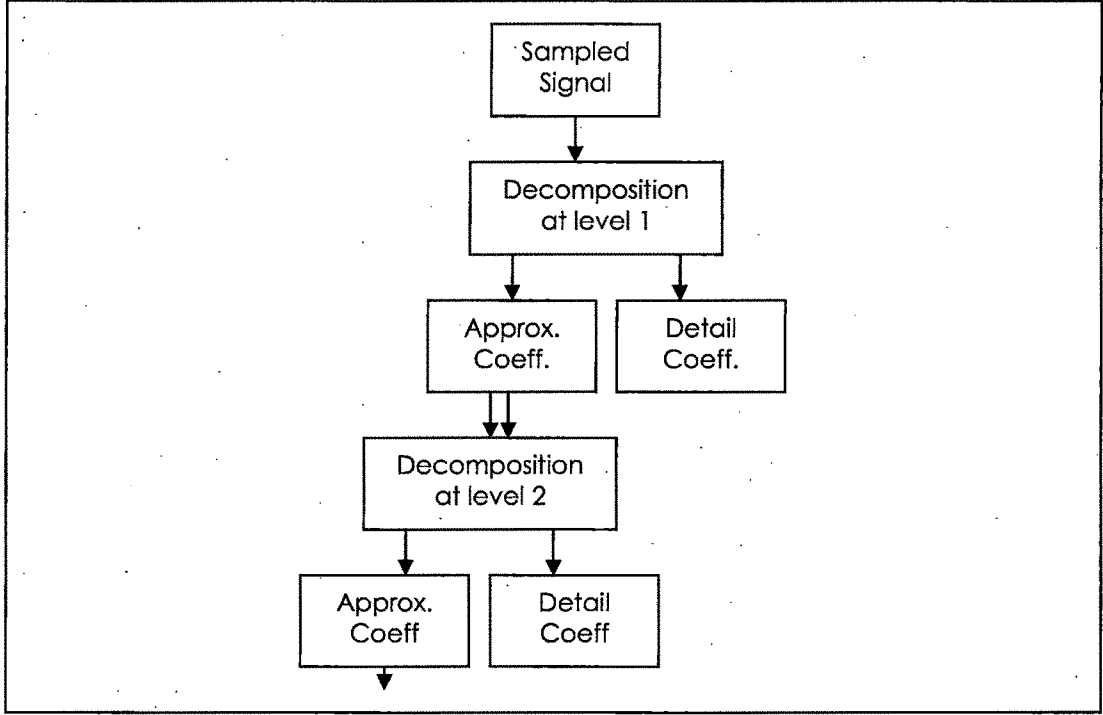


Figure 9. Wavelet - Multi-level Decomposition.

Maximum number of levels in DWT and SWT for given sequence of length N is $J = \log_2 (N)$ as given in [Nason, 1995].

The coefficients for DWT, as denoted in [Nason, 1995],

$$c^j = D_0 H c^{j+1} \text{ and } d^j = D_0 G c^{j+1} \text{ for } j = J-1, J-2, \dots, 0 \quad (3.4)$$

Where,

c^J is initialized with original sequence data.

D_0 is a binary decimation operator, keeping even indexed coefficients in the sequence.

H is a low pass filter producing approximate (smooth) coefficients c^j .

G is a high pass filter producing detailed coefficients d^j .

Decimation operation of DWT given in (3.4) causes length reduction of vectors c and d by a factor of 2 at every level which makes DWT unsuitable for the proposed method. The length reduction in the vectors introduces difficulties to exactly map the boundaries on to the image. So, the method incorporates SWT where there is no decimation

operation involved for coefficient computations resulting into same length vectors of coefficients at all levels.

The coefficients for SWT, as denoted in [Nason, 1995],

$$a^{j+1} = H^{[J-j]} a^j \text{ and } b^{j+1} = G^{[J-j]} a^j \text{ for } j = J, J-1, \dots, 1. \quad (3.5)$$

Where,

a^j is initialized with original sequence data.

$H^{[J-j]}$ is a low pass filter used at level j for producing approximate (smooth) coefficients a^j .

$G^{[J-j]}$ is a high pass filter used at level j for producing detailed coefficients b^j .

The H and G are required to be modified at every level so as to have length N for approximate and detailed coefficients, same as the length of original data.

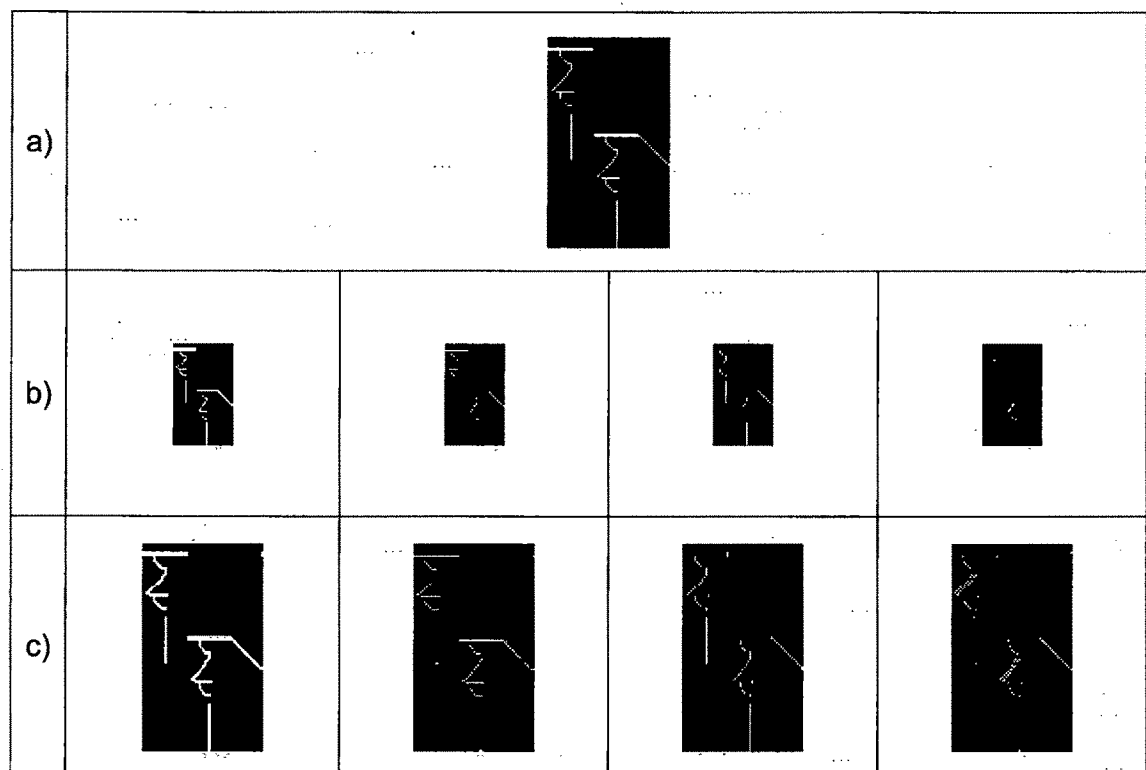


Figure 10. Decomposition with Discrete Haar Wavelet and Discrete Stationary Haar Wavelet. (a) Original Image. (b) DWT – Haar decomposition at level 1, from left to right - Approximate, Horizontal, Vertical and Diagonal coefficients (c) SWT –Haar decomposition at level 1, from left to right Approximate, Horizontal, Vertical and Diagonal coefficients

The qualitative comparison of results of SWT and DWT decompositions with Haar mother wavelet at level 1 are shown in Figure 10. The decimation operation of DWT halves the size of the coefficients discarding half of the information. The effect of

discarding the coefficients can be compared with corresponding coefficients achieved with SWT. The prominent boundary detection technique makes use of SWT with Haar. The performance comparison of other type of mother wavelet in prominent boundaries detection can be a future enhancement of the work.

3.3 CBIR Performance Measures

The precision, recall and F-measure are used to measure performance of CBIR algorithms / systems.

3.3.1 Precision

It is a measure of exactness (accuracy) in the retrieved results indicating how many retrieved images are relevant, given by a ratio of no. of retrieved relevant images over total no. of retrieved images. The maximum value 1 indicates all retrieved images are relevant.

$$p = \frac{rr}{\text{total}} \quad (3.6)$$

Where,

rr – No. of retrieved relevant images

total – Total no. of retrieved images

3.3.2 Recall

It is a measure of comprehensiveness of correctly retrieved images indicating how many relevant images are retrieved, given by a ratio of retrieved relevant images over total relevant images in the data base. The maximum value 1 indicates all relevant images in the database have been retrieved.

$$r = \frac{rr}{\text{Total}} \quad (3.7)$$

Where,

rr – No. of retrieved relevant images

Total – Total no. of relevant images in the database

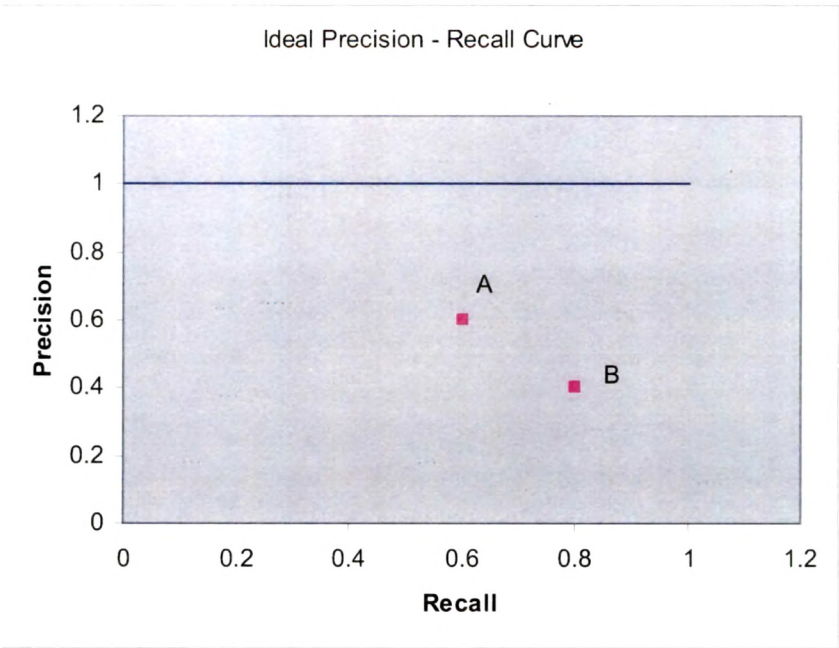


Figure 11. Ideal Precision – Recall Curve

3.3.3 P - R Curves

The precision – recall curve for a given query image is plotted for different parameters / thresholds of the retrieval algorithm. The ideal precision-recall curve is shown in Figure 11. It indicates that all retrieved images must be relevant giving precision as 1 for all values of recall. Recall value less than one indicate that some relevant images of database are missed, but all retrieved images are relevant as per ideal P – R curve. The best case scenario is precision 1 for recall 1 indicating all and only relevant images of the database have been retrieved. But, it is impossible to achieve ideal P-R curve for large image database. The precision falls non-linearly for higher values of recall in practical P – R curves. Increase in recall means an attempt to include missed relevant images. This attempt, in general, does not include only missed relevant but also includes new irrelevant images in the retrieval, decreasing the precision. Thus, better precision for lower recall and lower precision for higher recall is the characteristic of a practical P - R curves. So, precision and recall both cannot be high – ideally approaching to 1. Improvement in one measure compromises the other. The point B (0.8, 0.4) indicates that 80 % of relevant images form the database is retrieved with 40% of precision, i.e. out of total retrieved images, 60% images are irrelevant. The point A (0.6, 0.6) corresponds to better precision at the cost of recall compared to point B, i.e.

lesser irrelevant images are retrieved and missing more relevant images of the database.

3.3.4 F - measure

It is a harmonic mean, combining precision and recall to describe a single numerical value, given as [Petrakis, on line]

$$F = \frac{2}{(1/r + 1/p)} \quad (3.8)$$

Where,

r - recall

p - precision

The ideal value of precision and recall gives F measure as 1. Though F measure value does not indicate contribution of r and p, it is a single numeric value indicating performance measure of CBIR - higher the F-measure, better the performance.

The other important factor for P – R curve is the size of image database. For smaller databases, the P – R curves are likely to be better than that of larger image databases.

3.4 Watershed Regions and Issues

Watershed algorithm can be applied to segment intensity or binary images. The algorithm finds local minima forming catchment basins to determine segments of the image. The watershed function of Matlab R14 applied to segment intensity / binary images shown in Figure 12. The segmentation results are shown in the adjacent column. The segmentation results are appropriate for image of Figure 12 (a), where region boundaries are horizontal / vertical and of one pixel width. The watershed function does not segment intensity image of Figure 12 (b) appropriately. The regions surrounded by one pixel wide non-vertical and non-horizontal boundaries have not been resulted into segments despite considering 8- pixel connectivity (A Matlab-bug?). Figure 12 (c) is yet another typical image having two intensity regions surrounded by black boundaries. The watershed algorithm does not perform as expected for a region having same intensity value for all pixels belonging to the region. (?) So is the case for binary equivalent image of Figure 12 (c), shown in Figure 12 (d). The same binary image, if labeled with bwlable function, gives correct segmentation as shown in Figure 13. The

watershed algorithm implementation of Matlab R14 does produce expected segmentation for region boundaries of width greater than 1 pixel. But it is prone to erode a patch of region having same intensity values. Thus, watershed algorithm implementation of Matlab R14 adds to the challenges for achieving proper segmentation.

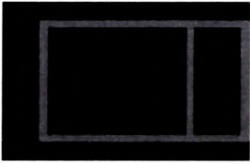







(a) Intensity Image & Corresponding Watershed Regions		
(b) Intensity Image & Corresponding Watershed Regions		
(c) Intensity Image & Corresponding Watershed Regions		
(d) Binary Image & Corresponding Watershed Regions		

Figure 12. Watershed Regions – Issues



Binary Image & Corresponding labeled regions without using watershed function		
--	---	--

Figure 13. Correctly Labeled Regions without using watershed transform

3.5 Concluding Remark

Major related technical issues have been briefly presented ...

4. Edges & Prominent Boundaries Detection

4.1 Introduction

The chapter covers proposed novel method for detecting perceptual-edges and the qualitative comparison of results with edge response of leading tools Adobe Photoshop, MS Photo Editor and ACD Photo Editor. The method results for detection of edges and thin edges, incorporating different levels of stationary Haar wavelet decompositions are compared, analyzed and presented. The method results outperform others detecting perceptually significant edges, proving its suitability for edge features extraction, object detection & identification. The thinned edges can be utilized further to reduce over-segmentation produced by watershed transformation, suggested as one of the future enhancements of the proposed work.

The later portion of the chapter includes a proposed novel method for categorizing visually prominent and non-prominent boundaries from candidate boundaries by considering prominence measures. The method results for various categories of images inclusive of standard databases [Fowlkes, on line] [Martin, 2001] [Wang, 2001] [SIMPLICity, on line] [Everingham, on line] [MedPics, on line] are presented and qualitatively compared with human segmented images of standard database [Fowlkes, on line] [Martin, 2001]. The reliable processing of low level color cues results into precise, well localized formation of prominent boundaries.

4.1.1 Key Terminologies

Contours: Contours are closed curves defining points of equal altitude (height/level). For a given channel, contours are generated by finding contour vertices (x_i, y_i) such that they form a closed curve and are at the same altitude. Here altitude for a channel under consideration refers to value of R / G / B / Gray component. For individual channel, the input matrix is treated as a regularly spaced grid, with each element connected to its all 3 neighbors forming a surface. These 4 neighbors constitute a cell.

At given height, contour vertices are found by performing a linear interpolation to locate the point at which the contour crosses the edges of the cell. Such contours at different and multiple heights are found & processed for all 4 channels. Figure 16 shows contours for respective channels.

Proximity influence: Proximity influence is a unit influence induced by a contour vertex to its nearest neighboring pixel. There can be multiple contour-vertices near a given pixel. Thus, prominence measure at a given pixel is proportionate to total proximity influence induced by all such contour vertices. Such measure is computed for all pixels of the image. E.g., let us consider two contour 1 & contour 2. Say, contour vertices A & B are on contour 1 and contour vertices P & Q are on contour 2. If M (x, y) is the nearest pixel of B; N(x-1, y-1) is the nearest pixel of say A & Q both; and O(x, y-1) is the nearest pixel of P then, B will induce proximity influence on M; A & Q will induce proximity influence on N; and similarly P will induce on O.

4.2 Block Diagram – Edges and Prominent Boundaries Detection

The block diagram for edges detection and prominent boundaries detection methods utilizing Stationary Haar wavelet based decomposition at various selected levels, contour detection at multiple levels and prominence measures is shown in Figure 14. The selection of wavelet level is performed with the help of GUI (Graphical User Interface). The level is to be selected based on image characteristics, categories, resolution and scale of segmentation. Lower level Haar results in to more number of contours / edges. E.g. Figure 20 (d) contains more edges compared to that of Figure 20 (g) detected with Haar SWT at level 1 and level 2 respectively.

After reading required input selected by user with GUI in first block, the next block performs basic operation of color channel separation. The wavelet decomposition at selected level is performed by the respective block. The contours at multiple levels are detected and processed for all four channels as explained in Section 4.1.1.

The prominence measure is utilized for edge detection and prominent boundaries detection as shown below. The block named Prominence Measure Computation of Figure 14 takes input from Contour Detection & Processing block and finds Prominence measure for all pixels for all four channels.

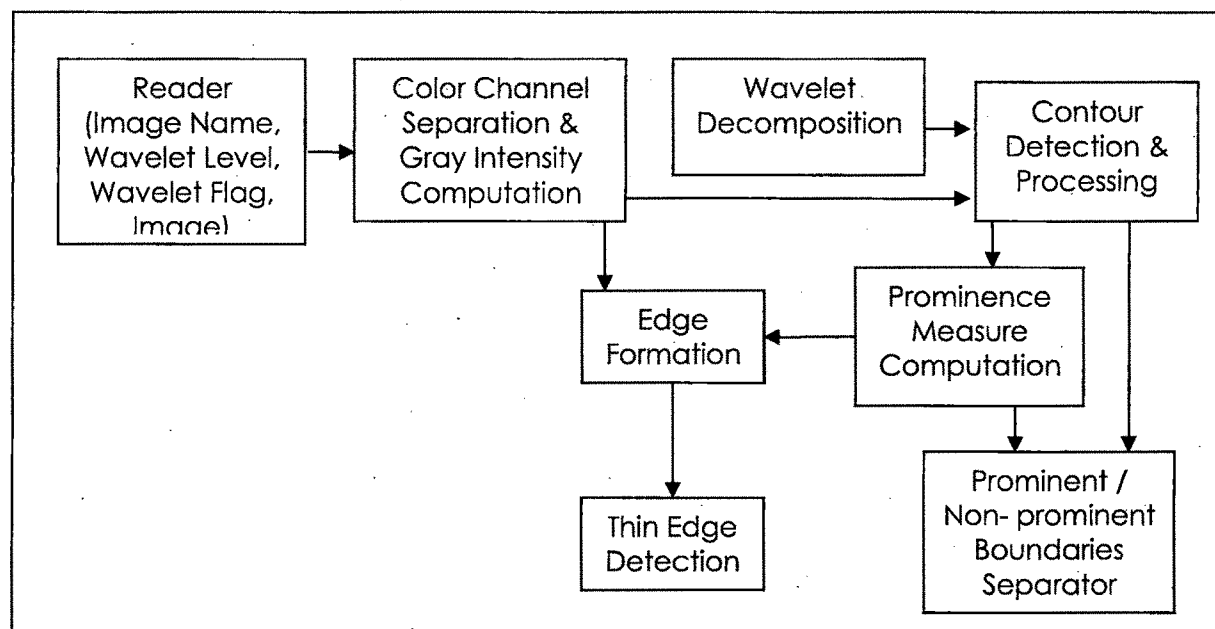


Figure 14. Block Diagram – Edge and Prominent Boundaries Detection.

4.3 Edge Detection

The local color cue based candidate boundary detection incorporating stationary Haar wavelet decomposition at various levels and thresholded prominence measures are used for detecting edges in color images in the proposed method. The non-homogeneous inter-tuples and non-uniform intra-tuple contributions of RGB tuples for conception of perceptual-edges are exploited in the edge detection method from candidate boundaries for color images. Refer Figure 16 and Figure 17 for example. The table-top boundaries in the original image are perceptually significant, as can be seen in the original image of Figure 15 (a). The intra tuple values and inter tuple values of pixels constituting these boundaries are such that the Red channel contributes the least for forming the perceptually significant edges as shown in the Figure 17 (a) left, corresponding to edges for Red channel where in table-top edges are not well defined. The Algorithm 1 exploits this characteristic of non-uniform contribution (without actually measuring it) to enforce four channel-processing to yield reliable processing.

Many state of the art techniques for image content analysis are enforcing reliable and precise processing of low level cues for extracting features, as that has been proved to be very critical for over-all performance of the applications. The edges in the image can be treated as primitive features useful for deriving other features like contours, regions & object boundaries, shapes etc. The performance of edge

detection method is challenged by image characteristics like image resolution, textures, variations in illumination etc.

The traditional gradient based edge detection techniques which examine a set of pixels for abrupt intensity changes are characterized by generation of large number of edges due to textures or color variations ending up into a difficult task of linking edges for forming boundaries by minimizing over segmentation. The state of the art techniques examine color and / or texture channels for edge detection forming boundaries / contours / regions finally leading to image segmentation. The proposed method addresses the problem in a hierarchical framework and incorporates stationary Haar wavelet decompositions at various levels, candidate boundaries and proximity influence for edge detection. The candidate boundaries and thresholded prominence measure produce edges which are not necessarily thin. These detected edges are thinned by performing a series of morphological operations of Matlab R14 on thresholded prominence measure. The local color cues used to form candidate boundaries ensures reliable processing of low level cues. The results are qualitatively compared with edge detection response of leading DTP and image processing tools for i) detection of perceptually significant edges ii) elimination of insignificant edges iii) detection of edges pertaining to region / object boundaries iv) preservation of continuities of detected edges on images of databases [Fowlkes, on line] [Wang, 2001] [SIMPLiCity, on line].

4.3.1 The Method

The multi-resolution signal decomposition as wavelet representation and the extension of orthogonal wavelet representation for images was proposed with mathematical model for computation and interpretation in [Mallat, 1989]. The classical Discrete Wavelet Transform (DWT) convolves the signal with appropriate low and high pass filters followed by decimation operation to keep generally even indexed elements and discard others by halving number of elements at each stage. The stationary Wavelet Transform (SWT) as proposed in [Nason, 1995], convolves signal with appropriate high pass and low pass filters without performing decimation operation for producing two sequences for the next level. The shortcoming of shift-invariantness of DWT is overcome in the SWT [Nason, 1995]. The proposed method makes use of stationary Haar wavelet transform at various levels [Mallat, 1989] [Nason, 1995]. Refer to

Section 3.2 for related technical details. Equations 3.1 and 3.2 describe the Haar mother wavelet function $\psi(x)$ and its scaling function $\phi(x)$ respectively.

The method uses RGB color model. As analyzed empirically, the contribution of RGB tuples for constituting boundaries is non-homogeneous inter-tuples wise and non-uniform intra-tuple wise. The candidate boundaries are closed contours of pixels forming perceptually significant and insignificant boundaries, incorporating decomposition of the image into approximate and detailed (vertical, horizontal & diagonal) coefficients by applying stationary Haar wavelet transform at various levels for RGB color and gray channels. The prominence measure based edge detection is followed by morphological operations to get thinned image of one pixel width. The steps of the proposed method for detecting edges and thinning of edges are as under. Refer Figure 14 for corresponding block diagram.

Step 1: Read Image name, wavelet decomposition level, and wavelet flag selected by a user with the help of Graphical User Interface.

Step 2: Read RGB color image $I(x, y, z)_{m \times n \times 3}$. Separate each color channel.

Compute intensity values for gray channel of the image.

Step 3: If wavelet flag is 1, resize image for height and width to make them integer power of 2 by zero padding.

Step 4: If wavelet flag is 1, apply stationary Haar wavelet transform at given level to decompose R, G B & Gray color channels, into approximate and detailed coefficients. Let us denote them as A_i , H_i , V_i and D_i as approximate coefficients, horizontal, vertical & diagonal detailed coefficients respectively at level j for given color channel c , where $0 < c < 5$.

$$Z_c^j = \{A_i, H_i, V_i, D_i\}, j > 0.$$

Step 5: Initialize prominence measure to zero

$$U(x, y) = 0$$

For First color channel and wavelet decomposed image,

Step 6: Find contours at multiple levels.

Let such set be $C_{ck} = \{(x_i, y_i)_{k}\}, i, k > 0$.

(Refer Figure 15 for the results.)

Here, k denotes index of a contour,

i denotes index of a vertex for a given contour C_k



Step 7: Merge all contours into one data structure.

Find length of each contour.

$L_{ck} = \text{length}(C_{ck})$.

Exclude contours having $L_{ck} < \text{contour_length_threshold}$.

Call remaining contours as candidate boundaries, denoted as C'_{ck} .

(Refer Figure 16 for the results.)

Here merging refers to storing of contours of all different levels into cell data structure.

Step 8: Update prominence values $U(x, y)$ at all pixels for all vertices of C'_{ck} .

Each vertex of C'_{ck} induce proximity influence to its nearest neighboring pixel.

Prominence value at given coordinate (x, y) gives total of induced proximity influence.

Step 9: Map C'_{ck} to produce binary image consisting of on pixels corresponding to the vertices of C'_{ck} . (Refer Figure 17 for the results.)

Step 10: Repeat steps 6 to 9 for all channels.

Step 11: Apply operator χ to threshold and map $U(x, y)$ on the image $I(x, y, z)$ to get edges-mapped image $I'(x, y, z)$ and binary image $BW(x, y)$, given as

$I'(x, y, z) = U(x, y) \chi I(x, y, z)$ such that

$I'(x, y, z) = I(x, y, z)$, if $U(x, y) > \text{prominence_threshold}$

and $I'(x, y, :) = \{255, 255, 255\}$, otherwise.

$BW(x, y) = 1$, if $U(x, y) > \text{prominence_threshold}$

and $BW(x, y) = 0$, otherwise.

(Refer Figure 18 (a) Right for the results.)

Step 12: Perform thinning and bridging morphological operations to get thinned image, given by

$BW(x, y) = \Lambda BW(x, y)$

Where Λ denotes thinning and bridging morphological operator. Thinning operation is to thin objects to lines by removing pixels so that an object without holes shrinks to a minimally connected stroke. Bridging operation bridges unconnected pixels, that is, sets 0-valued pixels to 1 if they have two nonzero neighbors that are not connected.

(Refer Figure 18 (b) Left for the results.)

Algorithm 1. Edge detection and thinning

The step 2 and step 3 are omitted if wave flag is set to zero and remaining steps are performed on RGB and gray channels without performing wavelet decomposition.

Thus, the method considers prominent boundaries and proximity influence induced of all four channels for edge and thin-edge detection. The method is novel for detecting edges from candidate boundaries by considering proximity influence. The approach eliminates insignificant edges and detects significant ones. Its suitability for hierarchical approach using SWT permits multi-resolution analysis required for images of different characteristics. The produced results are better than those produced with professional softwares for detecting visually significant edges.

4.3.2 Step-wise Results of the Method

The stepwise results are shown below for an image of Pascal image database [Everingham, on line].

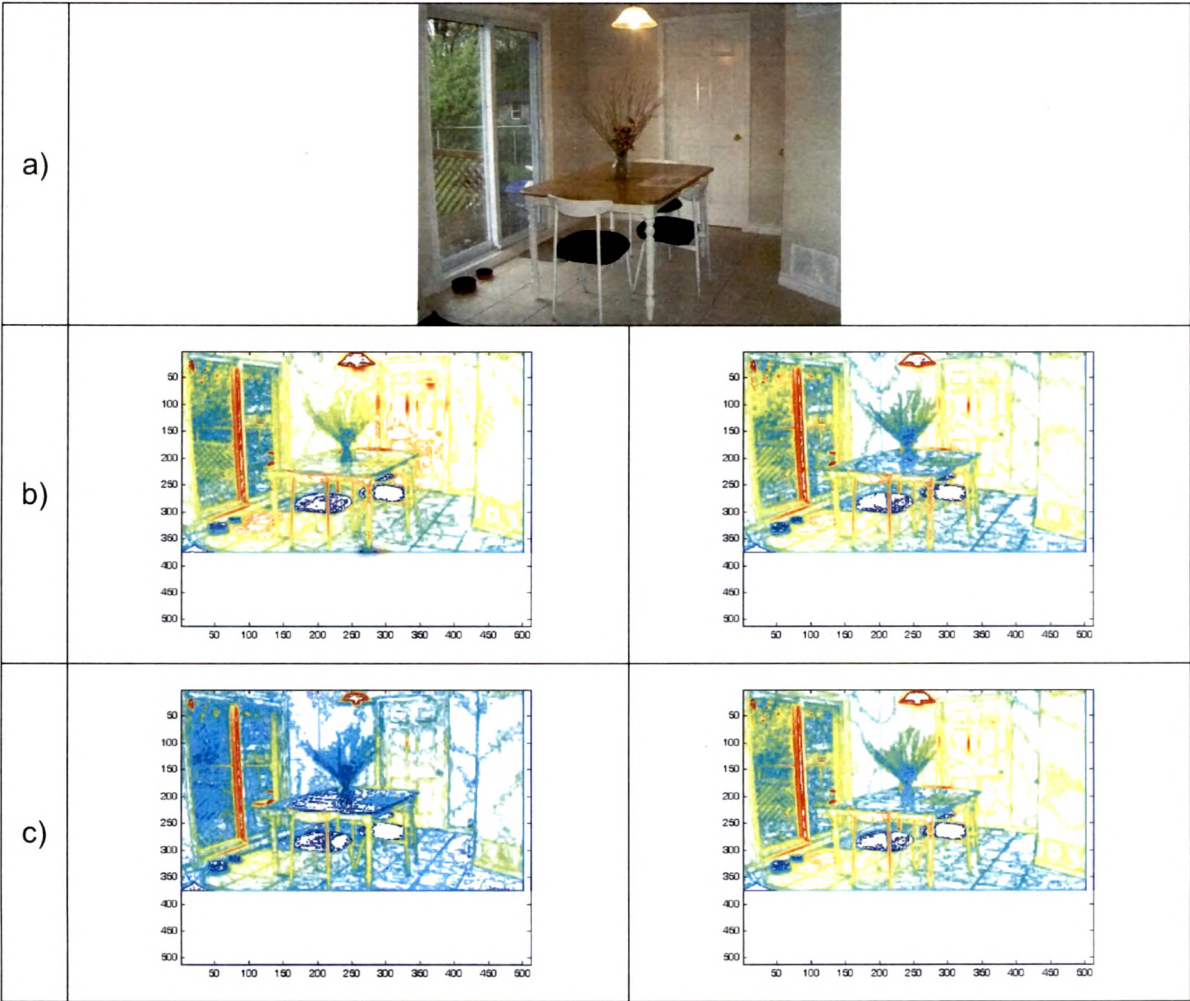


Figure 15. Contour Detection . (a) Original Image [Everingham, on line]. (b) Left, (b) Right, (c) Left, (c) Right: Contours of Red, Green, Blue and Gray channels respectively.

The image depicted to illustrate the results possesses multiple typical challenging characteristics like multiple sources of lights – point and distributed, producing illumination variations, shadows & reflection of light, different texture zones with illumination, color & color tone variations, combination of typical colors of regularly & irregularly shaped natural & man-made objects. The detected contours for RGB and gray channels are shown in Figure 15. The detected contours are large in numbers and densely placed, particularly in textured regions. The location changes of many contours produced in different channels should be noted, implying i) non-uniform contribution of different color channels for constituting prominent-real boundaries ii) one of the causes for over-segmentation.

The Figure 16 shows results of processed contours of Figure 15 obtained by eliminating very small contours produced due to textures or slight variations in intensity. The elimination of such small contours is visually apparent in the textured zones of the image.

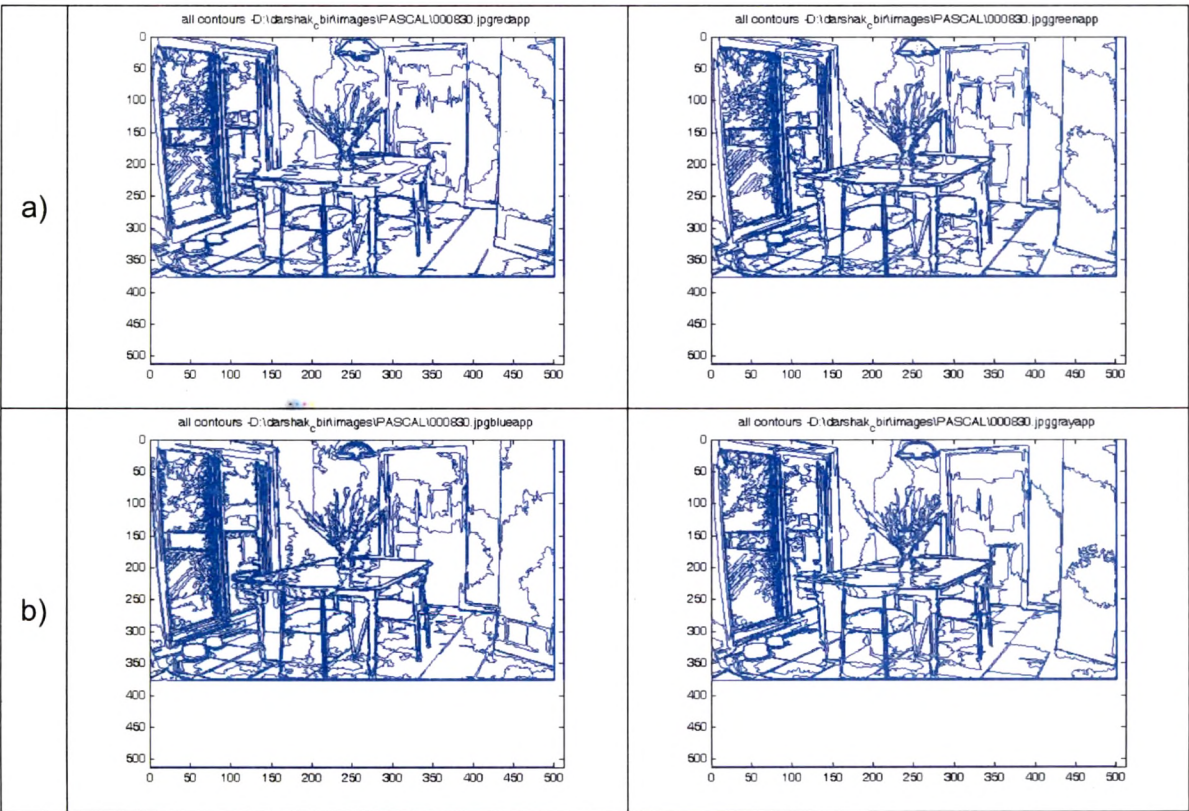


Figure 16. Processed Contours – Candidate boundaries. (a) Left, (a) Right, (b) Left, (b) Right: for Red, Green, Blue and Gray channels respectively.

The vertices of processed contours of all channels are mapped to form binary images as shown below in Figure 17. An attempt to combine these binary output images to a single image for segmentation will lead to over-segmentation of the image.

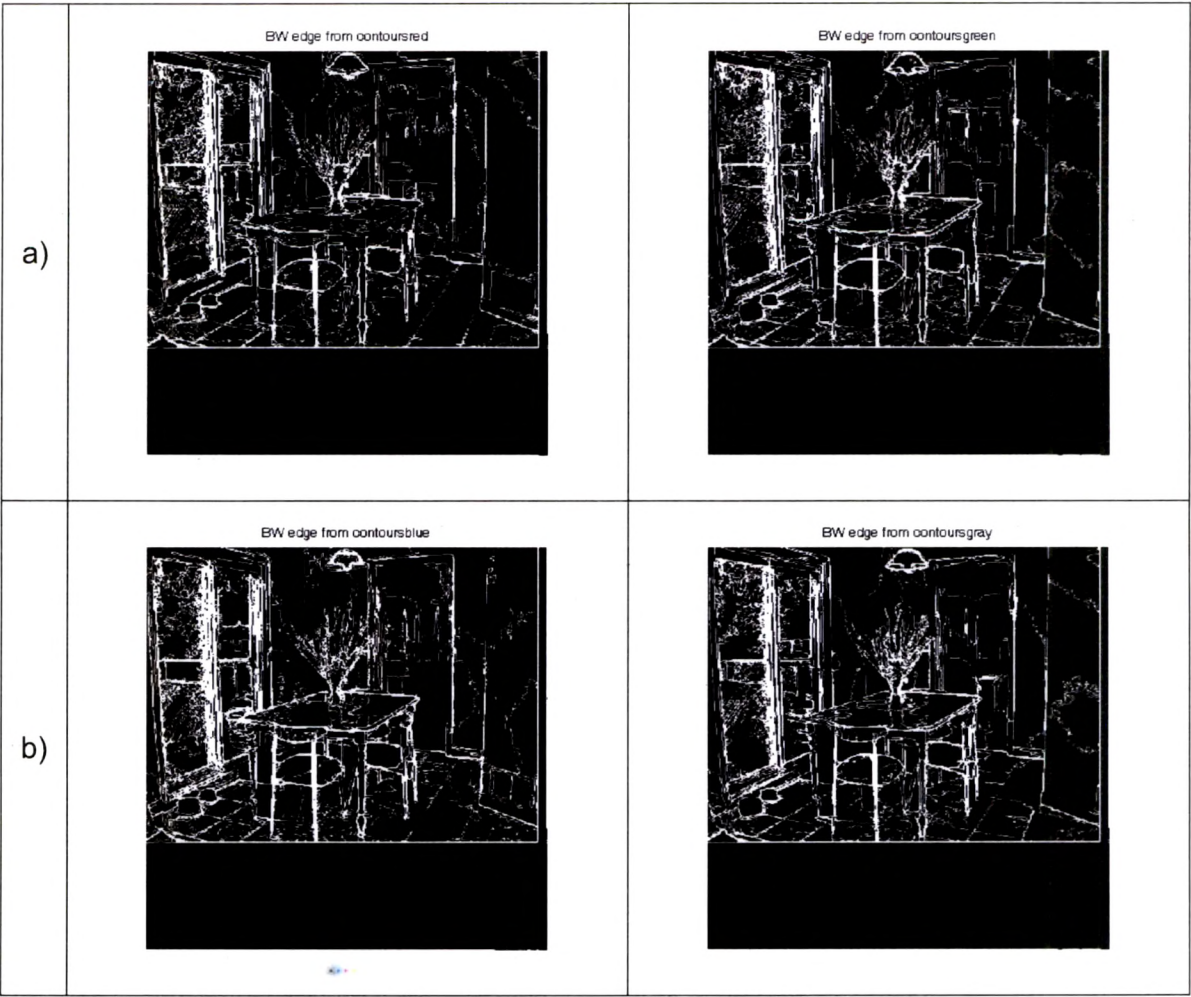


Figure 17. Binary Images: Edges from contours. (a) Left, (a) Right, (b) Left, (b) Right: for Red, Green, Blue and Gray channels respectively.

The Figure 18 (a) Left is the binary image produced by thresholding prominence measure which is mapped to image as shown in Figure 18 (a) Right. Note that the contrasts of Figure 18 (b) are altered manually for clarity in presentation. The morphological operations – thinning and bridging produces thinned image of one pixel width as shown in Figure 18 (b) Left. The Canny edge detection response on binary equivalent of thresholded prominence measure is shown in Figure 18 (b) Right for the comparison.

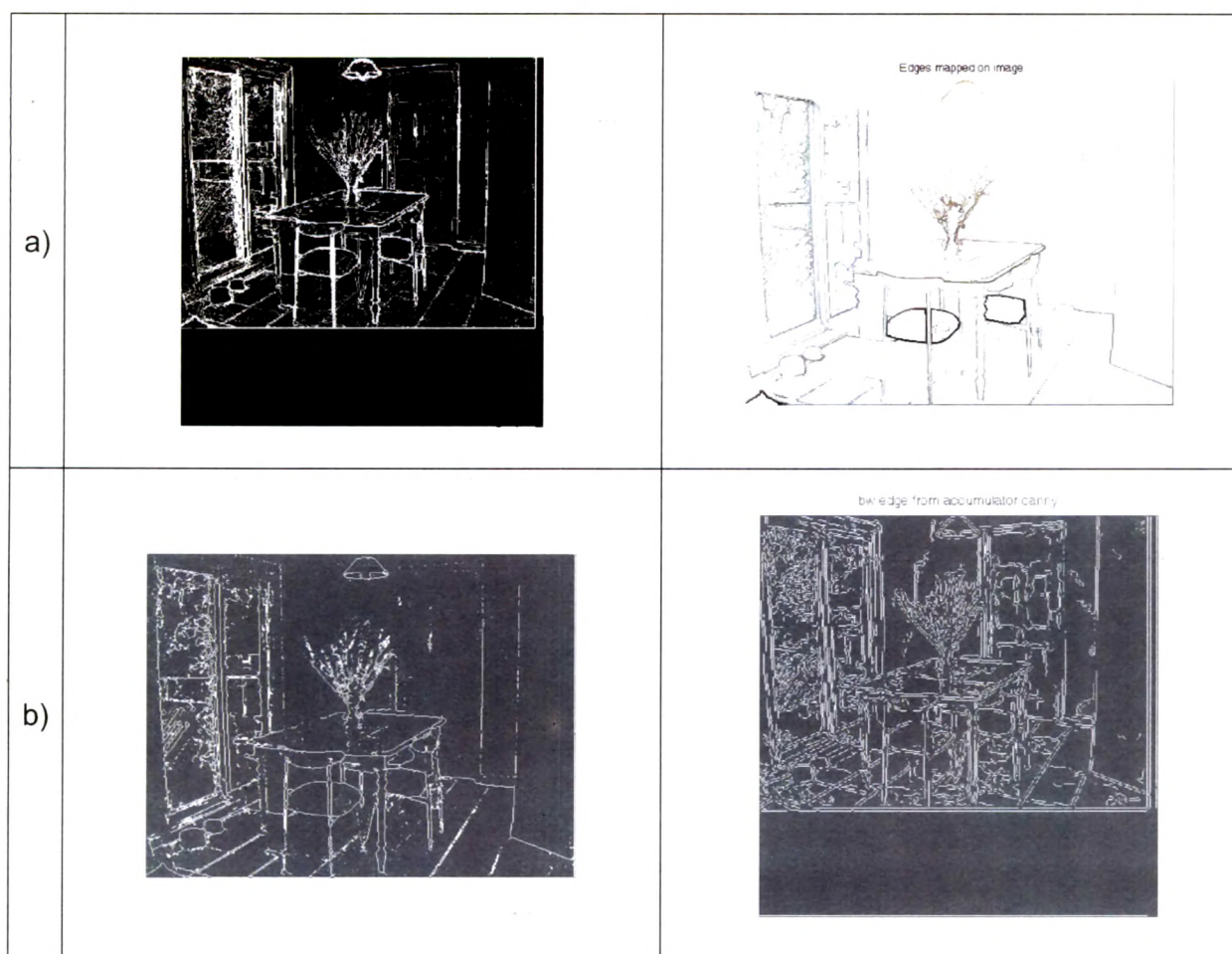


Figure 18. Edges. (a) Left: Thresholded prominence measure. (a) Right: Edges- mapped on image. (b) Left: Thinned edges. (b) Right: Canny edge detection.

4.3.3 Results – Edge Response Comparisons

The results for representative test images are shown in Figure 19 to Figure 21 for qualitative comparisons with the edge detection response of various leading software tools. The candidate boundaries of only gray channel, mapped on images are shown in all Figures. The edge and thin edge detection response take into account candidate boundaries and proximity influence of all channels. Figure 19 and Figure 20 show result-comparison for stationary Haar wavelet decompositions at levels 1 & 2 whereas Figure 21 gives qualitative comparison of results for stationary Haar wavelet decomposition at levels 2 & 3 with those of ACD Photo Editor, Adobe Photoshop and MS Photo Editor. Figure 20 (d) and Figure 20 (g) are corresponding thinned edges obtained from detected edges by applying Stationary Haar wavelet at level 1 and level 2 respectively.

Similarly, Figure 21 (d) and Figure 21 (g) are corresponding thinned edges obtained from detected edges by applying Stationary Haar wavelet at level 2 and level 3 respectively.

The salient characteristics of representative test images are listed in Table 2.

Table 2. Test Images & Their Performance Challenging Salient Characteristics.

Figures	Salient characteristics
19 (a) - Left	Low resolution natural image of SIMPLIcity [Wang, 2001] [SIMPLIcity, on line] database; Presence of two distinct background regions; one forming high contrast with the fore-ground object whereas the second forming low contrast with the fore-ground object; Inter-region illumination variations.
19 (a) Right	Low resolution natural image of SIMPLIcity [Wang, 2001] [SIMPLIcity, on line] database; Presence of significant intra-object edges; Smooth color variations.
20 (a)	Resized image – 1/10 of original high resolution image captured by an amateur; Textured back-ground; Smooth color variations in the foreground objects.
21 (a)	Higher resolution image [Fowlkes, on line] [Martin, 2001]; Presence of variety of texture zones. Presence of large number of perceptually significant as well as insignificant edges; Inter-region illumination variations.

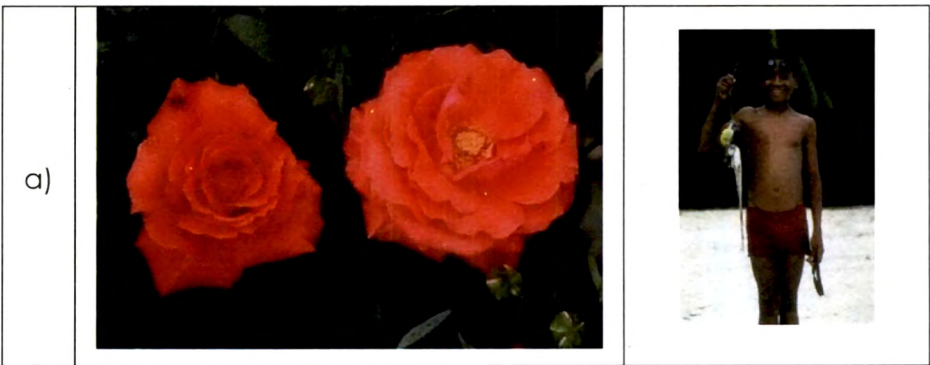


Figure 19. Edge Response Comparison. (a) Original images [Wang, 2001] [SIMPLIcity, on line].

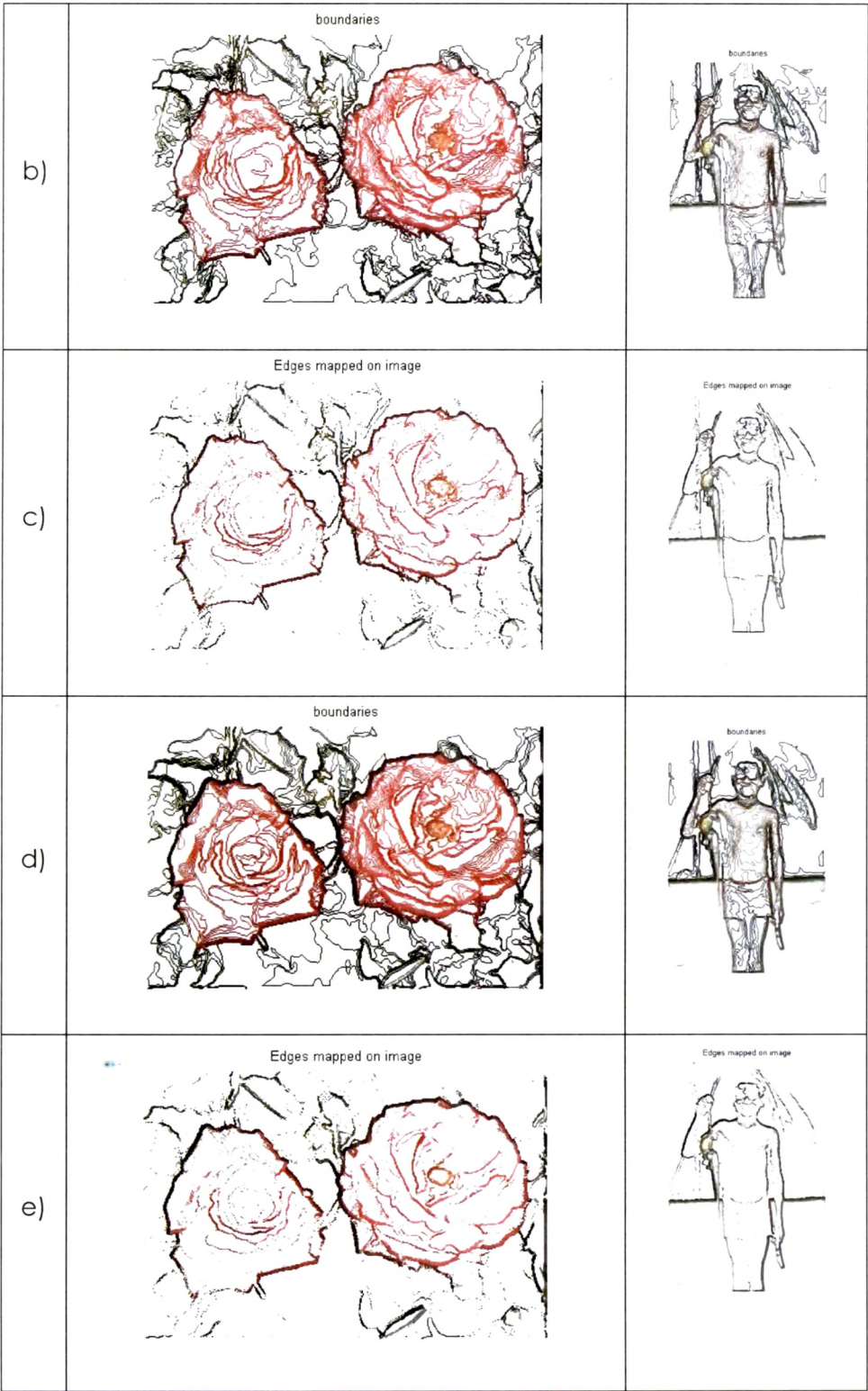


Figure 19 (Contd.). Edge Response Comparison. (b) Candidate boundaries incorporating stationary Haar decomposition at level 1. (c) Detected edges from (b). (d) Candidate boundaries incorporating stationary Haar decomposition at level 2. (e) Detected edges from (d).

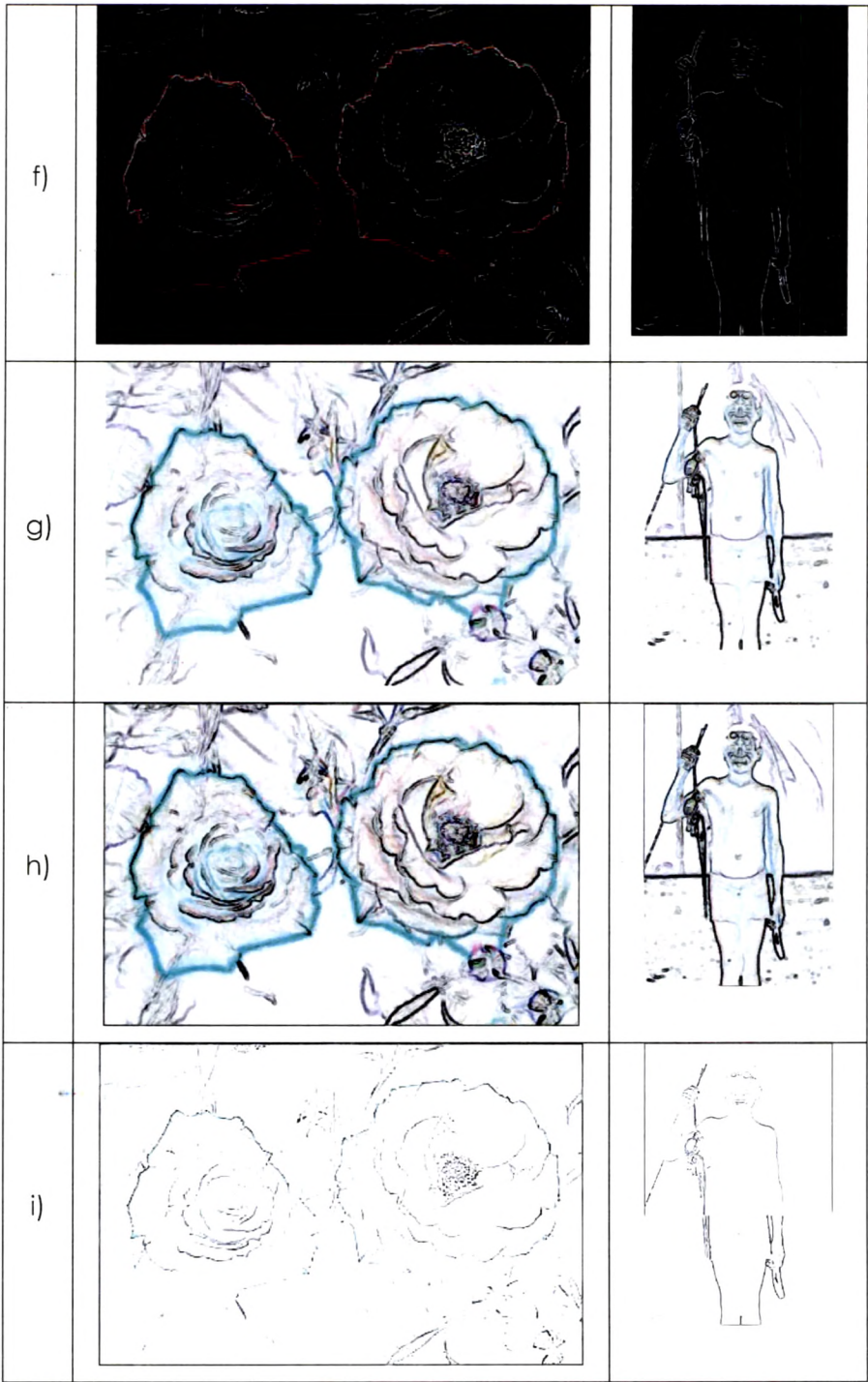


Figure 19 (Contd.). Edge Response Comparison. (f) Edge detection with ACD Photo Editor. (g) Edge detection with Adobe Photoshop. (h) Thick Edge detection with MS Photo Editor. (i) Thin Edge detection with MS Photo Editor.

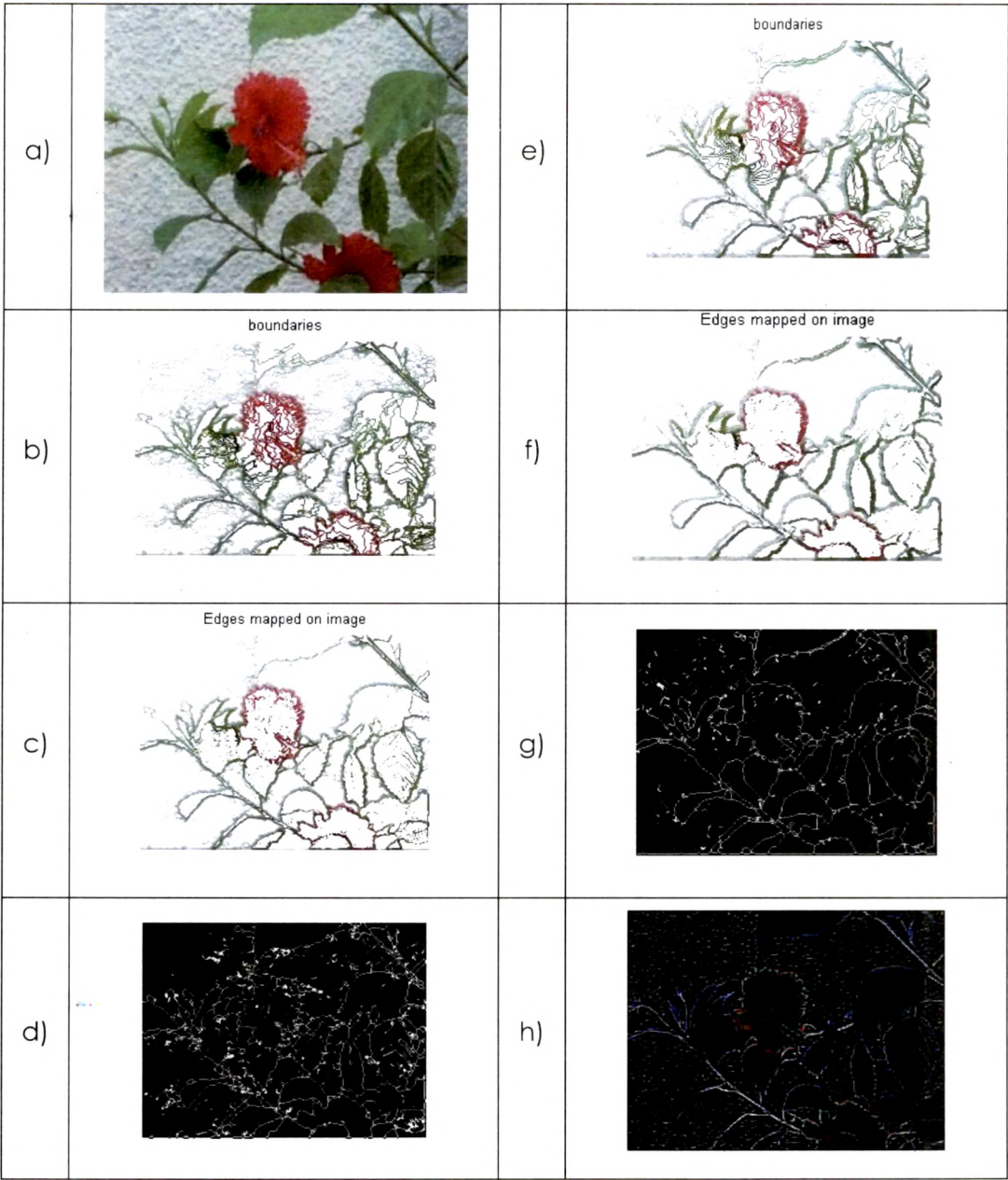


Figure 20. Edge Response Comparison. (a) Original image. (b) Candidate boundaries incorporating stationary Haar decomposition at level 1. (c) Detected edges from (b). (d) Thinned edges corresponding to (c). (e) Candidate boundaries incorporating stationary Haar decomposition at level 2. (f) Detected edges from (e). (g) Thinned edges corresponding to (f). (h) Edge detection with ACD Editor.

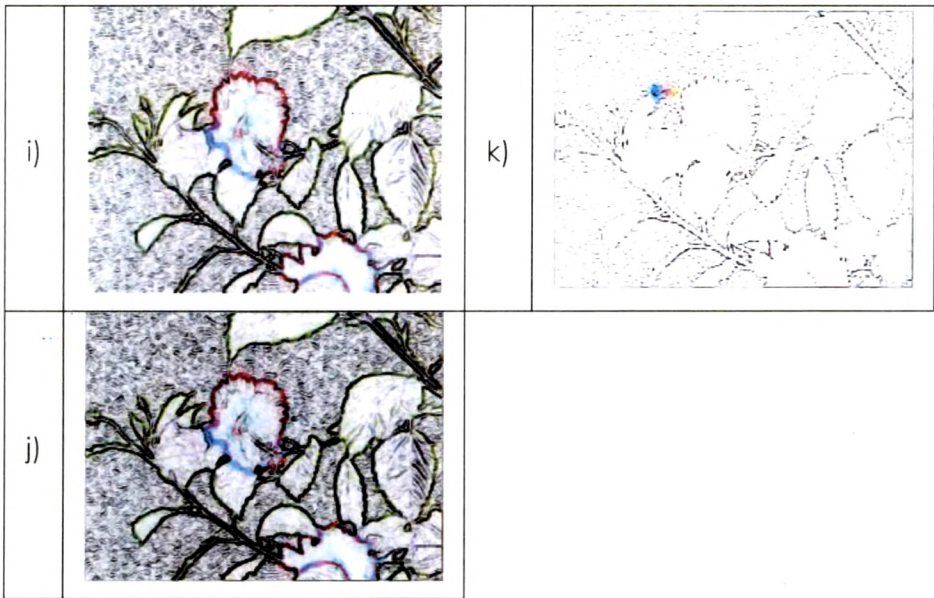


Figure 20 (Contd.). Edge Response Comparison. (i) Edge detection with Adobe Photoshop. (j) Thick Edge detection with MS Photo Editor. (k) Thin Edge detection with MS Photo Editor.

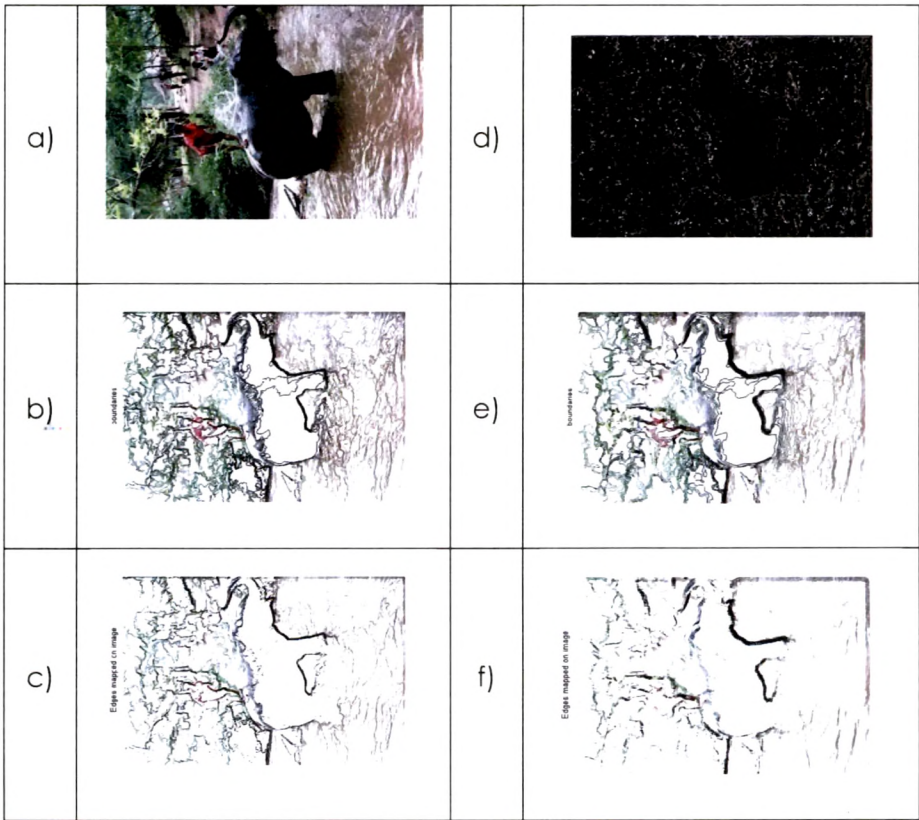


Figure 21. Edge Response Comparison. (a) Original image [Fowlkes, on line] [Martin, 2001]. (b) Candidate boundaries incorporating stationary Haar decomposition at level 2. (c) Detected edges from (b). (d) Thinned edges corresponding to (c). (e) Candidate boundaries incorporating stationary Haar decomposition at level 3. (f) Detected edges from (e).

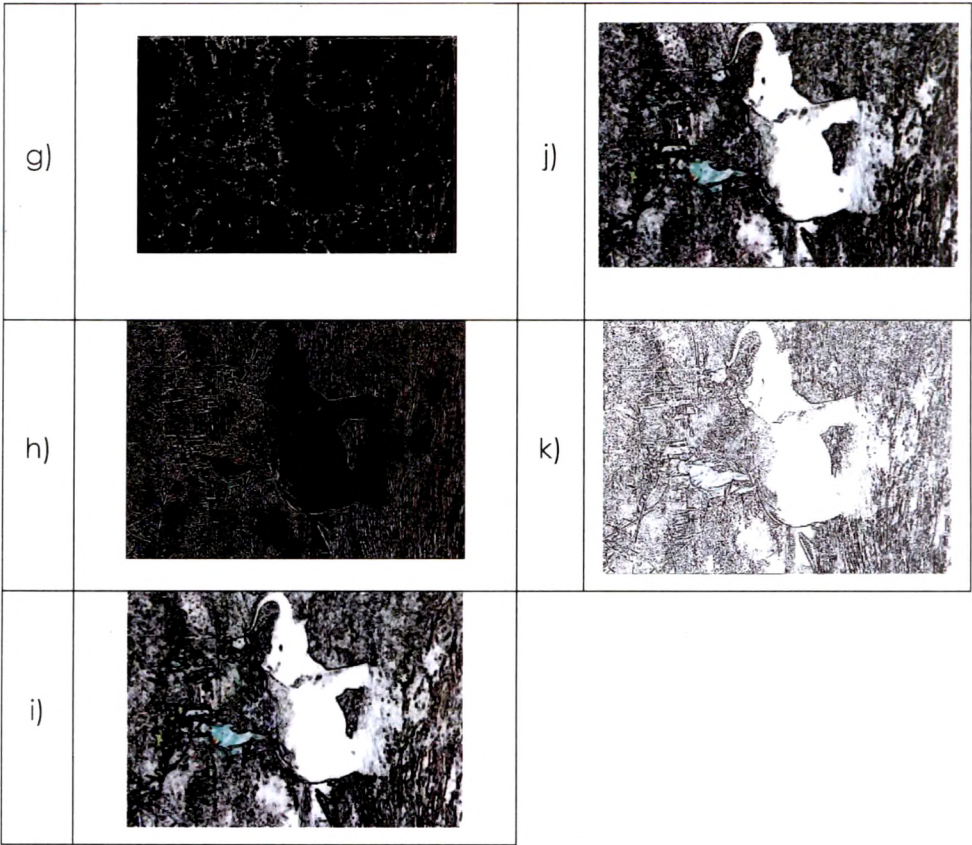


Figure 21 (Contd.). Edge Response Comparison. (g) Thinned edges corresponding to (f). (h) Edge detection with ACD Editor. (i) Edge detection with Adobe Photoshop. (j) Thick Edge detection with MS Photo Editor. (k) Thin Edge detection with MS Photo Editor.

4.3.4 Discussion

- The edge detection method based on candidate boundaries and proximity influence of all four channels detects perceptually significant edges, which are well localized and delineated.
- The detected edges are not necessarily thinned. The morphological operations are incorporated to produce thinned edges of one pixel width.
- The edges mapped on the image and thinned edges incorporating different levels of Haar wavelet decompositions are suitable as inputs to contour-detectors for producing continuity preserving closed contours and regions.
- The edge features may be used to derive shape features of objects.

- The method is well-suited for hierarchical multi-resolution approach for edge detection leading to image segmentation and object detection & identification for image content analysis and content based image retrieval.
- The selection of decomposition level of SWT decides the characteristics of the significant edges to be detected for meeting requirements of different categories of images, showing its versatility for applications.
- The detected edges and thin-edges can be incorporated to further reduce over-segmentation produced in prominent boundaries detected images, cited as future enhancement of the proposed work.
- The edge detection response analysis of the proposed method, ACD Photo Editor, MS Photo Editor and Adobe Photoshop was carried out on data set consisting of more than 300 images of different categories, resolutions and characteristics.
- The professional software packages are enough-sensitive (or over sensitive?) to detect edges constituted due to abrupt changes in color channels / textures. The sensitivity results into detection of a large number of edges, of them, many may not be visually significant. The algorithms for linking and processing of these large numbers of edges for contour generation / boundary detection / image segmentation may be not only complex but also computationally expensive.
- The results of the proposed method outperform others for i) detection of significant perceptual edges ii) elimination of insignificant edges corresponding background and foreground textures iii) better preservation of continuity.
- Quantitative analysis for comparison of edge responses of ACD Photo editor, Adobe Photoshop & MS Photo editor with proposed technique is presented in Annexure 4 showing better F – measure for proposed method.

4.4 Prominent Boundaries Detection

Image segmentation is a process of grouping region-forming pixels, satisfying single or multiple constraints on various direct or inferred cues of image attributes. Automatic segmentation of images is not only a crucial but also a challenging task. The precision-recall measures of the image segments significantly reflect the overall performance of image processing and computer vision applications involving post-

segmentation phases. Any segmentation algorithm faces the biggest challenge of avoiding over and under segmentation – subjective and image category dependent criterions. Variety of image categories, variations in required segmentation-scales, vast variations of inter & intra object textures, multiple and occluded objects, changes in the inter region illumination conditions and different image resolutions make automatic image segmentation a difficult challenge. Hence, user specified parameters, user interactions or parameters tuning are generally inevitable for better performance of any algorithm capable of segmenting variety of images.

The bottom-up segmentation approaches rely on pixel level cues or inferred cues for forming and then merging regions. So, any mistake committed at low level processing of cues, propagates without giving chance for correction, imposing stringent demand on reliability of low level cues processing mechanisms [Kass, 1988]. Similarly, as hinted in [Arbel'aez, 2009], over segmentation is a common problem across feature clustering based approaches and lack of mechanism for enforcing contour closures may cause under segmentation resulted by joining regions of leaky contours.

The proposed novel method categorizes candidate boundaries into visually-prominent and non-prominent boundaries, considering local intensity cues of multiple color channels and pixel-prominence-values measured as a function of proximity-influence. The results of the method with and without incorporating a wavelet transform are compared for images of different characteristics, types and resolutions. The method has been exhaustively tested on textured, medical, natural, biometric and synthesized images for prominent boundaries detection and the results are qualitatively compared with those of human segmented images of benchmark image-segmentation dataset. The results presented show suitability and compatibility of the method for detecting prominent boundaries in various images. These boundaries - forming regions, make the basis for object detection and identification.

The proposed bottom-up approach for prominent boundaries detection with reliable processing of low level cues that preserves non-prominent boundaries. The approach emphasizes continuity preservation and minimizes chances of contours being leaky. The separation of non-prominent boundaries eliminates visually insignificant details as far as the segmentation is concerned. Preserved non-prominent boundaries may be used for corrections of mistakes and for a multi-scale hierarchical approach of automatic image segmentation.

4.4.1 The Method

The proposed method works on RGB color model and produces prominent and non-prominent contours by classifying candidate boundaries for all three color & gray channels. Refer to Figure 14 for the block diagram. The steps of the proposed methods are as follows:

Step 1: Detect Candidate boundaries and prominence measures, as described in edge detection method proposed in [Algorithm 1, Section 4.3.1](#).

(Step 1 to Step 10 of Algorithm 1.)

Step 2: For all candidate boundaries C'_{ck} (of all four channels),

 Compute number_of_prominent_vertices and total_vertices

 Compute ratio = number_of_prominent_vertices / total_vertices

A vertex is called prominent-vertex if prominence_measure at the corresponding image coordinates is greater than threshold. The subscript c of C'_{ck} stands for the color channel under consideration and subscript k indicates contour identification number.

Step 3: Apply a classifier function λ to mark and separate prominent contours.

λ classifies contours based on contour length and ratio.

If ratio computed in above step is greater than 0.5 and a contour consists of more than 5 vertices (length), mark the contour as prominent and otherwise as non-prominent. Let P_c and N_c be a set of prominent and non-prominent contours respectively, containing classified contours.

Algorithm 2. Prominent boundaries detection.

The feature-preserving non-prominent boundaries are maintained in N_c is a salient characteristic of the method. Though N_c is not utilized at present, its usage may be explored.

The Figure 22 shows separated prominent and non-prominent boundaries for all color channels. As observed, the contribution of different color channels for constituting real, prominent boundaries is not homogeneous. Processing of any one channel may not detect a portion of real boundaries leading to under-segmentation of the image. The method not only provides reliable processing, but also takes advantage of redundancy of cues for precise, well localized detection of prominent boundaries.

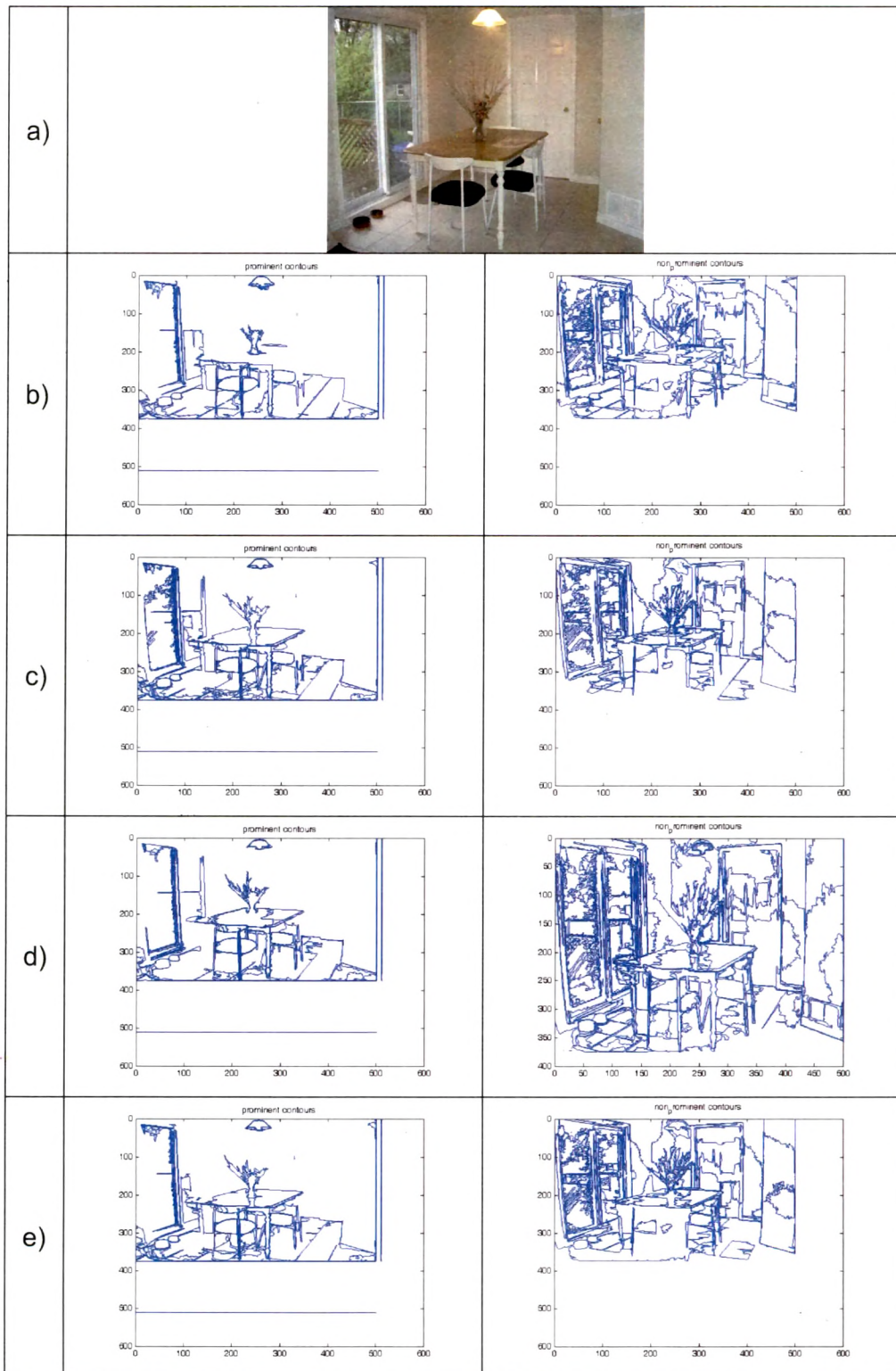


Figure 22. Prominent & non-prominent boundaries. (a) Original Image [Everingham, on line]. (b) Left, (c) Left, (d) Left, (e) Left: prominent boundaries of Red, Green, Blue and Gray channels respectively. (b) Right, (c) Right, (d) Right, (e) Right: non-prominent boundaries of Red, Green, Blue and Gray channels respectively.

4.4.2 Results

The salient characteristics of some of the images of test-set are tabulated below.

Table 3. Categorical Representative Test Images & Characteristics.

Description and salient characteristics of images	
Figure 23, left	The BSDb test image [Fowlkes, on line] [Martin, 2001], multiple objects, textured background, natural image.
Figure 23, right	The BSDb test image [Fowlkes, on line] [Martin, 2001], multiple objects covering significant portion of the image, objects touching image boundaries, natural image.
Figure 24, left	A standard test image.
Figure 24, right	The BSDb test image [Fowlkes, on line] [Martin, 2001], single central main object, captured with background softening filter, homogeneous background.
Figure 25, left	A synthesized image, smooth variations of color tones, light reflections and alphabets of typical colors & type on typical background.
Figure 25, middle	A close-up, resized to 1/8 th of original, image captured with high resolution device having inbuilt image processor.
Figure 25, right	Single central main object [Wang, 2001] [SIMPLicity, on line], typical background.
Figure 26, left	An image of regular, repeated shapes [Wang, 2001] [SIMPLicity, on line].
Figure 26, right	A biometric image - a finger-print.
Figure 27, left	The BSDb test image [Fowlkes, on line] [Martin, 2001], multiple similar objects, occluded objects in the object group, partially textured background, natural image.
Figure 27, right	An image of man-made object [Wang, 2001] [SIMPLicity, on line].
Figure 28 left, right	The BSDb test images [Fowlkes, on line] [Martin, 2001].
Figure 29 left, right	The PASCAL challenge, 2008, images [Everingham, on line].
Figure 30, left	A medical image with less prominent boundaries [MedPics, on line].
Figure 30, right	The BSDb test image [Fowlkes, on line] [Martin, 2001], single central main object, reflections & whirls in the water.

The prominent boundaries and non-prominent boundaries of gray channel, mapped on images are presented here in Figure 23 to Figure 30 for various categorical test images for the qualitative comparisons of detected prominent boundaries with those of human segmented images of BSDB [Fowlkes, on line] [Martin, 2001]. The prominent boundaries detection results with Stationary Haar wavelet decomposition at various levels are shown for the comparison. Figure 28 (c) shows detected prominent boundaries without incorporating Stationary Haar wavelet decomposition.

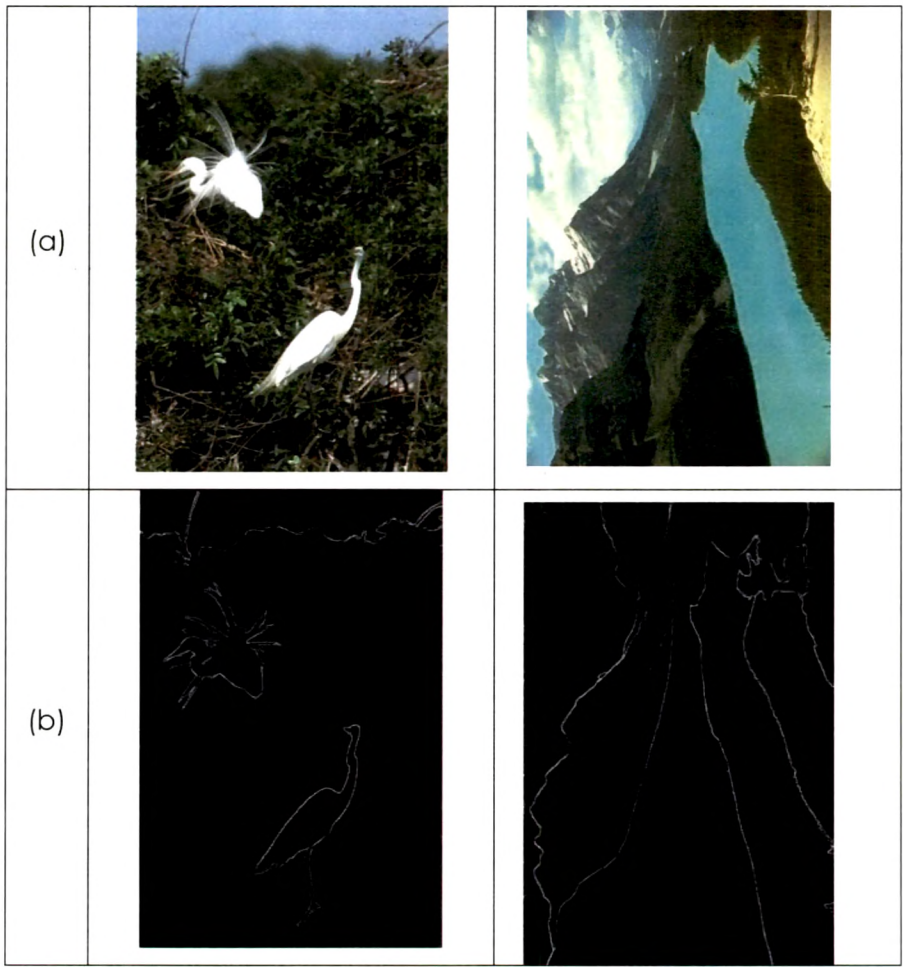


Figure 23. Comparison of human segmented image results with detected prominent boundaries. (a) Original image BSDB [Fowlkes, on line] [Martin, 2001]. (b) Human segmented image BSDB [Fowlkes, on line] [Martin, 2001].

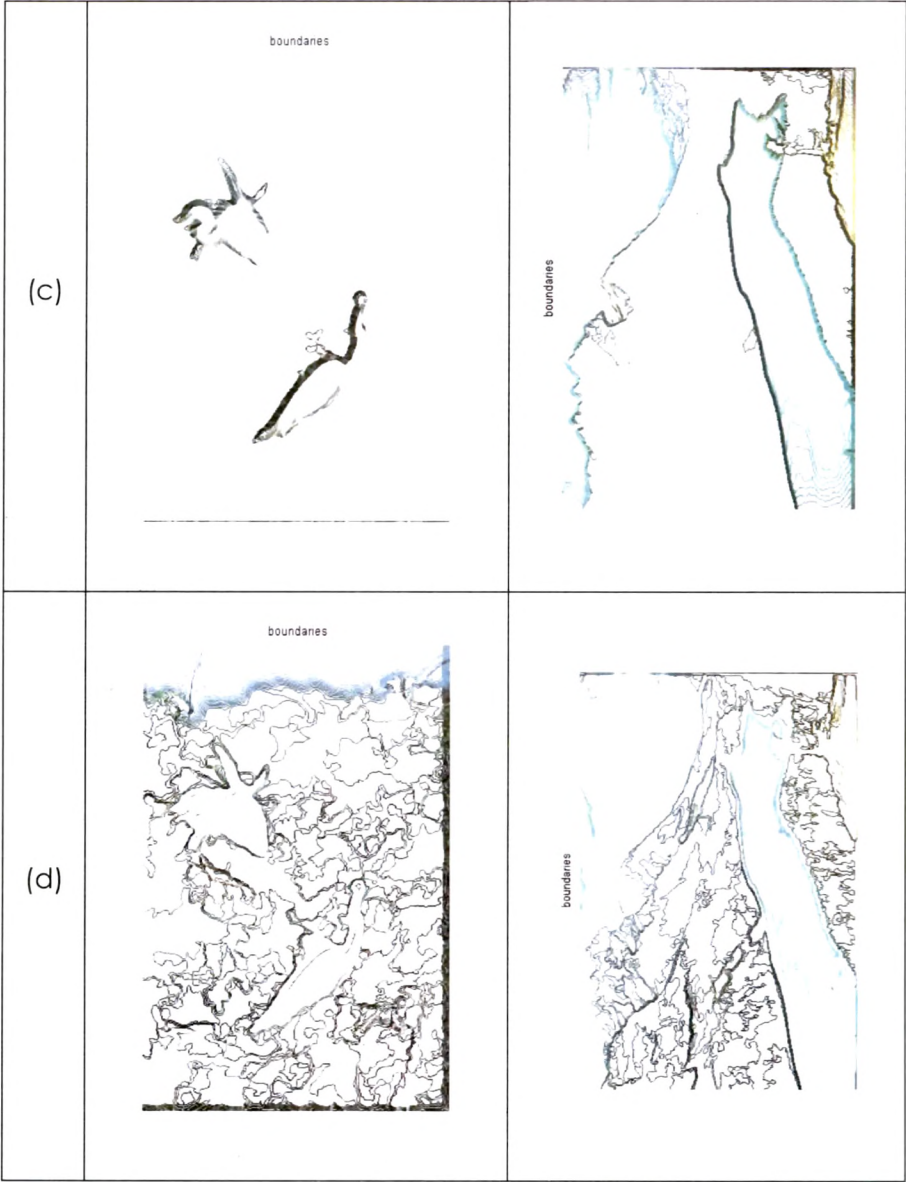


Figure 23 (Contd.). Comparison of human segmented image results with detected prominent boundaries. (c) Prominent boundaries incorporating Stationary Haar decomposition at level 2. (d) Non-prominent boundaries.

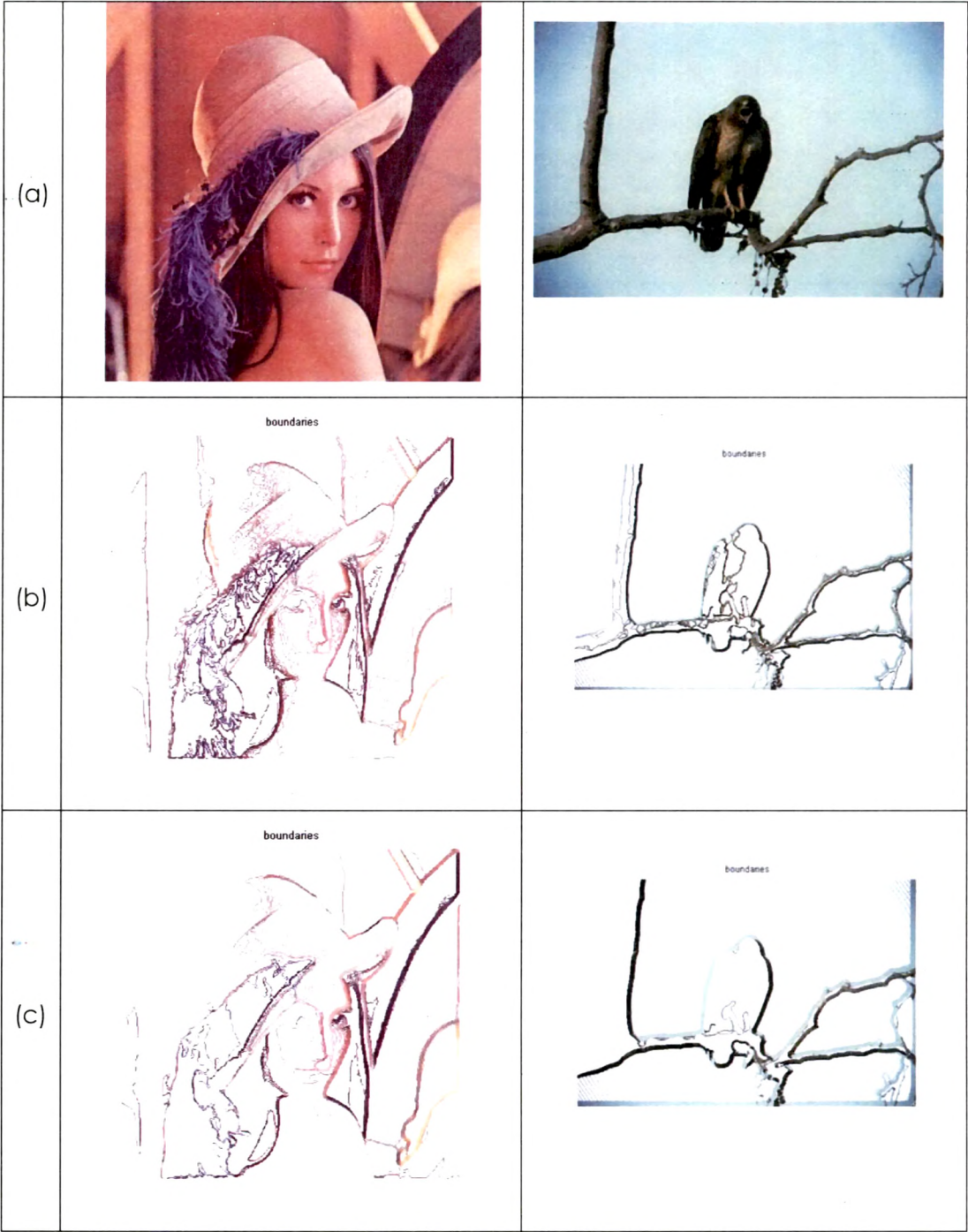


Figure 24. Prominent boundaries detection, incorporating different levels of Stationary Haar decompositions. (a) Left - Original image. (a) Right - Original Image BSDB [Fowlkes, on line] [Martin, 2001]. (b) Prominent boundaries incorporating Stationary Haar decomposition at level 2. (c) Prominent boundaries incorporating Stationary Haar decomposition at level 3.

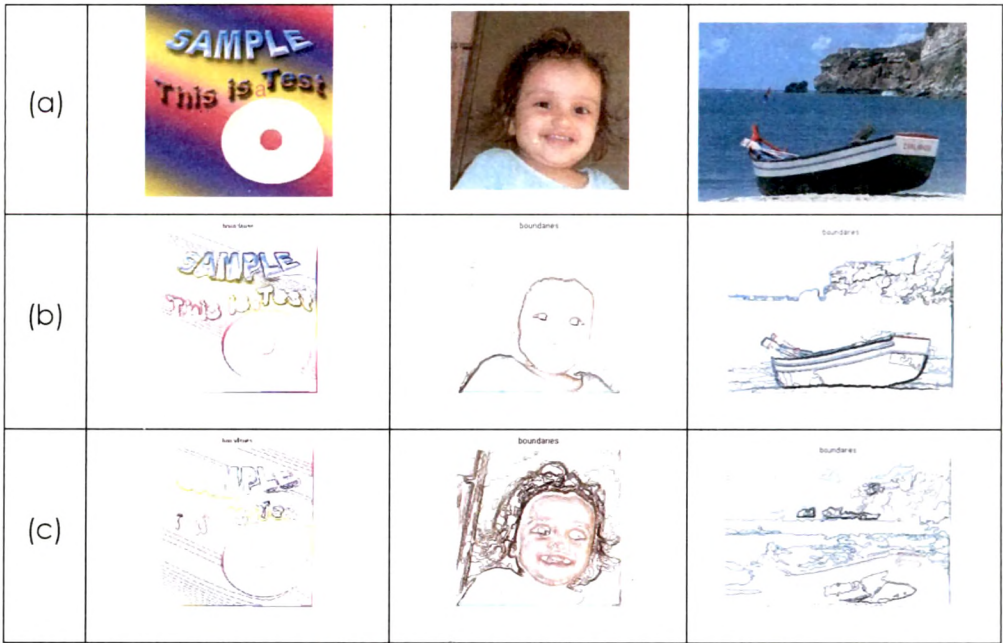


Figure 25. Prominent boundaries detection results - Different categorical images. (a) Left & Middle - Original images. (a) Right - Original image [Wang, 2001] [SIMPLIcity, on line]. (b) Prominent boundaries incorporating Stationary Haar decomposition at level 2. (c) Non-prominent boundaries.

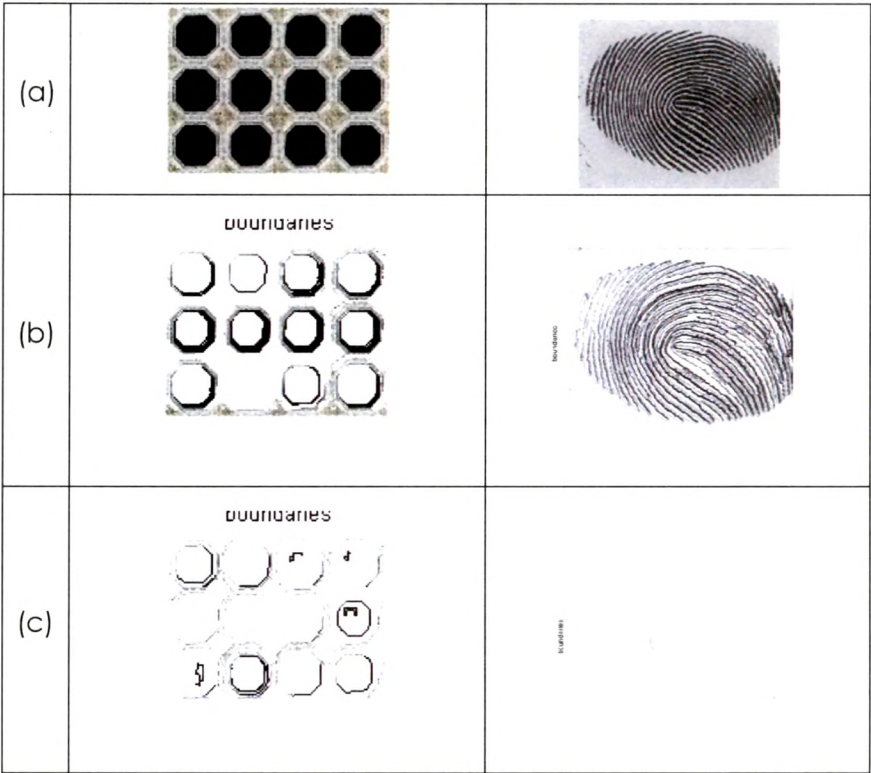


Figure 26. Prominent boundaries detection results - Different categorical images. (a) Left - Original image [Wang, 2001] [SIMPLIcity, on line]. (a) Right - Original image. (b) Prominent boundaries incorporating Stationary Haar decomposition at level 2. (c) Non-prominent boundaries.

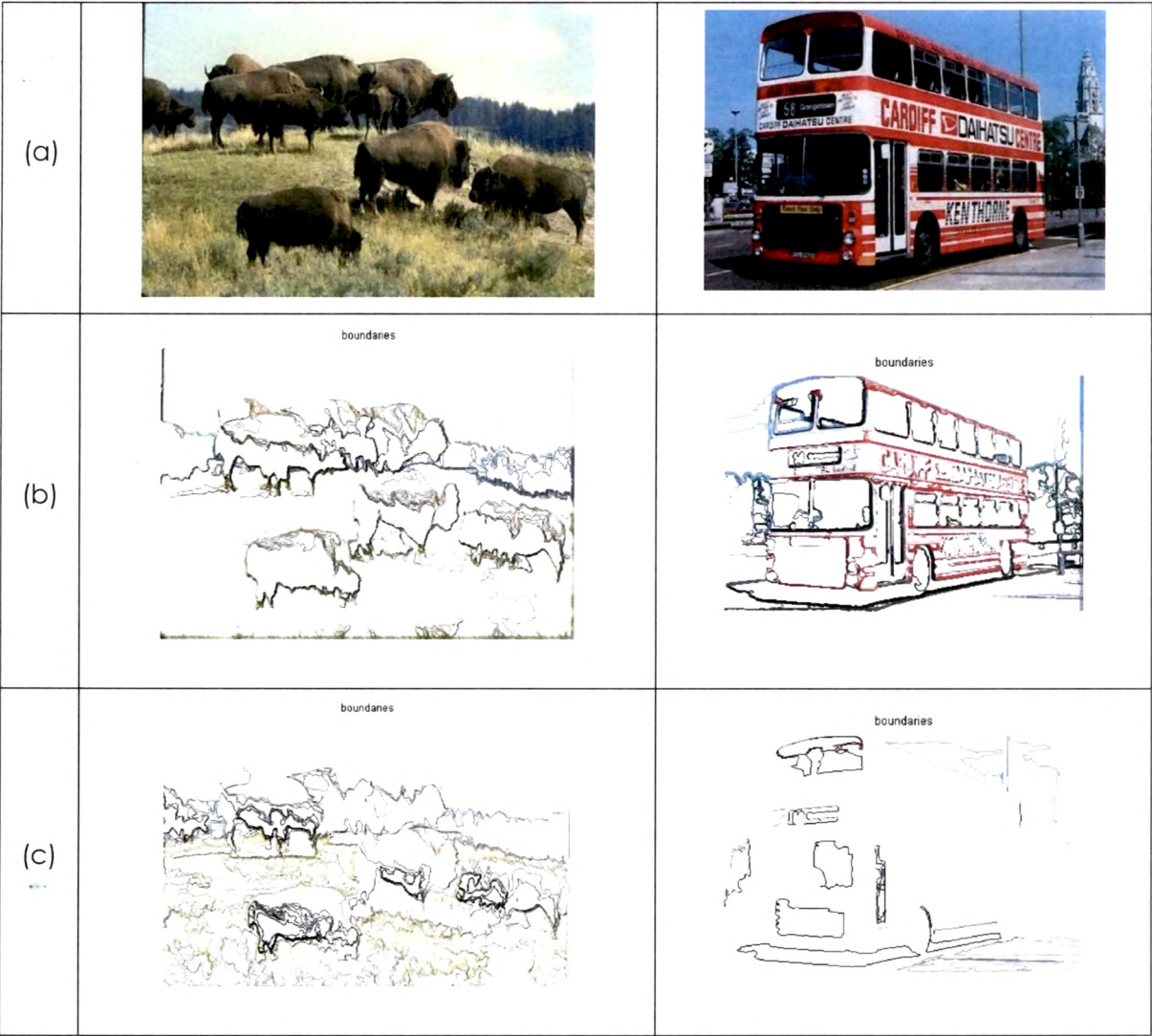


Figure 27. Prominent boundaries detection results - Different categorical images. (a) Left - Original image [Fowlkes, on line] [Martin, 2001]. (a) Right - Original image [Wang, 2001] [SIMPLIcity, on line]. (b) Prominent boundaries incorporating Stationary Haar decomposition at level 2. (c) Non-prominent boundaries.

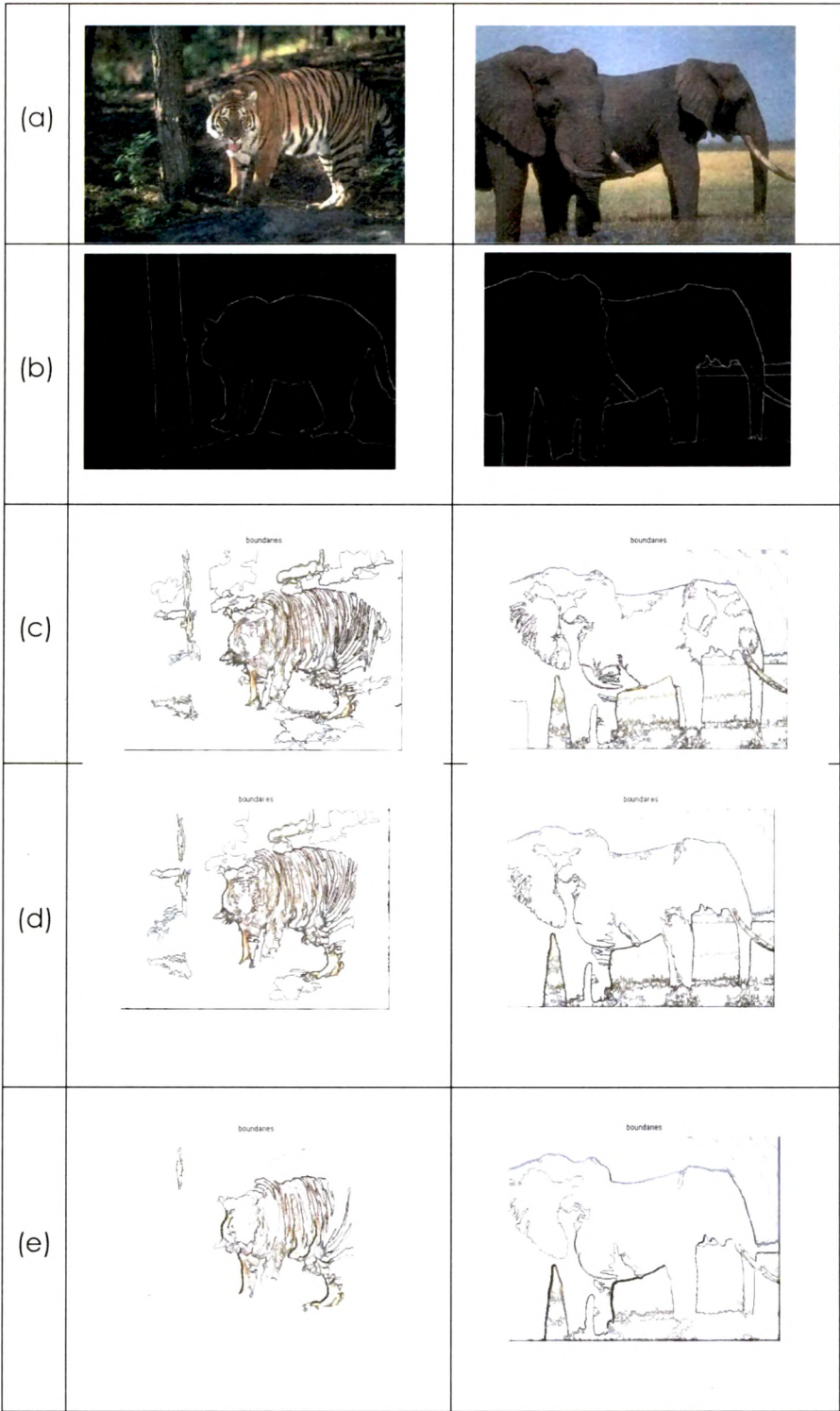


Figure 28. Comparison of human segmented image results with detected prominent boundaries. (a) Original image BSDB [Fowlkes, on line] [Martin, 2001]. (b) Human segmented image BSDB [Fowlkes, on line] [Martin, 2001]. (c) Prominent boundaries without incorporating Stationary Haar decomposition. (d) Prominent boundaries incorporating Stationary Haar decomposition at level 1. (e) Prominent boundaries incorporating Stationary Haar decomposition at level 2.

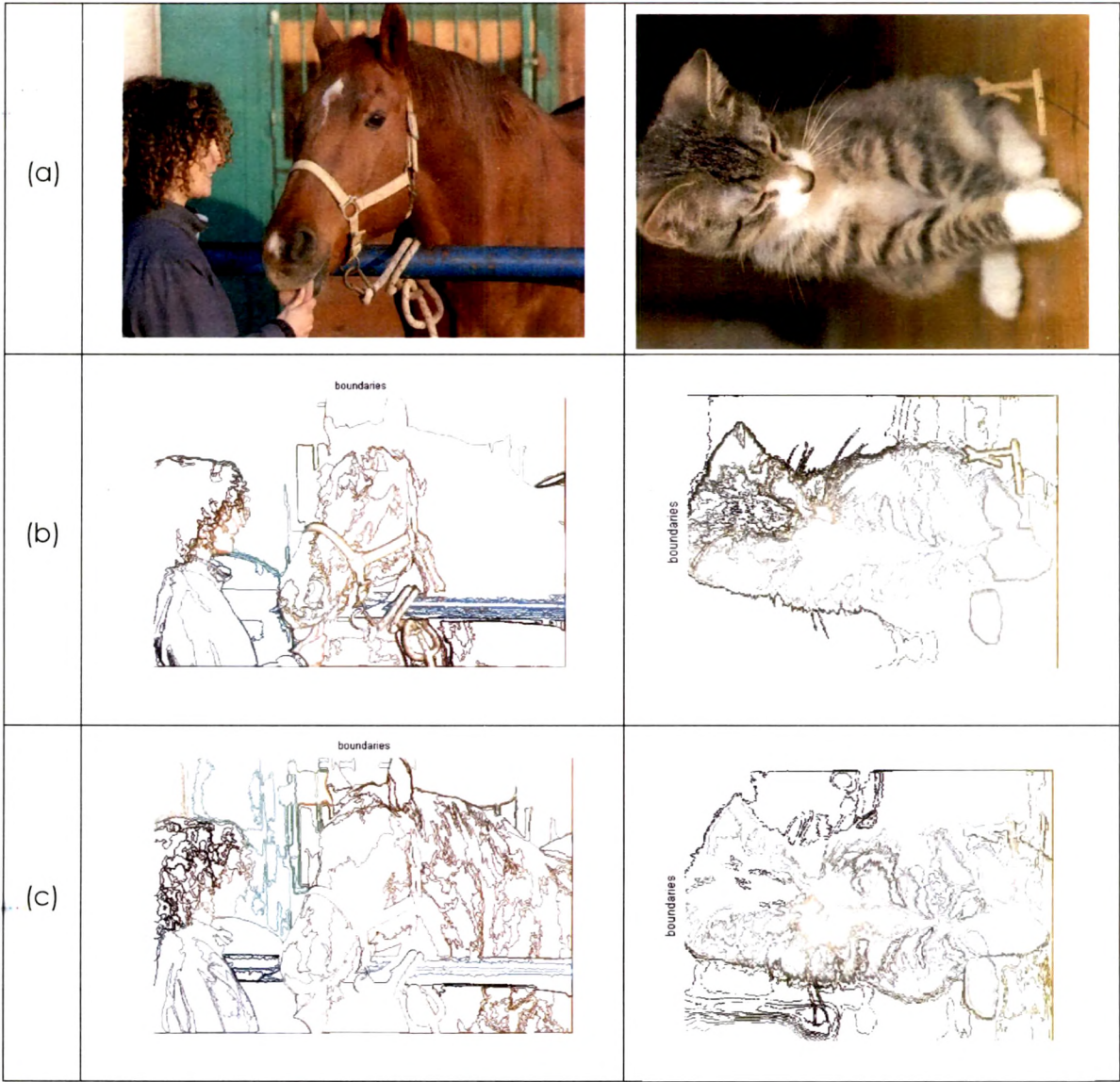


Figure 29. Prominent boundaries detection results - Different categorical images of PASCAL challenge 2008. (a) Original images [Everingham, on line]. (b) Prominent boundaries incorporating Stationary Haar decomposition at level 1. (c) Non-prominent boundaries.



Figure 30. Prominent boundaries detection results - Different categorical images. (a) Left - Original image [MedPics, on line]. (a) Right - Original image BSDb [Fowlkes, on line] [Martin, 2001]. (b) Detected prominent boundaries incorporating Stationary Haar decomposition at level 1. (c) Non-prominent boundaries.

4.4.3 Discussion

- The detected prominent boundaries are well-localized and well-delineated.
- The processing of low level cues generating multiple candidate boundaries enforces reliability for categorizing prominent boundaries with continuity preservation.
- Categorization resulting into exclusion of non-prominent boundaries separates insignificant features from significant ones (from segmentation point-of-view).
- The Stationary Haar wavelet decomposition and detected prominent boundaries at various levels make the approach suitable for multi-scale hierarchical image segmentation.

- o The method produces significantly comparable results for a test-set consisting of more than 300 images, covering representative images of different categories, possessing performance challenging, wide variety of salient characteristics (as presented in Table 3).
- o The prominent boundaries detection results are effective for Berkeley Segmentation Dataset images [Fowlkes, on line] [Martin, 2001], PASCAL challenge 2008 database images [Everingham, on line] and SIMPLicity database images [Wang, 2001] [SIMPLicity, on line].
- o The development of effective methodology for identifying regions pertaining to objects make the proposed wavelet based hierarchical method suitable for applications like object detection, object identification, automatic image tagging, content based image retrieval and visual scene analysis.

4.5 Concluding Remark

The continuity preserving well localized prominent boundaries form the basis for segmentation, feature extraction and foreground separation for CBIR ...

5. Foreground Objects Detection & Background Separation

5.1 Introduction

The chapter covers a novel method of background separation, revealing foreground objects for color images. The approach effectively incorporates visually prominent boundaries, prominence measure of pixels and local color & region cues along with Watershed transform. The proposed method addresses the issues for avoiding under segmentation and over segmentation by enforced reliable processing of local cues used for producing continuity preserving prominent boundaries. The method results, incorporating different levels of stationary Haar wavelet decompositions are compared, analyzed and presented. The effectiveness, suitability and versatility of the method for well localization of prominent boundaries leading to foreground extraction are shown by qualitative comparisons of method results with that of human segmented images of benchmark-image-dataset [Fowlkes, on line] [Martin, 2001]. In the last portion of the chapter, segmentation results of JSEG [Deng, on line] [Deng, 2001] are qualitatively compared with the results of proposed method for observing effects of illumination changes and texture variations.

The reliable and precise processing of low level cues in pixel domain is very important and crucial for over all performance of any image processing applications handing out image features derived or inferred from these low level cues. The precise and reliable processing of low level cues for feature extraction is tested by various factors like image resolutions, intra-image illumination variations, non-homogeneity of intra-region and inter-region textures, multiple and occluded objects etc. The foreground objects revealing by separating the background is, in general, a subsequent phase of image segmentation. Image segmentation is a process of identifying and then grouping region-forming pixels, satisfying single or multiple constraints. The diversities of image characteristics enforce requirements of parameter selection and / or parameter tuning or user interaction for better performance of

generic segmentation algorithms. The performance of the segmentation algorithm is hence evaluated with Precision-Recall measures for different segmentation scales. So, for a given scale, any segmentation algorithm faces the biggest challenge of avoiding over and under segmentations – subjective and image category dependent criterions. Thus, the subjectivity in the human perception for segmentation, required scale of segmentation and diversified image characteristics play decisive role for justifying proper segmentation of the images. The foreground objects detection is severely affected by a lapse generated during segmentation phase.

Similarly, difficulties do exist with watershed algorithms. The watershed algorithms find catchments basins by locating local minima, resulting into artifacts and over segmentation. The conversion of color image into gray scale image produces local minima leading to artifacts and conversion of gray scale image into binary introduces breaks likely to cause under segmentations. As illustrated in top-right of Figure 33, the watershed regions of gray scale converted image are large in numbers and small in sizes, producing over segmented image. Similarly, the watershed regions obtained from dithered image exhibits relatively lesser number of segments due to lossy conversion, giving still few over segmented zones in the image, as shown in bottom-right of Figure 33. Thus, for distinction of objects using resulted watershed regions, additional low level cues must be incorporated for selective merging of these regions. Hence, marking of well localized object boundaries becomes the necessary condition for proper object detection.

The proposed approach is based on precisely detected prominent boundaries with continuity preservation for minimizing chances of them being leaky. The problem of foreground objects revealing of prominent boundaries detected images has been addressed in the hierarchical frame work incorporating wavelet decomposed images at various levels along with proximity influence measure and watershed transform.

5.2 The Method

The method exploits stationary Haar wavelet for decomposing RGB images at various levels. The well localized, visually prominent, continuous boundaries encompassing various regions called prominent boundaries are detected as proposed in [Algorithm 2, Section 4.4.1](#). The prominent boundaries are categorized candidate boundaries. The non-prominent boundaries resulted because of smooth variations in textures and colors are excluded because of the categorization. The proposed method

is novel for its prominent boundary features, proximity influence measures and its usage for watershed regions. Many variational techniques to reduce watershed artifacts were tried out for development of proposed method. The watershed transform of Matlab R14, a modification of basic watershed algorithm of [Vincent, 1991] has been utilized in following algorithm with 8 neighbor connectivity.

The steps of the proposed method are:

Step 1: Detect prominent boundaries by applying proposed method of [Algorithm 2, Section 4.4.1](#).

Denote the set of prominent contours as $P_c = \{V_i\}$, $i > 0$, where V_i is a vector, corresponding to i^{th} contour of color channel c consisting of coordinate-pairs denoted as $\{(x_j, y_j)\}$, $j > 0$.

Apply operator χ to map P_c on the image $I(x, y, z)$ to get prominent-boundaries-mapped image, given as

$$I'(x, y, z) = P_c \chi I(x, y, z) \text{ such that}$$

$$I'(x, y, z) = I(x, y, z), \text{ if } x, y \in P_c$$

$$\text{and } I'(x, y, :) = \{255, 255, 255\}, \text{ otherwise.}$$

Step 2: For all pixels on prominent boundaries, compute total proximity influence value i.e. prominence measure induced by vertices of prominent contours of channel under considerations on nearest neighboring prominent pixels. Refer Section 4.1.1 for prominence measure computation. Denote it as $UP_c(x, y)$. At the given point, higher the value stronger is the boundary strength. Refer Figure 31 (c) for the results.

Step 3: Perform watershed transformation on $UP_c(x, y)$. Refer Figure 31 (d) for the results.

Step 4: Perform morphological operation on $UP_c(x, y)$ to get $UP'_c(x, y)$. Refer Figure 31 (e) for the results.

Step 5: Perform watershed transformation on $UP'_c(x, y)$.

Label watershed region. Refer Figure 31 (f) for the results.

Step 6: Repeat previous steps for all channels.

Step 7: Find composite prominence measure UP'' and watershed transform. Label watershed regions. Refer Figure 32 (a) for the results.

Composite prominence measure is total of prominence measure of all 4 channels.

Step 8: Let $SP = \{(x_i, y_i)\}$ such that $x_i, y_i \notin P_c$, a fixed set of seed points determined empirically.

$$BG(x, y) = ((I'(x, y, z), UP''(x, y)) \cup SP) \zeta I(x, y, z).$$

Where,

\cup - Denotes morphological operations involving region growing algorithm based image filling implementation of Matlab R14.

ζ - Denotes mapping operator to extract the background.

And, find foreground as

$$FG(x, y) = I(x, y) \sim BG(x, y).$$

Where,

\sim - Denotes exclusion operator to reveal the foreground from the original image by excluding the background.

Refer Figure 32 (b) for the results.

Step 9: Find watershed pixels constituting foreground object boundaries. The watershed pixels are the pixels not belonging to any regions. Two rows and two columns on the image boundaries are excluded.

Refer Figure 32 (c) for the results.

Step 10: Find the region attributes of foreground objects.

Algorithm 3. Foreground objects detection & background separation.

5.3 Results

The intermediate results produced by the method, qualitative comparison of results on typical representative images including images of standards databases and qualitative comparison of results of proposed method with that of JSEG [Deng, on line] [Deng, 2001] and human segmented images of standard databases [Fowlkes, on line] [Martin, 2001] are presented in this section.

5.3.1 Step-wise Results of the Method

The original image [Everingham, on line] and detected prominent boundaries of all four channels are shown in Figure 31 (a) and Figure 31 (b) respectively. Figure 31 (c) shows the results of prominence measure for prominent boundaries of each channel. Figure 31 (d) is the results of labeled watershed transform regions corresponding to Figure 31 (c). Though prominent boundaries were detected precisely, the generated watershed regions do not depict all regions of foreground. The morphological

operations are performed to get enhanced prominence measures as shown in Figure 31 (e).

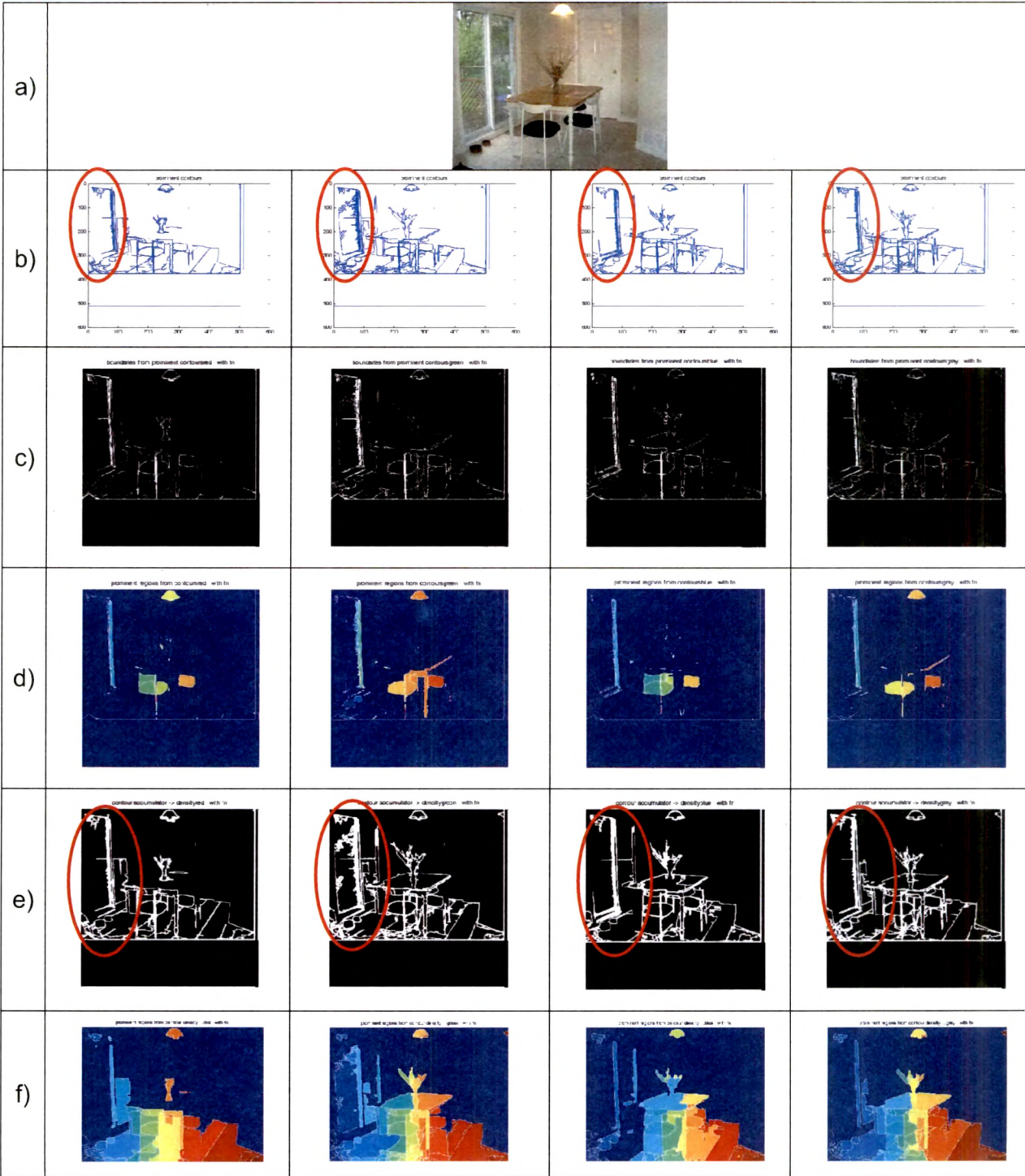


Figure 31. Foreground objects detection & background separation - Step-wise results – 1. (a) Original image [Everingham, on line]. (b) Prominent boundaries of R, G, B and Gray channels.(c) Respective prominence measures. (d) Respective labeled watershed regions of (c). (e) Respective enhanced prominence measures after morphological operations. (f) Respective labeled watershed regions of (e).

As the contribution of each channel for defining real prominent boundaries is non-homogeneous, individual prominent boundaries and corresponding prominence measures of each channel still do not cover all real prominent boundaries, as illustrated by encircling such portion of boundaries in Figure 31 (e). The labeled watershed regions corresponding to Figure 31 (e) are shown in Figure 31 (f).

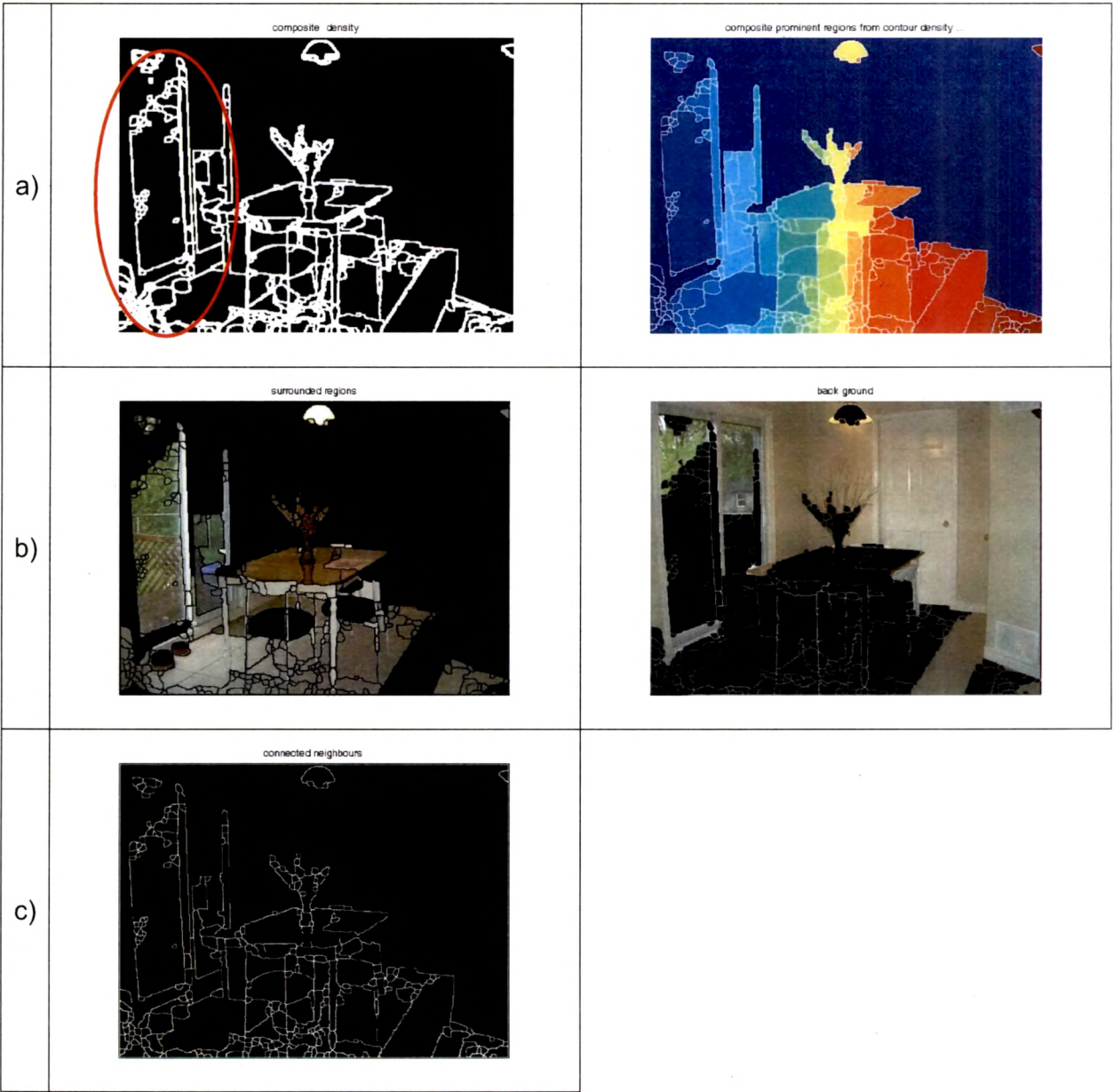


Figure 32. Foreground objects detection & background separation - Step-wise results – 2. (a) Left – Composite watershed regions from Figure 31 (e). (a) Right – Labeled watershed regions. (b) Left - Extracted foreground. (b) Right – background. (c) Watershed pixels of foreground regions.

The composite prominence measure obtained is shown in Figure 32 (a) - Left and respective labeled watershed regions are presented in Figure 32 (a) Right. The extracted foreground and background are shown in Figure 32 (b) – Left and Figure 32 (b) – Right respectively. The watershed pixels corresponding to foreground and foreground regions are shown in Figure 32 (c).

5.3.2 Qualitative Comparisons

Table 4. Categorical Representative Test Images & Their Performance Challenging Salient Characteristics.

Figure 34	A natural image [Wang, 2001] [SIMPLicity, on line]; textured background; foreground object having intra-object texture variations.
Figure 35, Left	A typical image [Wang, 2001] [SIMPLicity, on line] of a human face; textured background;
Figure 35, Right	A synthesized image [Wang, 2001] [SIMPLicity, on line]; a single central main foreground object with a typical texture; homogeneous background;
Figure 36, Left	An image resized to 1/8 th of the original size; image captured by an amateur with high resolution SONY device having inbuilt bionz image processor.
Figure 36, Right	A natural image [Wang, 2001] [SIMPLicity, on line]; multiple-textured background; foreground object having intra-object texture variations; shadowed foreground object.
Figure 37, Left	A natural image [Wang, 2001] [SIMPLicity, on line]; multiple similar partially touching foreground objects.
Figure 37, Right	An image [University of Washington, on line] with multiple textures; self reflection attached to foreground object;
Figures 38 - 41	The BSDb test images & associated human segmented images [Fowlkes, on line] [Martin, 2001]; covering multiple objects, occluded and shadowed objects, natural and man-made objects; distant and small objects; inter region and intra-region texture variations in foreground objects and background;

The image segmentation by human being is intrinsically characterized by grouping of regions based on 'some' perceptual similarities. Though there exists subjectivity in the perception of visual similarities, human beings do not over-segment

the objects contained in the images. That is, the focus of segmentation by humans is objects and not regions. Hence, the results of proposed method for foreground detection have been qualitatively compared with the human segmented images of Berkeley Segmentation Dataset and Benchmark database of natural images [Fowlkes, on line] [Martin, 2001].

The results on categorical representative test images, possessing performance challenging salient characteristics listed in Table 4, are shown in Figure 34 to Figure 41. The results presented are for prominent boundaries detection, separation of background, revealing of foreground objects, region features using watershed transform and corresponding watershed pixels of foreground objects and artifacts reduced regions and corresponding boundaries by incorporating different levels of stationary Haar wavelet decomposition. Figure 34 demonstrates results with different levels of SWT on an image [Wang, 2001] [SIMPLicity, on line] having texture variations in the foreground object as well as in the background. The detected prominent boundaries incorporating SWT with Haar wavelet at level 1, 2 and 3 are shown in Figure 34 (b) from left to right respectively. The corresponding separated background is shown in Figure 34 (c), where pure black regions do not belong to the background. The corresponding foreground is shown in Figure 34 (d), where pure black regions do not belong to the foreground. The watershed regions and watershed pixels are respectively shown in Figure 34 (e) and Figure 34 (f). The watershed regions and watershed pixels with reduced artifacts are respectively shown in Figure 34 (g) and Figure 34 (h). Figure 35 - Left shows the results on the image [Wang, 2001] [SIMPLicity, on line] having poorly defined object boundaries where the Haar stationary wavelet decompositions of step 1 of the method are omitted. Figure 35-Right to Figure 37 demonstrates the results incorporating Haar wavelet decompositions at level 2 on different categorical representative test images. Figure 38 and Figure 39 show comparison of the method results with stationary Haar wavelet decompositions at level 2 and level 3 on images [Fowlkes, on line] [Martin, 2001] with human segmented images of Segmentation Dataset and Benchmark [Fowlkes, on line] [Martin, 2001]. Similarly, Figure 40 and Figure 41 show comparison of the method results with stationary Haar wavelet decompositions at level 2 on different categorical representative test images [Fowlkes, on line] [Martin, 2001] with human segmented images of segmentation dataset and benchmark [Fowlkes, on line] [Martin, 2001].

The results are to be observed and compared for

- Localization of detected boundaries.
- Effect of complex background, textures on segmentation / foreground extraction.
- Effect on segmentation / foreground extraction due to smooth changes.
 - in textured regions
 - of colors within regions
 - of intensities within region
- Suitability of segmentation results for foreground separation / object detection.
- Human segmented images of BSDB [Fowlkes, on line] [Martin, 2001].

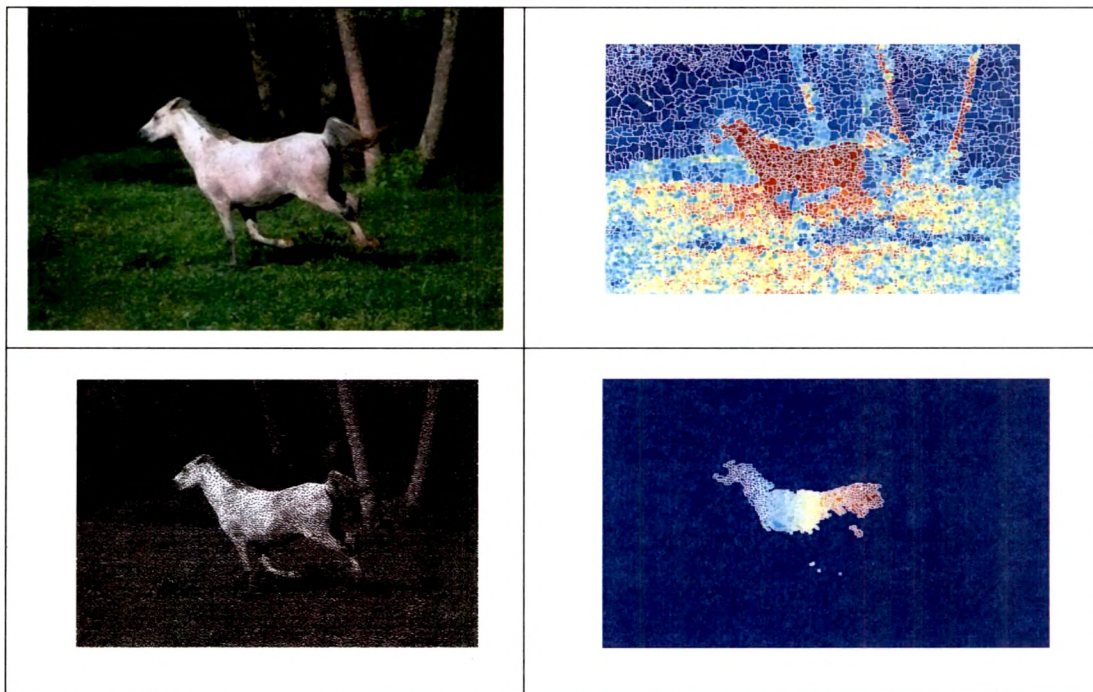


Figure 33. Watershed transform based segmentation without applying proposed method. Top-left: Original image [Wang, 2001] [SIMPLcity, on line]. Top-right: Watershed regions of gray scale converted image of top-left. Bottom-left: Dithered image of top-left. Bottom –right: Watershed regions of dithered image.

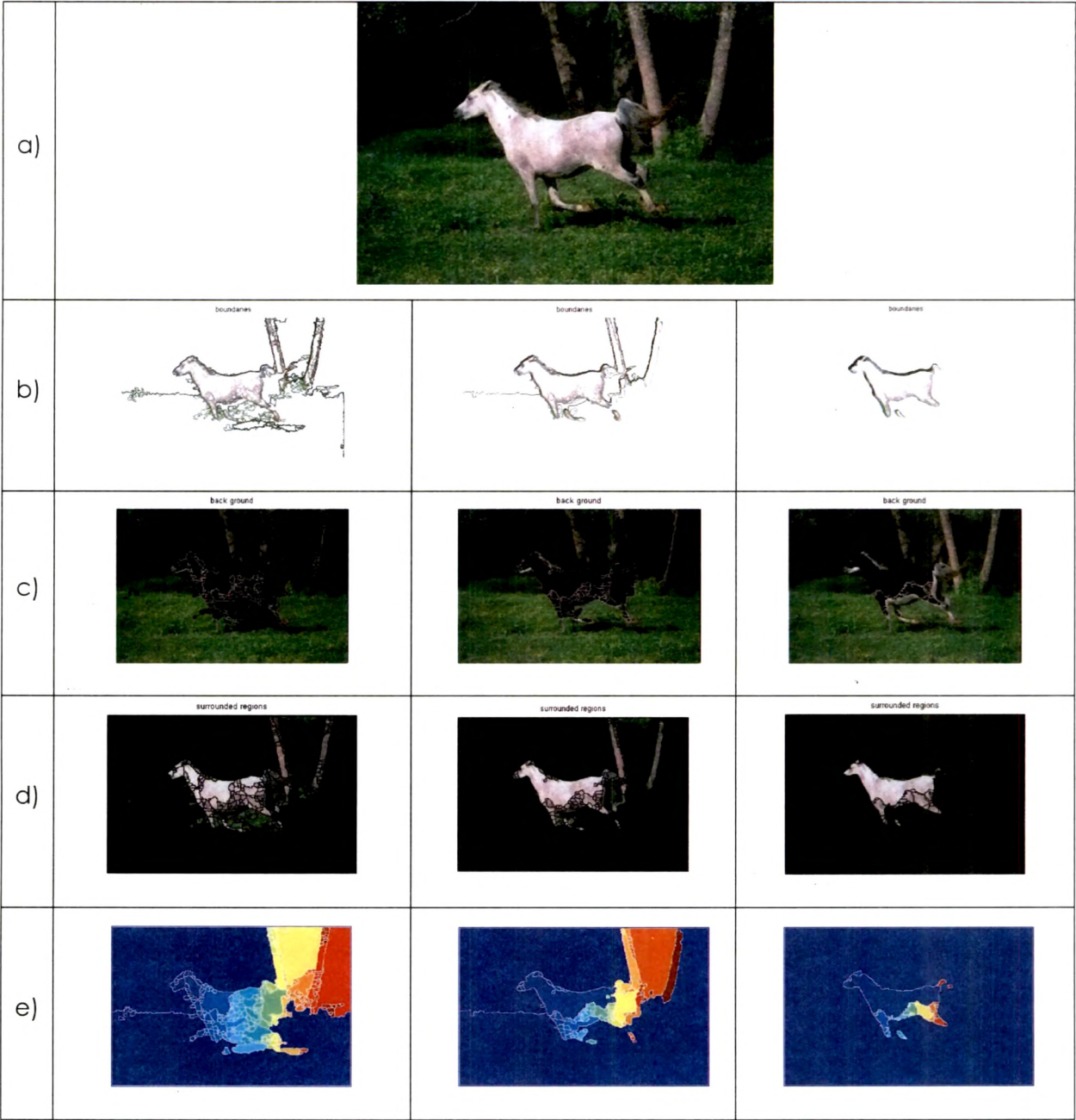


Figure 34. Revealed fore ground objects & background, incorporating different levels of wavelet decompositions. (a) Original image [Wang, 2001] [SIMPLIcity, on line]. (b) Detected prominent boundaries. Left: Incorporating stationary Haar wavelet decomposition at level 1. Middle: Incorporating stationary Haar wavelet decomposition at level 2. Right: Incorporating stationary Haar wavelet decomposition at level 3. (c) Separated background. (d) Revealed foreground. (e) Watershed regions.

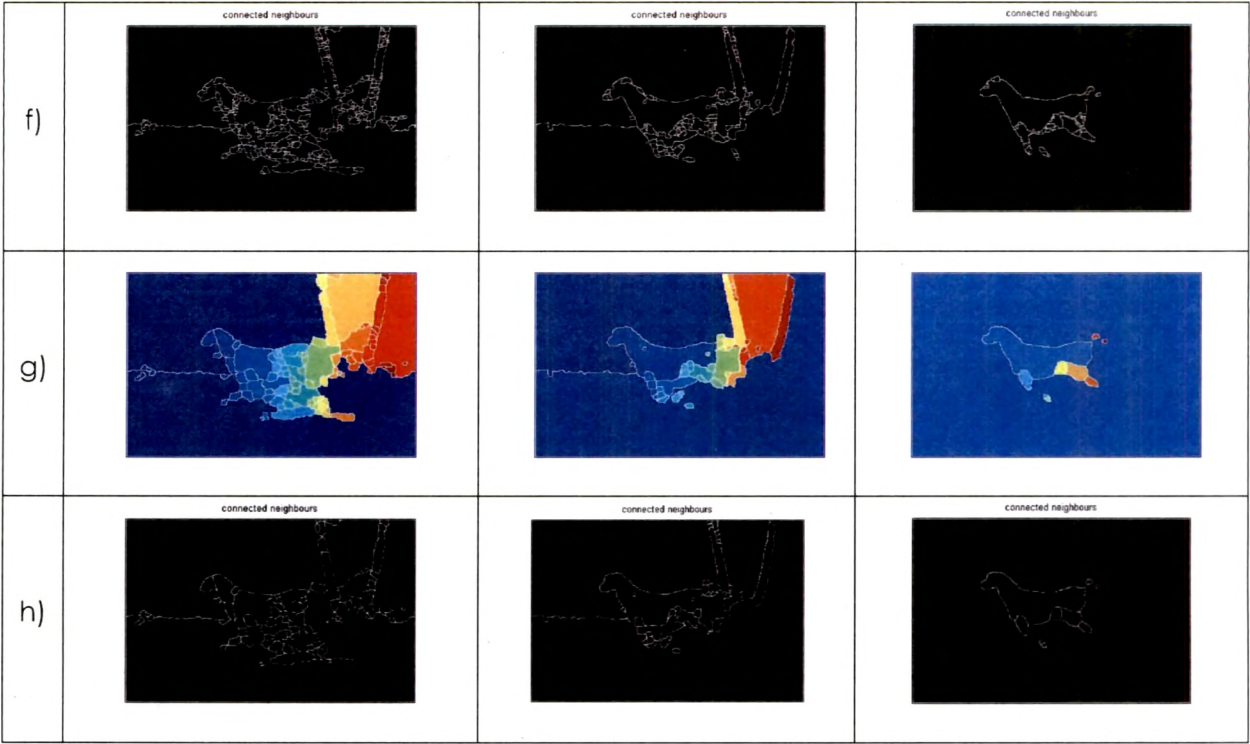


Figure 34. (Contd.). Revealed fore ground objects & background, incorporating different levels of wavelet decompositions. (f) Corresponding watershed pixels of foreground object boundaries. (g) Watershed regions with reduced artifacts. (h) Corresponding watershed pixels of foreground object boundaries.

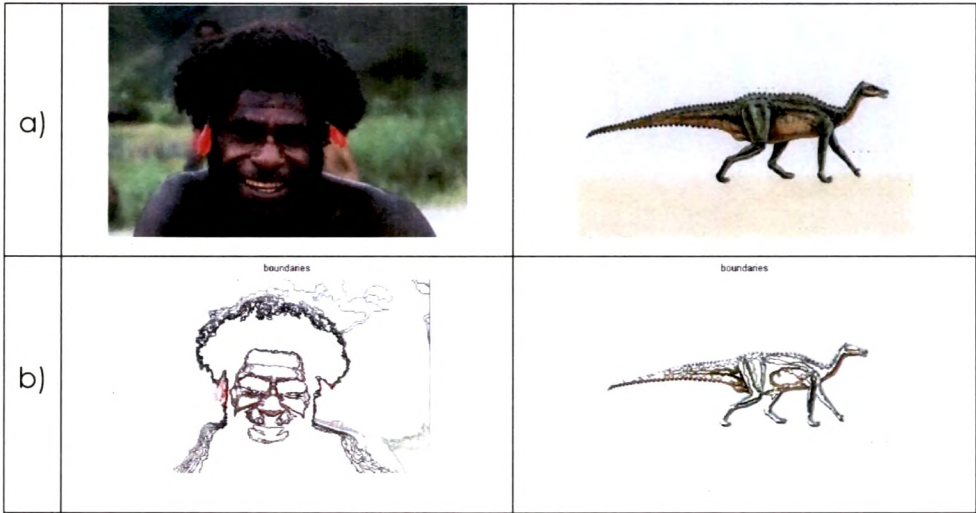


Figure 35. Revealed fore ground objects & background for images with typical textures. (a) Original images [Wang, 2001] [SIMPLIcity, on line]. (b) Detected prominent boundaries. Left: Without wavelet decomposition. Right: Incorporating stationary Haar wavelet decomposition at level 2.

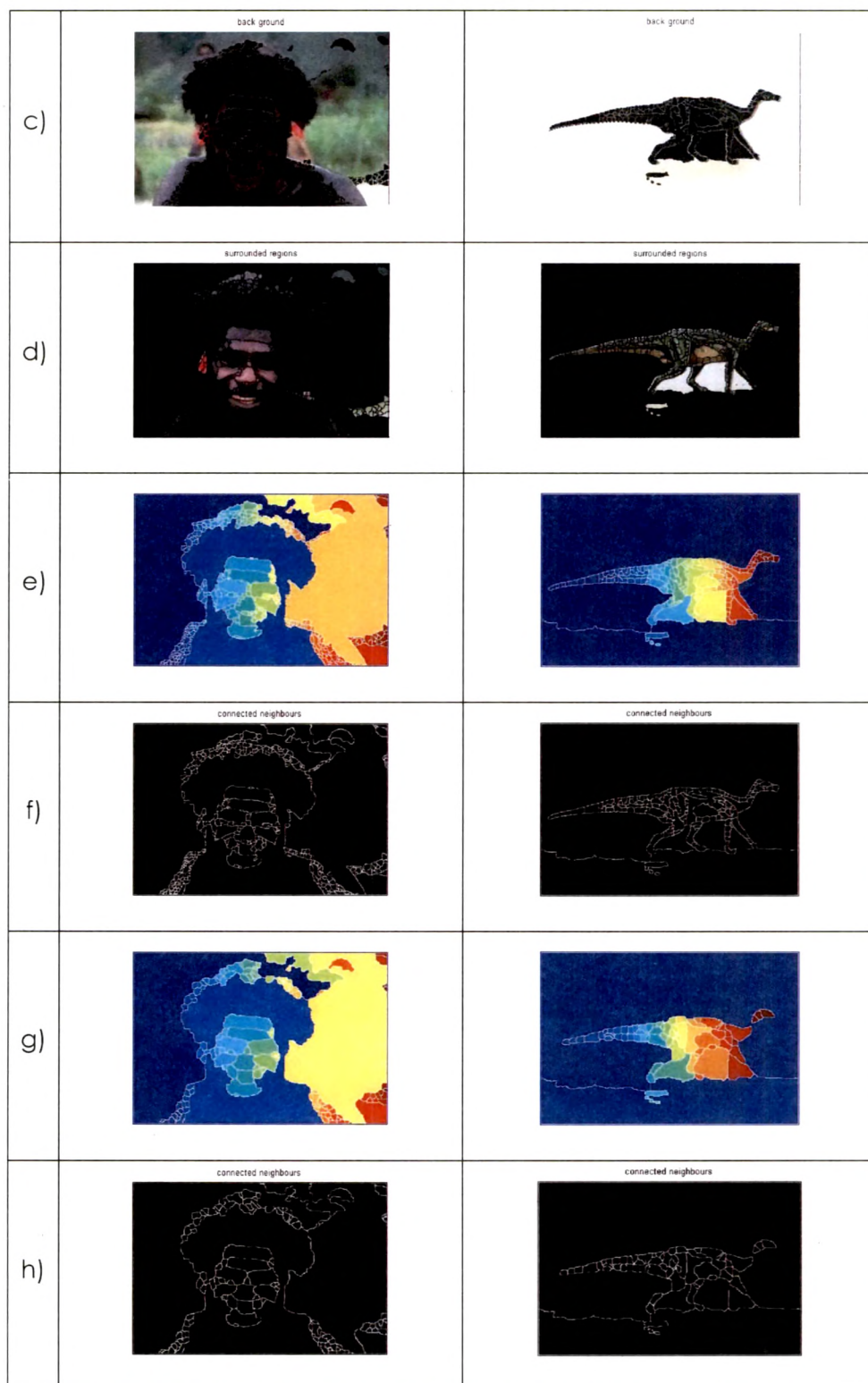


Figure 35 (Contd.). Revealed fore ground objects & background for images with typical textures. (c) Separated background. (d) Revealed foreground. (e) Watershed regions. (f) Corresponding watershed pixels of foreground object boundaries. (g) Watershed regions with reduced artifacts. (h) Corresponding watershed pixels of foreground object boundaries.

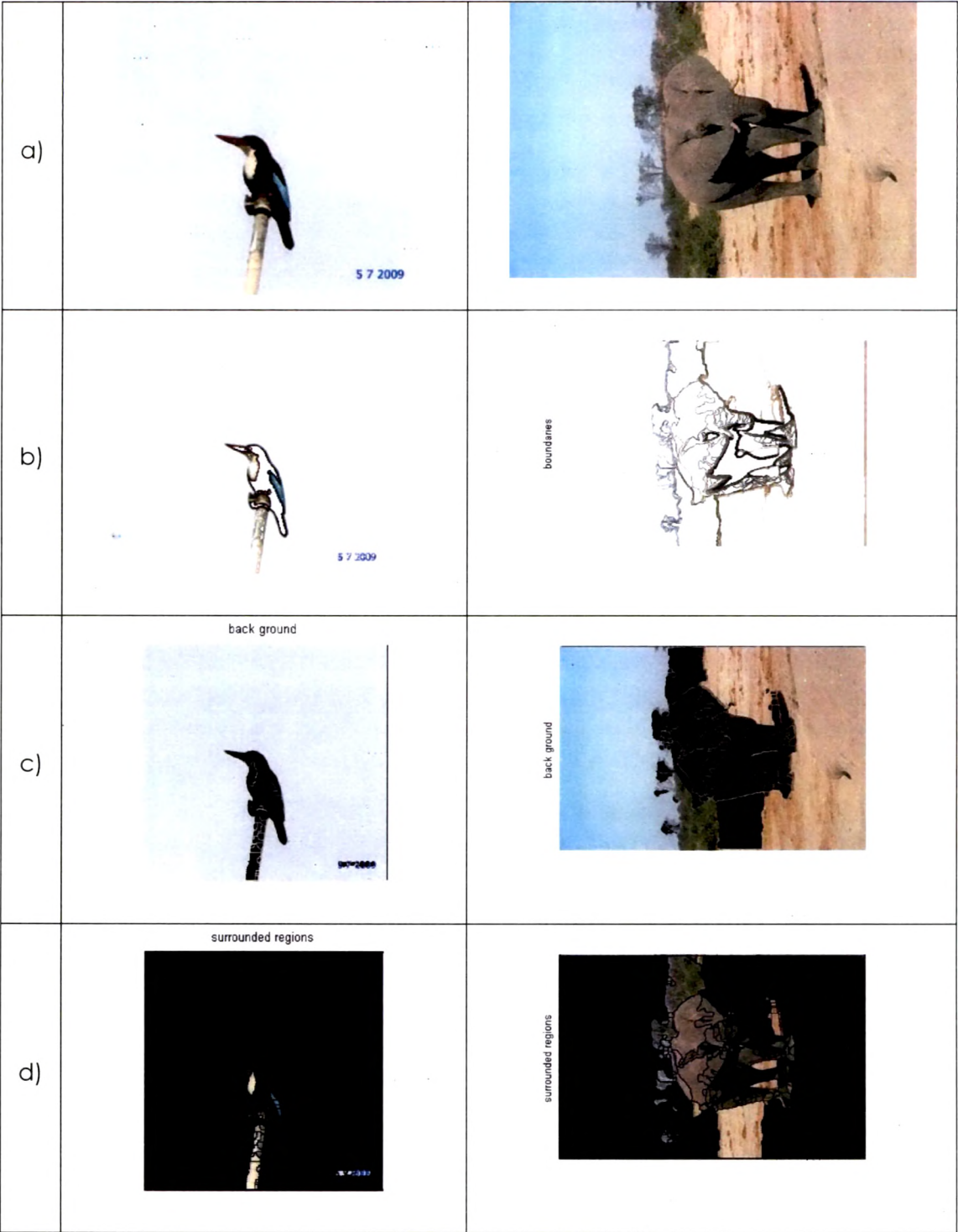


Figure 36. Revealed foreground objects & background, incorporating stationary Haar wavelet decompositions at level 2. (a) Left: Original image. Right: Original image [Wang, 2001] [SIMPLIcity, on line]. (b) Detected prominent boundaries. (c) Separated background. (d) Revealed foreground.

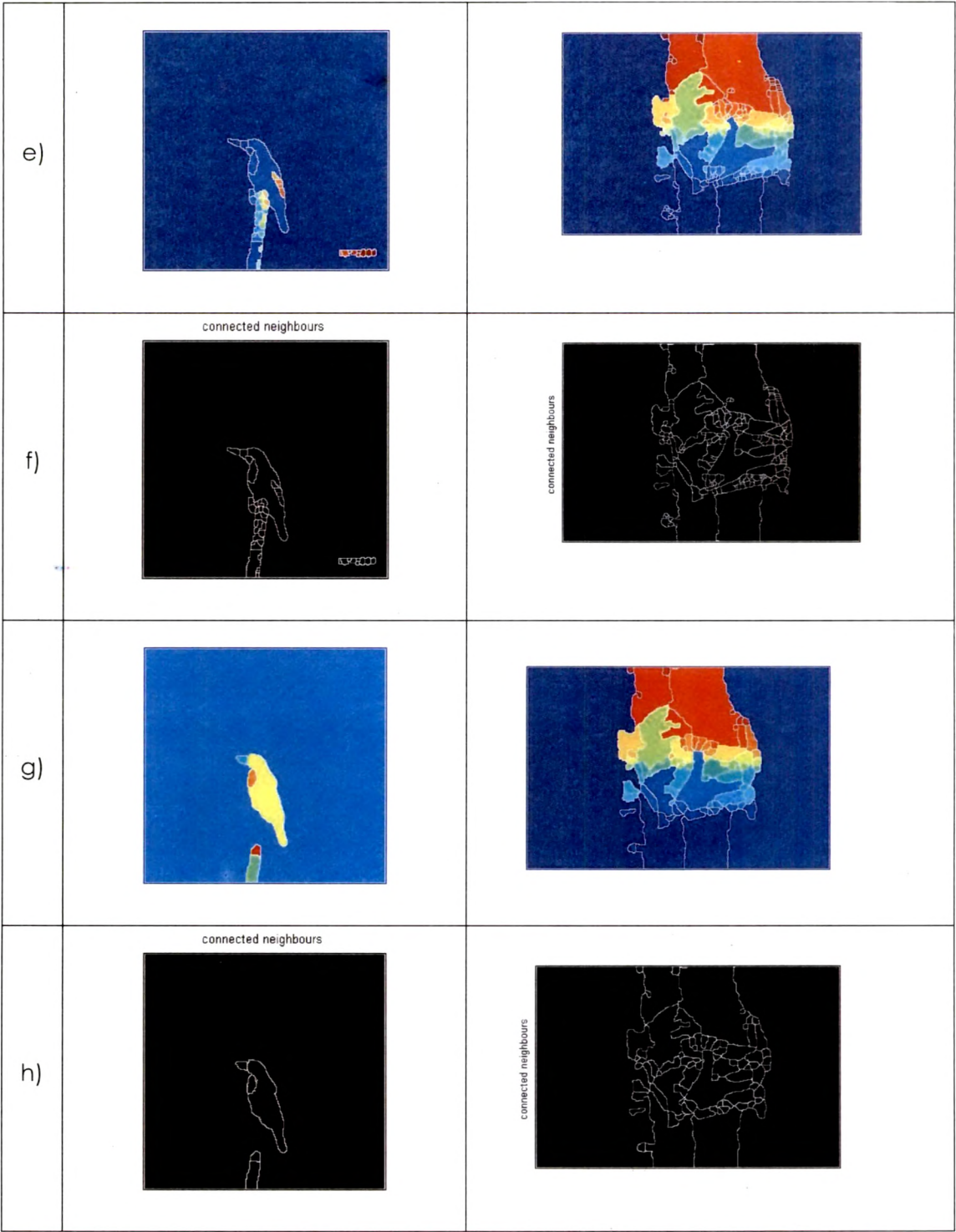


Figure 36. (Contd.). Revealed foreground objects & background, incorporating stationary Haar wavelet decompositions at level 2. (e) Watershed regions. (f) Corresponding watershed pixels of foreground object boundaries. (g) Watershed regions with reduced artifacts. (h) Corresponding watershed pixels of foreground object boundaries.

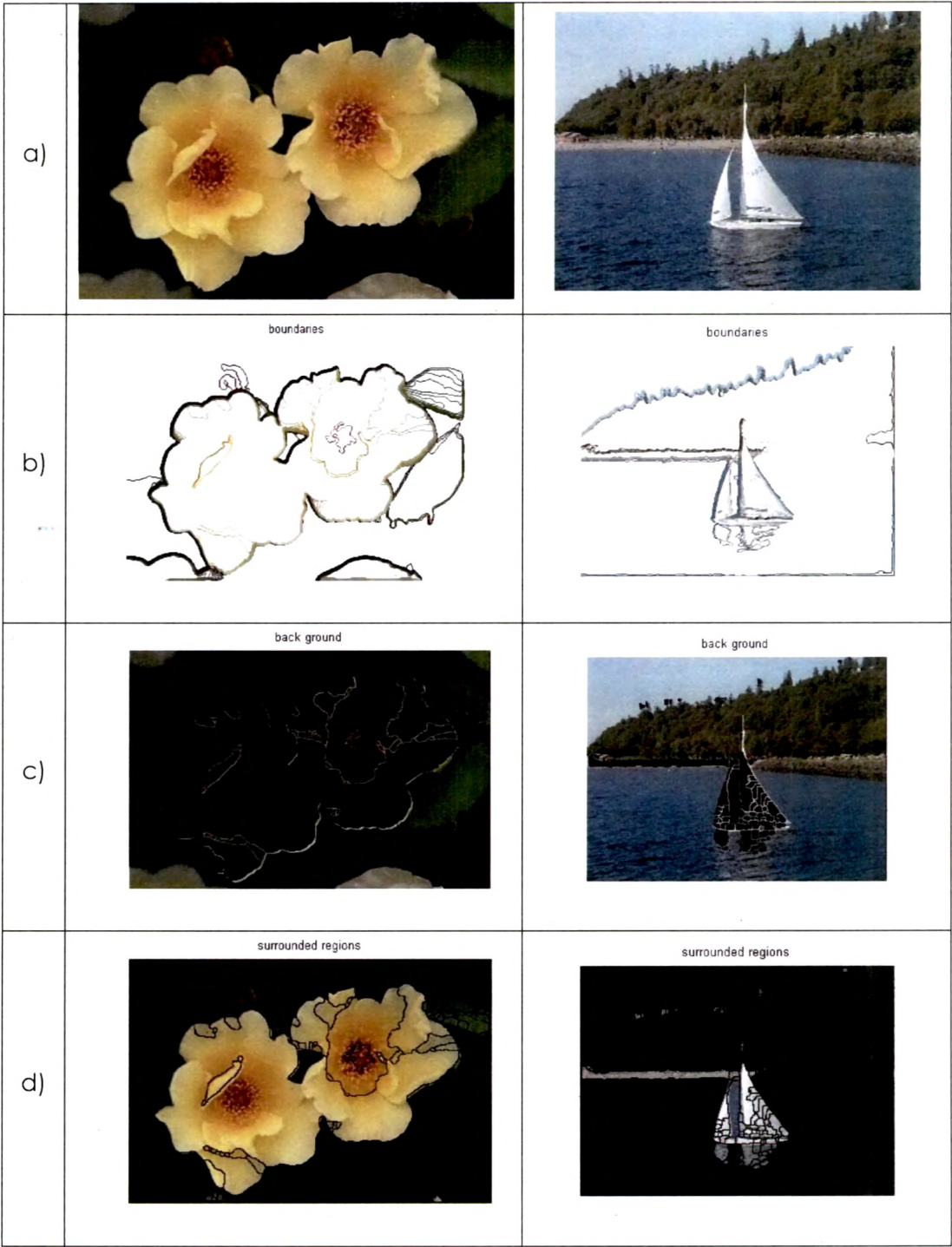


Figure 37. Revealed foreground objects & background, incorporating stationary Haar wavelet decompositions at level 2. (a) Left: Original Image [Wang, 2001] [SIMPLcity, on line]. Right: Original image [University of Washington, on line]. (b) Detected prominent boundaries. (c) Separated background. (d) Revealed foreground.

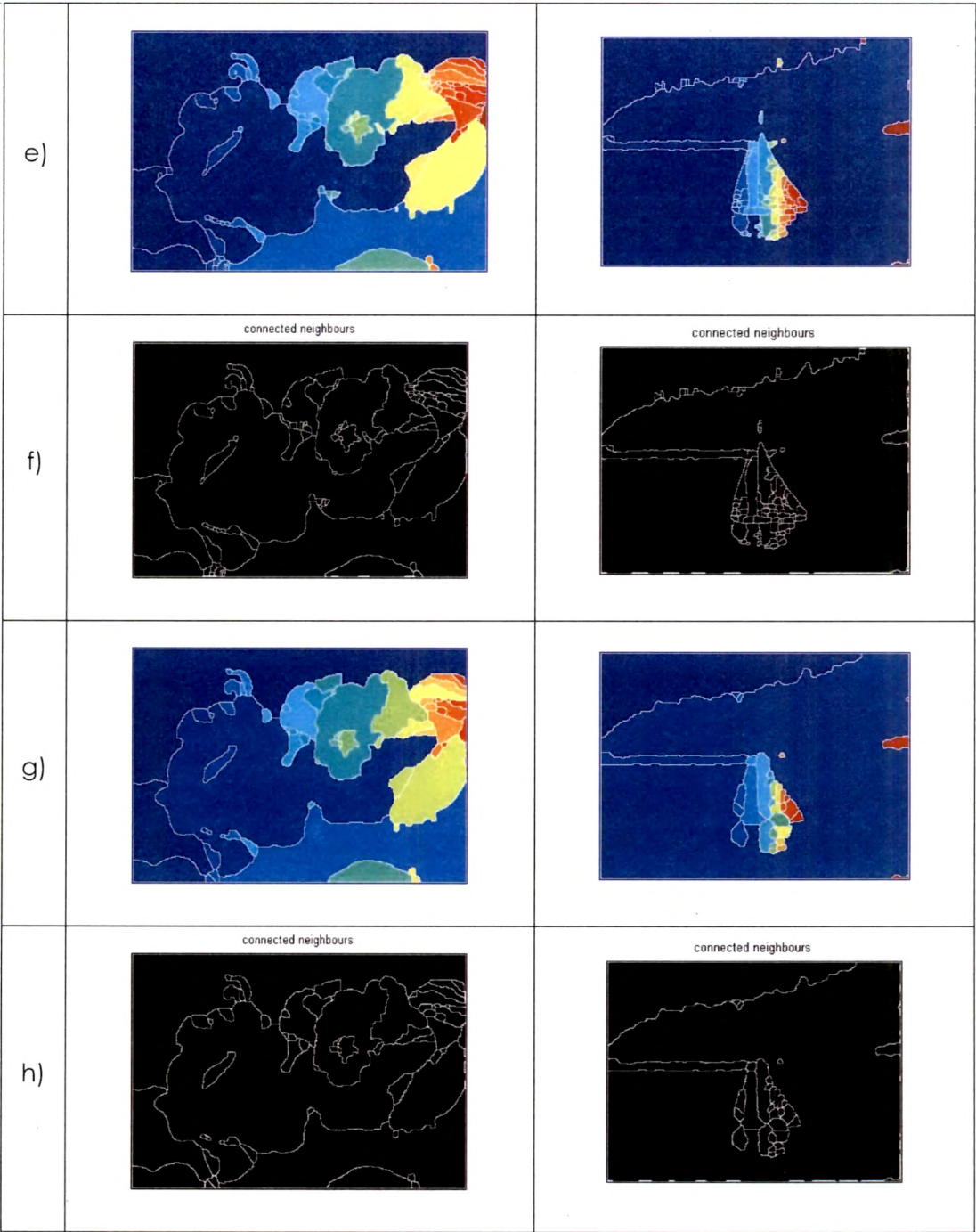


Figure 37 (Contd.). Revealed foreground objects & background, incorporating stationary Haar wavelet decompositions at level 2. (e) Watershed regions. (f) Corresponding watershed pixels of foreground object boundaries. (g) Watershed regions with reduced artifacts. (h) Corresponding watershed pixels of foreground object boundaries.

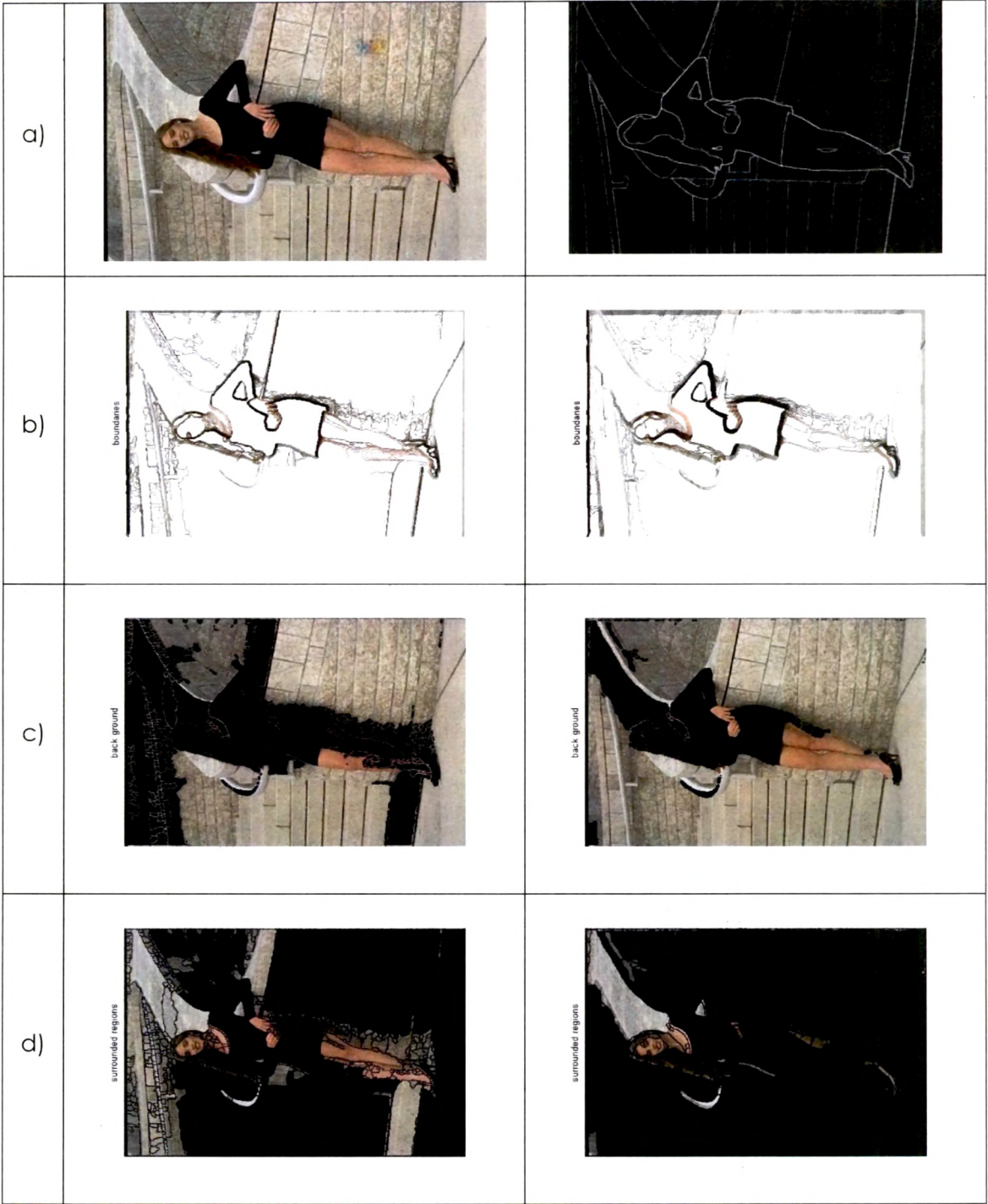


Figure 38. Result-comparison: Human segmented image with revealed foreground objects & background, incorporating different levels of stationary Haar wavelet decompositions. (a) Left: Original image [Fowlkes, on line] [Martin, 2001]. Right: Human segmented image [Fowlkes, on line] [Martin, 2001]. (b) Detected prominent boundaries. Left: Incorporating stationary Haar wavelet decomposition at level 2. Right: Incorporating stationary Haar wavelet decomposition at level 3. (c) Separated background. (d) Revealed foreground.

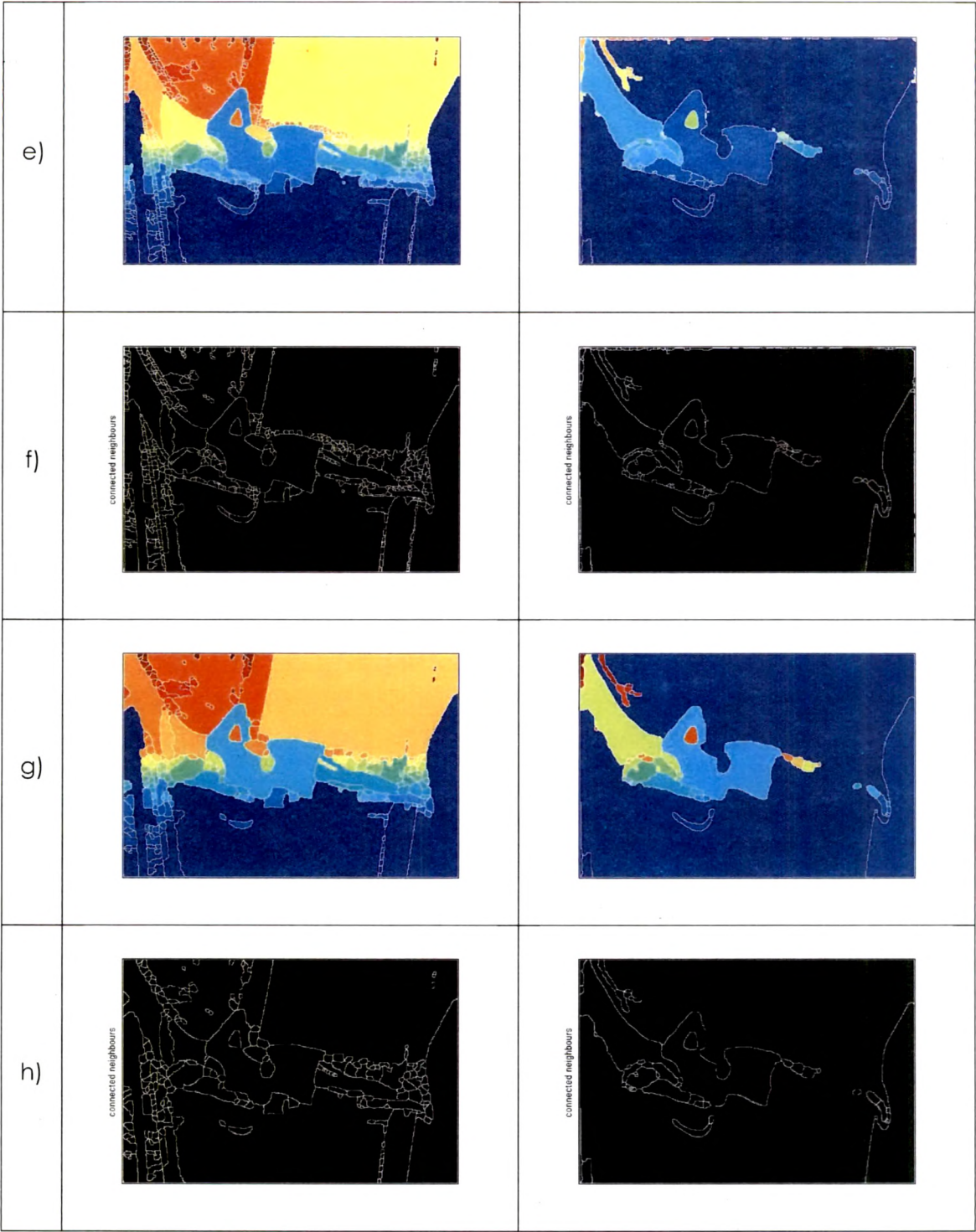


Figure 38 (Contd.). Result-comparison: Human segmented image with revealed foreground objects & background, incorporating different levels of stationary Haar wavelet decompositions. (e) Watershed regions. (f) Corresponding watershed pixels of foreground object boundaries. (g) Watershed regions with reduced artifacts. (h) Corresponding watershed pixels of foreground object boundaries.

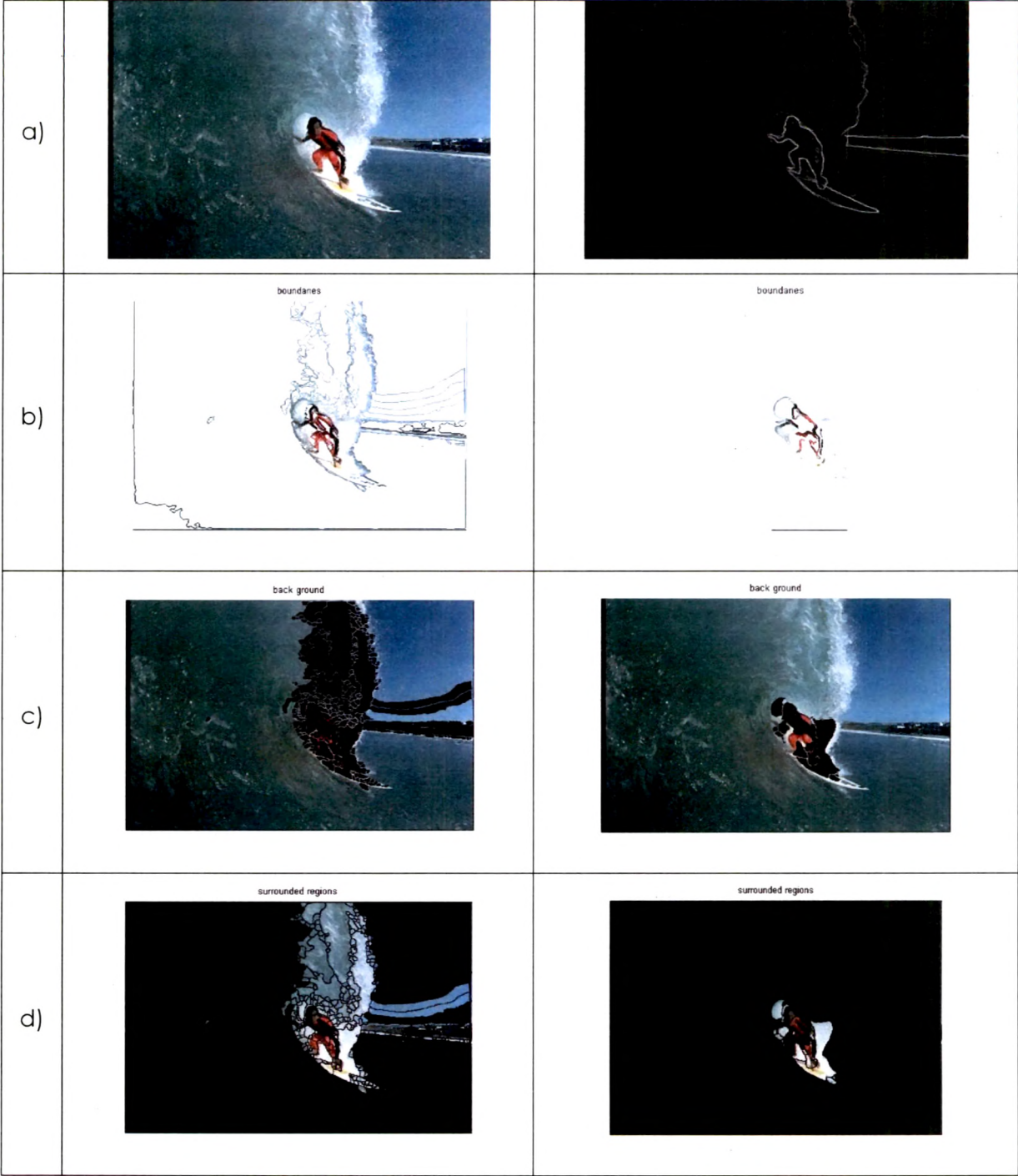


Figure 39. Result-comparison: Human segmented image with revealed foreground objects & background, incorporating different levels of stationary Haar wavelet decompositions. (a) Left: Original image [Fowlkes, on line] [Martin, 2001]. Right: Human segmented image [Fowlkes, on line] [Martin, 2001]. (b) Detected prominent boundaries. Left: Incorporating stationary Haar wavelet decomposition at level 2. Right: Incorporating stationary Haar wavelet decomposition at level 3. (c) Separated background. (d) Revealed foreground.

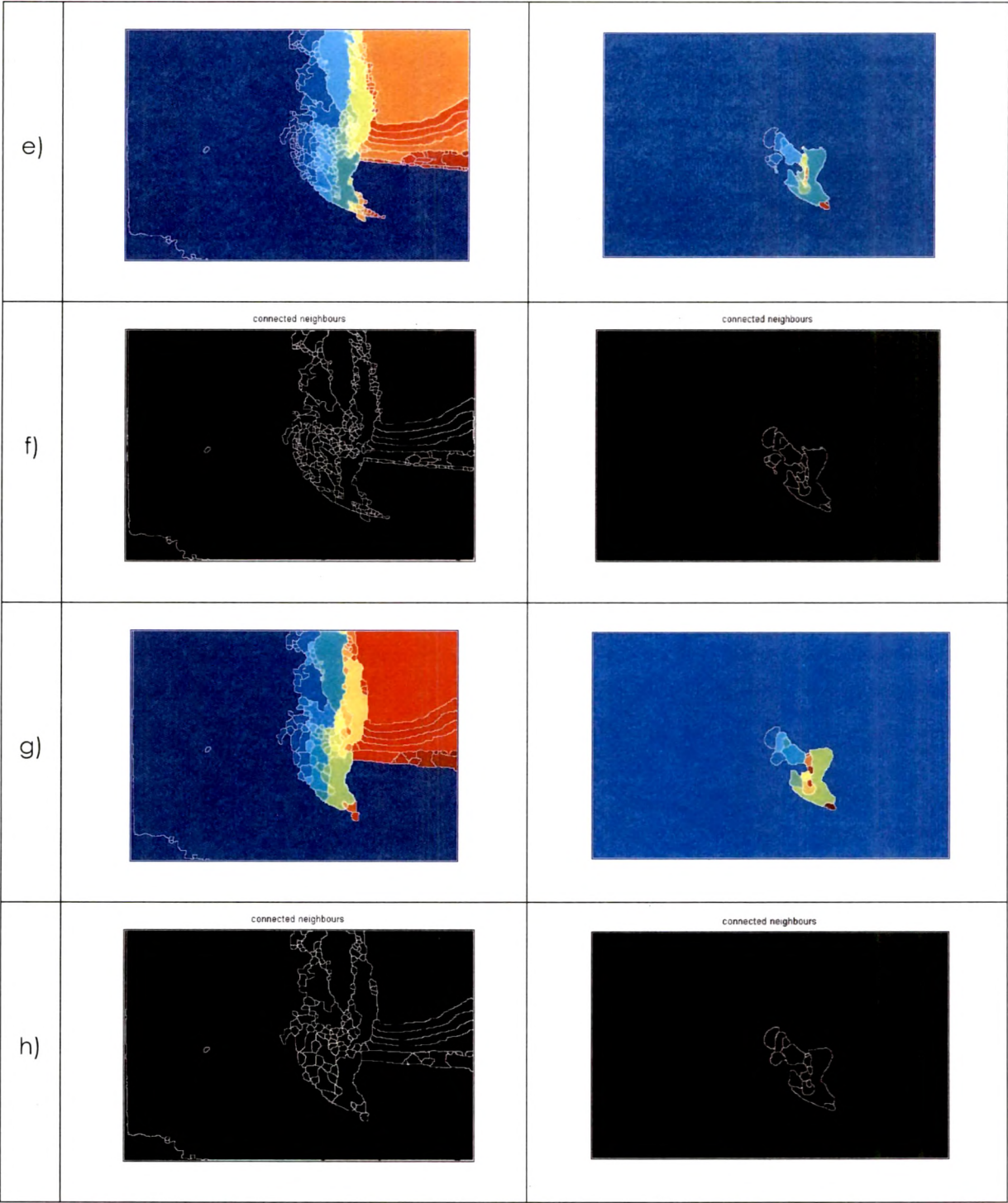


Figure 39 (Contd.). Result-comparison: Human segmented image with revealed foreground objects & background, incorporating different levels of stationary Haar wavelet decompositions. (e) Watershed regions. (f) Corresponding watershed pixels of foreground object boundaries. (g) Watershed regions with reduced artifacts. (h) Corresponding watershed pixels of foreground object boundaries.

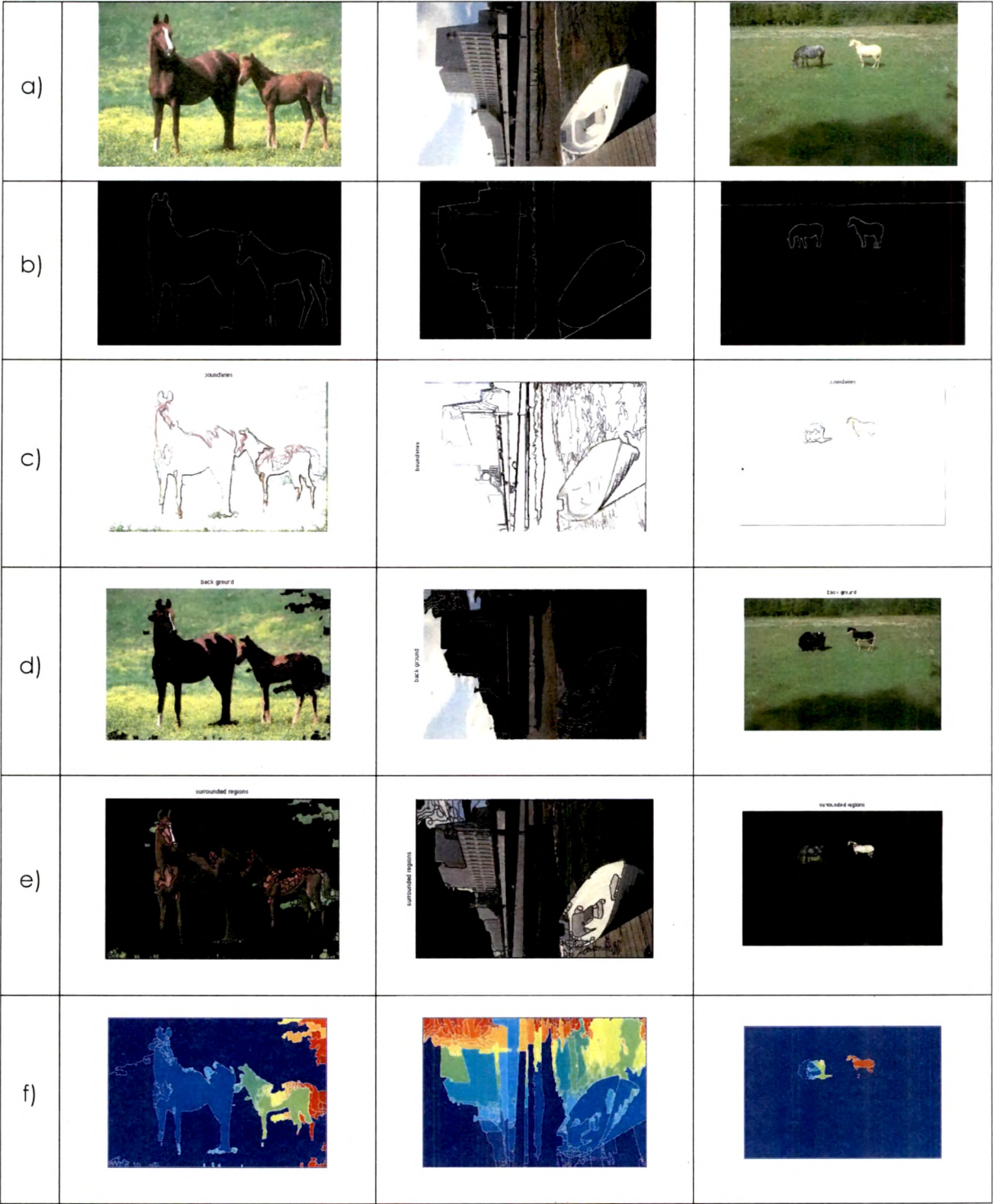


Figure 40. Result comparison: Human segmented images with revealed foreground objects & background, incorporating stationary Haar wavelet decompositions at levels 2. (a) Original images [Fowlkes, on line] [Martin, 2001]. (b) Human segmented images [Fowlkes, on line] [Martin, 2001]. (c) Detected prominent boundaries. (d) Separated background. (e) Revealed foreground. (f) Watershed regions.

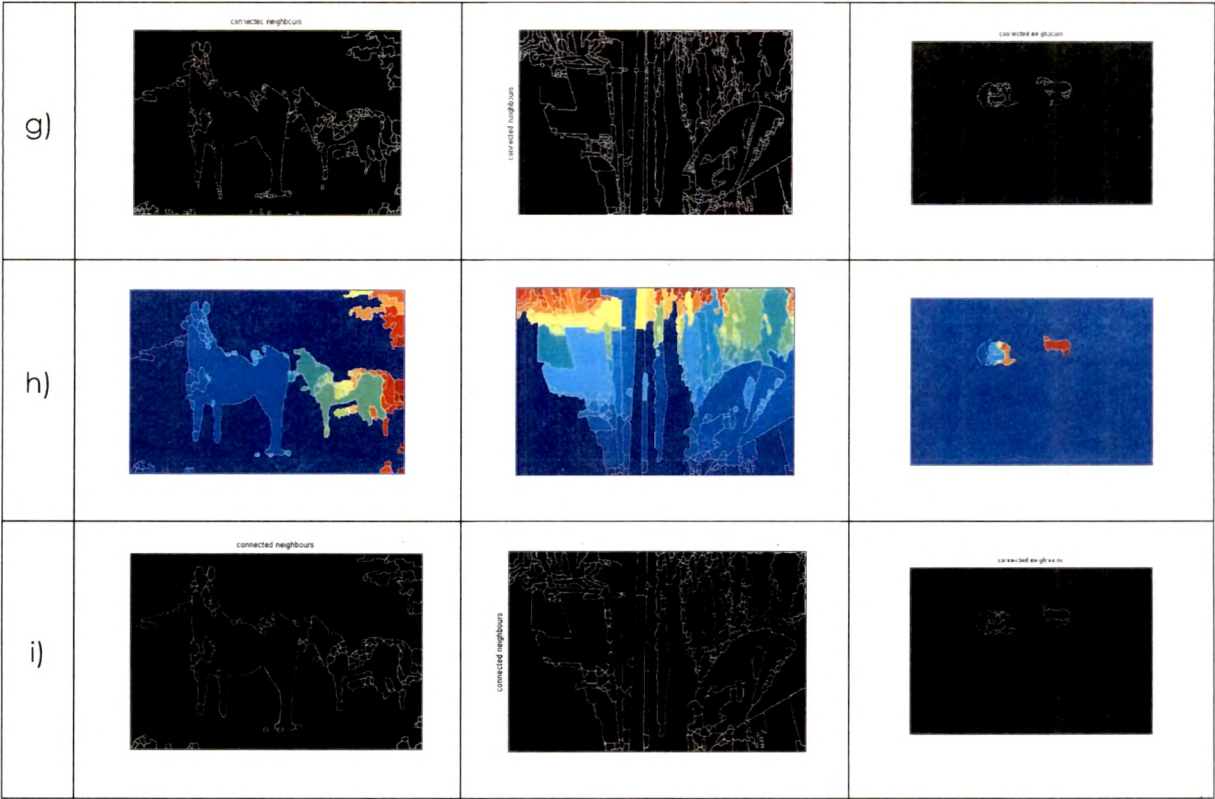


Figure 40 (Contd.). Result comparison: Human segmented images with revealed foreground objects & background, incorporating stationary Haar wavelet decompositions at levels 2. (g) Corresponding watershed pixels of foreground object boundaries. (h) Watershed regions with reduced artifacts. (i) Corresponding watershed pixels of foreground object boundaries.

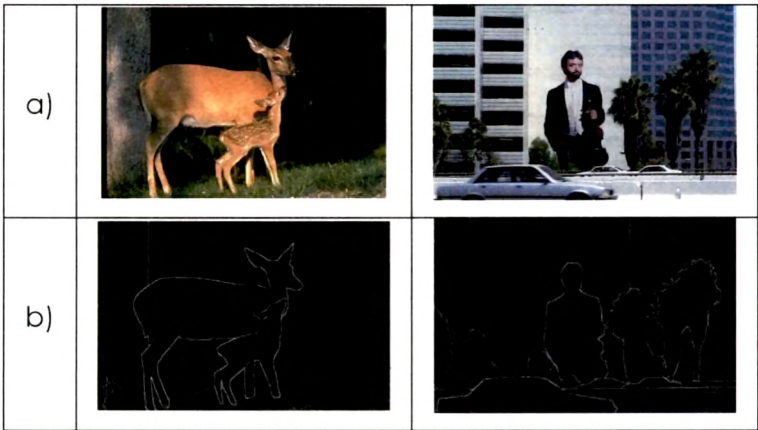


Figure 41. Result comparison: Human segmented images with revealed foreground objects & background. (a) Original images [Fowlkes, on line] [Martin, 2001]. (b) Human segmented images [Fowlkes, on line] [Martin, 2001].

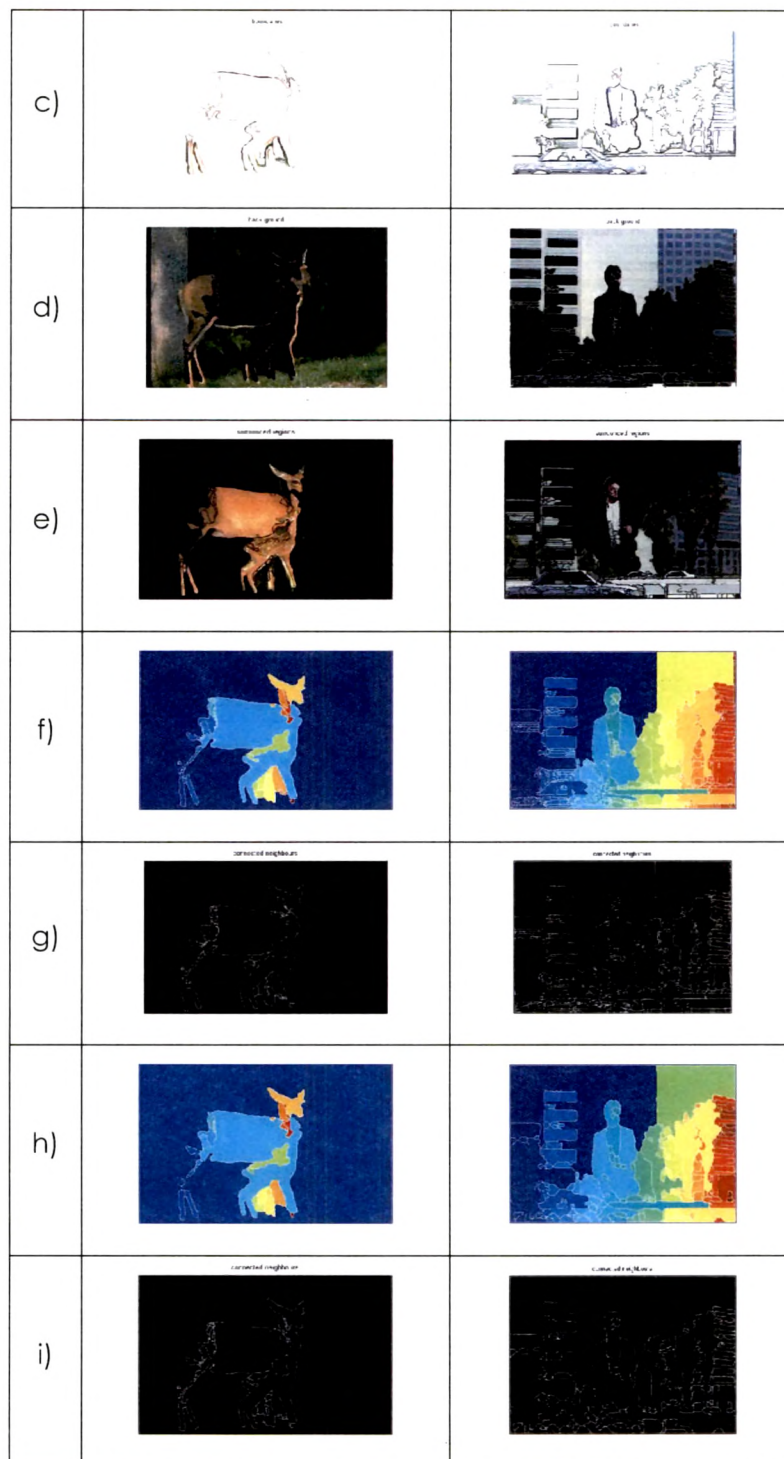


Figure 41 (Contd.). Result comparison: Human segmented images with revealed foreground objects & background. (c) Detected prominent boundaries. Left: Incorporating stationary Haar wavelet decomposition at level 3. Right: Incorporating stationary Haar wavelet decomposition at level 2. (d) Separated background. (e) Revealed foreground. (f) Watershed regions. (g) Corresponding watershed pixels of foreground object boundaries. (h) Watershed regions with reduced artifacts. (i) Corresponding watershed pixels of foreground object boundaries.

5.3.3 Qualitative Result-comparisons: Proposed Method, JSEG & Human

Segmented Images of BSDb [Martin, 2001] [Fowlkes, on line]

JSEG algorithm, proposed in [Deng, on line] [Deng, 2001] incorporates color quantization and spatial segmentation stages, where colors of the images are quantized in several color classes to produce class-map of the image in the first step. The class image contains the pixels labeled with the color class. The local window based spatial criterion is applied on class-map to generate multi-scale J-Images which are processed for region growing method leading to segmentation. The values of J-Image represent possible boundaries and interiors of color-texture regions. The algorithm can be applied to still images and video. Despite being a fast-computationally efficient algorithm, JSEG segmentation results are sensitive to various parameters - quantization scale, number of scales and region merge thresholds, a pointed mentioned and illustrated qualitatively in [Deng, 2001] with the help of segmentation results at different parameters. The major limitation of the algorithm, cited in the paper [Deng, 2001] with the remark **"There seems to be no easy solution"**, is to handle shades and smooth transitions due to illumination variations caused by intensity or spatial changes of illumination source.

Following Figures of the section compares results of proposed methods with that of JSEG [Deng, on line] [Deng, 2001] segmentation and human segmented natural images of BSDb [Fowlkes, on line] [Martin, 2001] to illustrate the suitability of proposed method to eliminate the background for extracting the foreground from images possessing variety of background and characteristics. *The JSEG results are obtained with the software available at JSEG site [Deng, on line] for default value of number scale and region merge threshold with quantization threshold specified as 255.*

The well localization of prominent boundaries and effectiveness of foreground extraction has been illustrated in Figure 42 and Figure 43 for images of ALOI database [ALOI, on line] [Geusebroek, 2001] captured under controlled conditions for varied viewing angles, illumination angles and illumination color. Refer [Section 6.4](#) for description and characteristics of the database. The Figure 42 contains images of objects with finer well defined details with illumination changes. And, Figure 43 contains objects with illumination changes producing shadows and intensity variations. Though images included in Figure 42 and Figure 43 contain monotonous gray background making foreground extraction simple, they are included here to show effects of

illumination variations produced under controlled conditions on segmentation / foreground extraction.

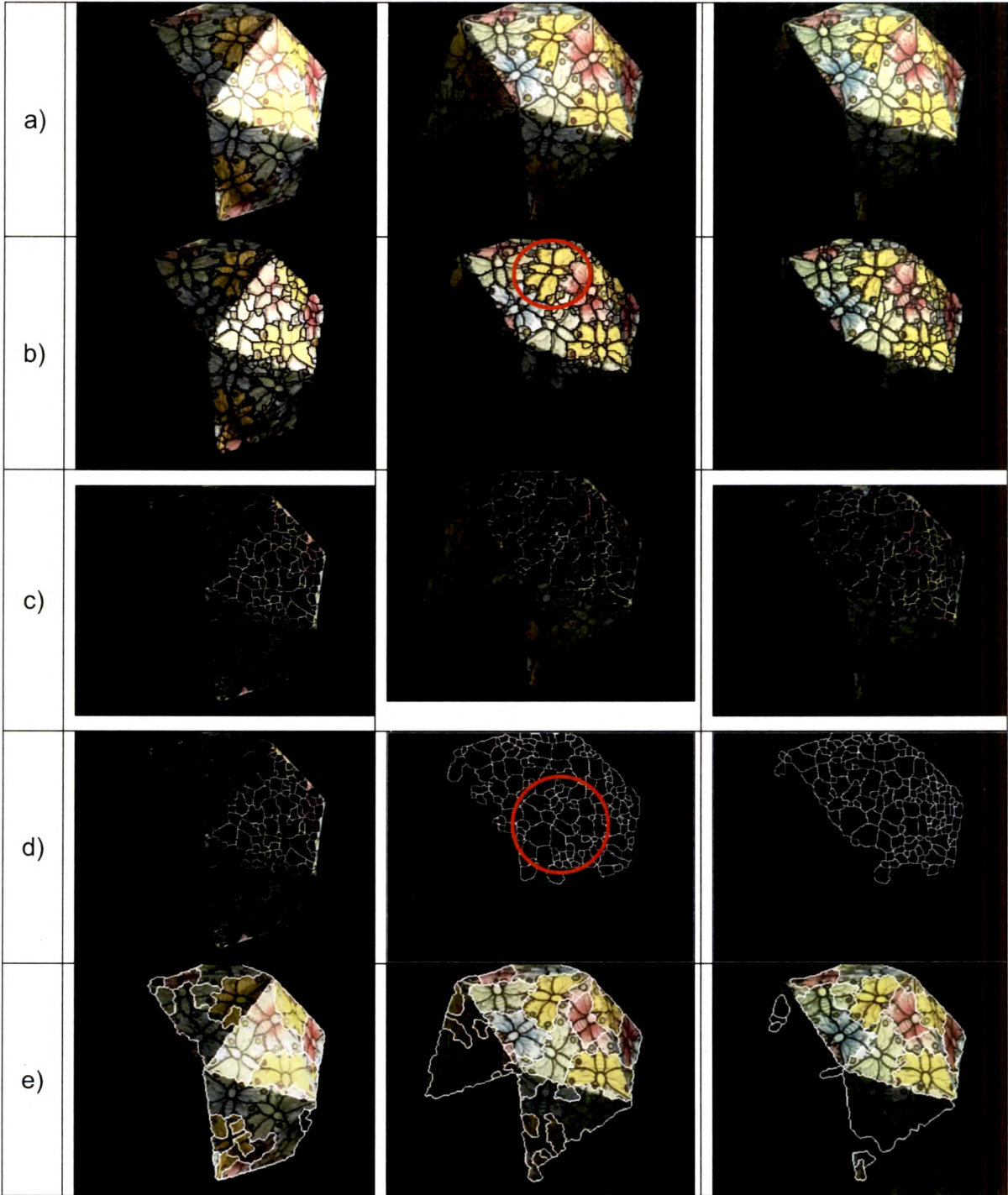


Figure 42 Effect of illumination variations on segmentation / foreground extraction on object with fine details. (a) Original Images [ALOI, on line] [Geusebroek, 2001]. (b) Extracted foreground of (a). (c) Background. (d) Watershed pixels corresponding to (b). (e) Segmentation with JSEG [Deng, on line] [Deng, 2001].

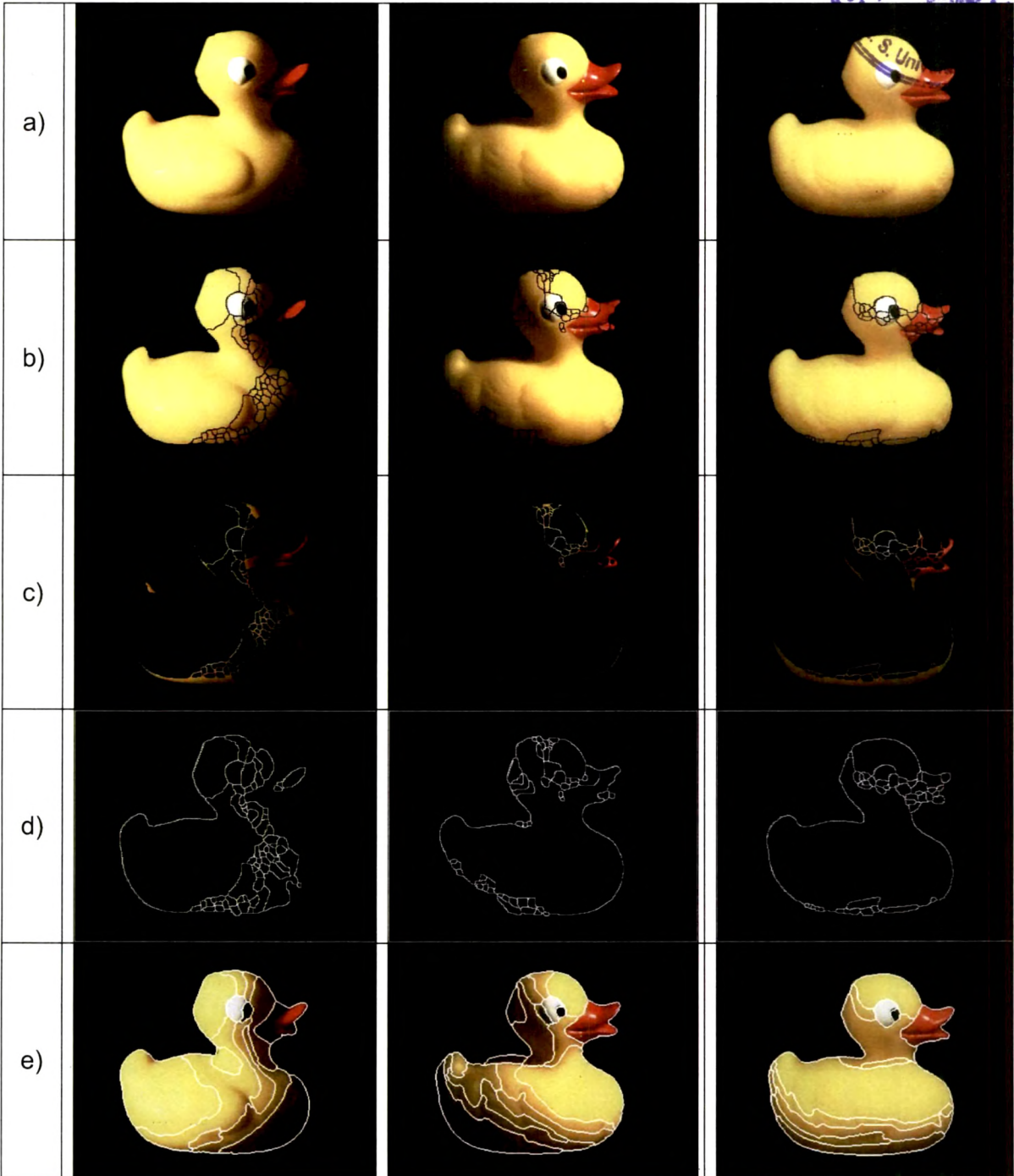


Figure 43. Effect of illumination variations on segmentation / foreground extraction on non-textured object. (a) Original Images [ALOI, on line] [Geusebroek, 2001]. (b) Extracted foreground of (a). (c) Background. (d) Watershed pixels corresponding to (b). (e) Segmentation with JSEG [Deng, on line] [Deng, 2001].

As png file format is not supported by JSEG software [Deng, on line], the jpg equivalent of ALOI images are used in Figure 42 and Figure 43 for JSEG segmentation. The segmentation / foreground extraction results of the proposed method are

compared with the results produced with JSEG [Deng, on line] [Deng, 2001]. The finer prominent details are not detected by JSEG - Figure 42 (e), whereas they are marked as prominent boundaries by the proposed method as shown in Figure 42 (d). The boundaries enclosing butterflies are well detected and one such case is encircled and illustrated in Figure 42 (b) and Figure 42 (d).

As illustrated in Figure 43 (e), smooth intensity changes / shadows tend to over-segment regions with JSEG segmentation (the cited limitation in [Deng, 2001]). The spatial and intensity variations resulting into shadows and shades of smooth variations do not significantly alter outer boundaries of the foreground detected with proposed method. The over-segmentation produced around some boundaries due to watershed transformation has no effect on foreground extraction.

The smooth changes and diversities in colors & textures are performance challenging characteristics of natural images for segmentation. The results of segmentation / foreground extraction with proposed method for natural images are shown in Figure 44 to Figure 46 for comparisons with Human segmented images of BSDB [Fowlkes, on line] [Martin, 2001] and segmentation results of JSEG [Deng, on line] [Deng, 2001]. The results and comparison for standard Baboon image is also shown in Figure 44 – Right.

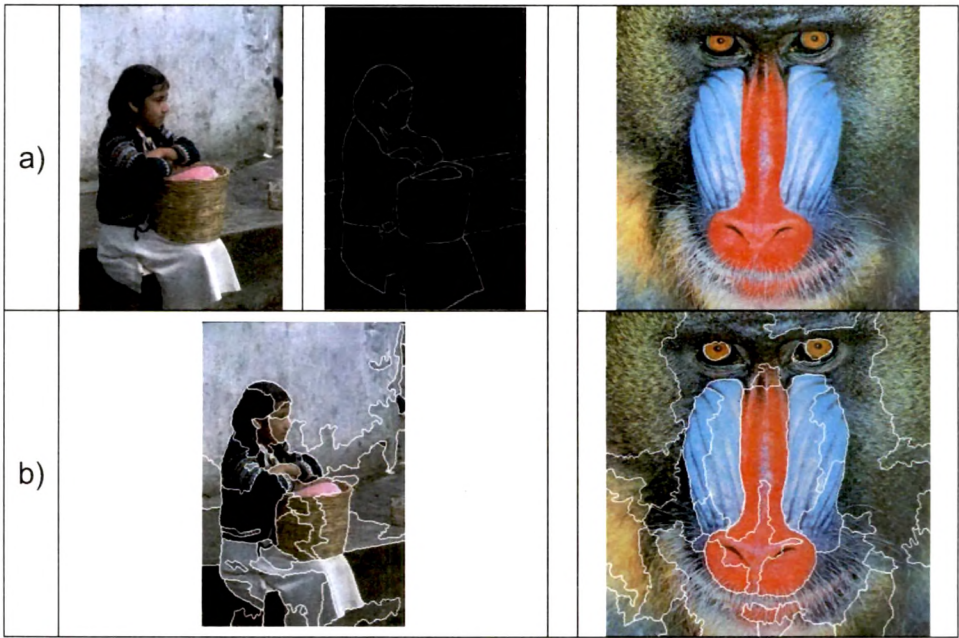


Figure 44. Comparison of segmentation results. (a) Left & Middle - Original and Human segmented Images respectively BSDB [Fowlkes, on line] [Martin, 2001]. (a) Right – Original Baboon image. (b) Segmentation with JSEG [Deng, on line] [Deng, 2001].

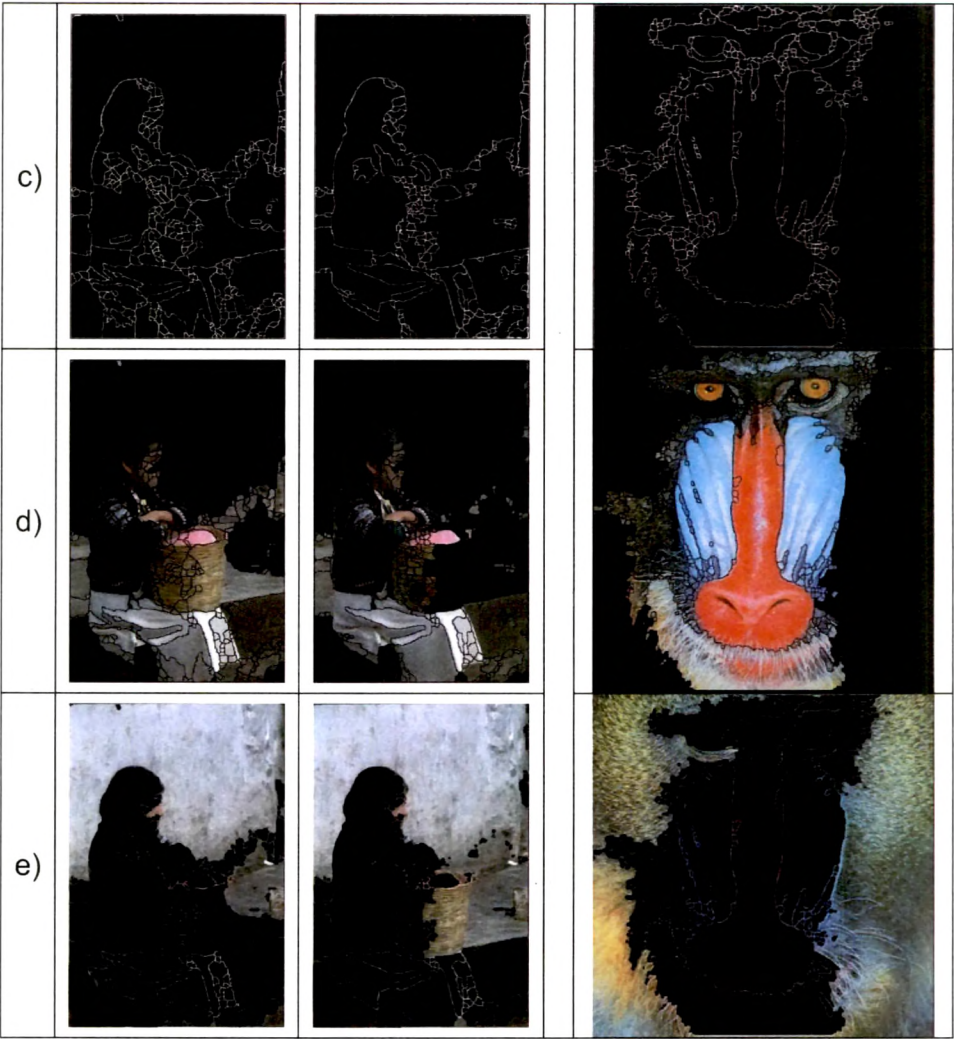


Figure 44 (Contd.). Comparison of segmentation results.. (c) Left & Middle - Watershed pixels with Stationary Haar decomposition at level 1 and level 2 respectively. (c) Right - Watershed pixels with Stationary Haar decomposition at level 2. (d) Extracted foreground. (e) Background.

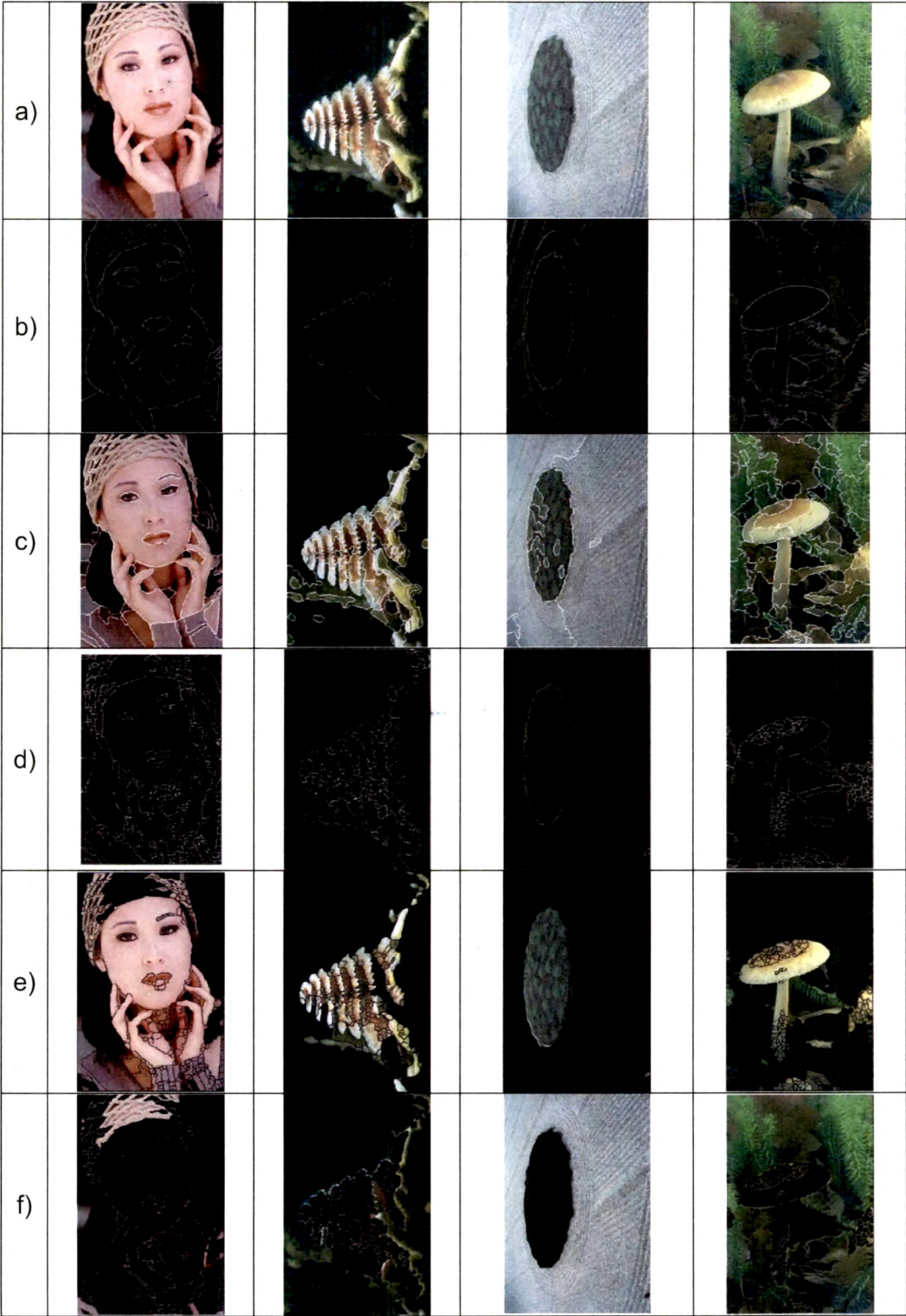


Figure 45. Comparison of segmentation results. (a) Original images BSDB [Fowlkes, on line] [Martin, 2001]. (b) Human segmented Images BSDB [Fowlkes, on line] [Martin, 2001]. (c) Segmentation with JSEG [Deng, on line] [Deng, 2001]. (d) Watershed pixels with Stationary Haar decomposition at level 2. (e) Extracted foreground. (f) Background.

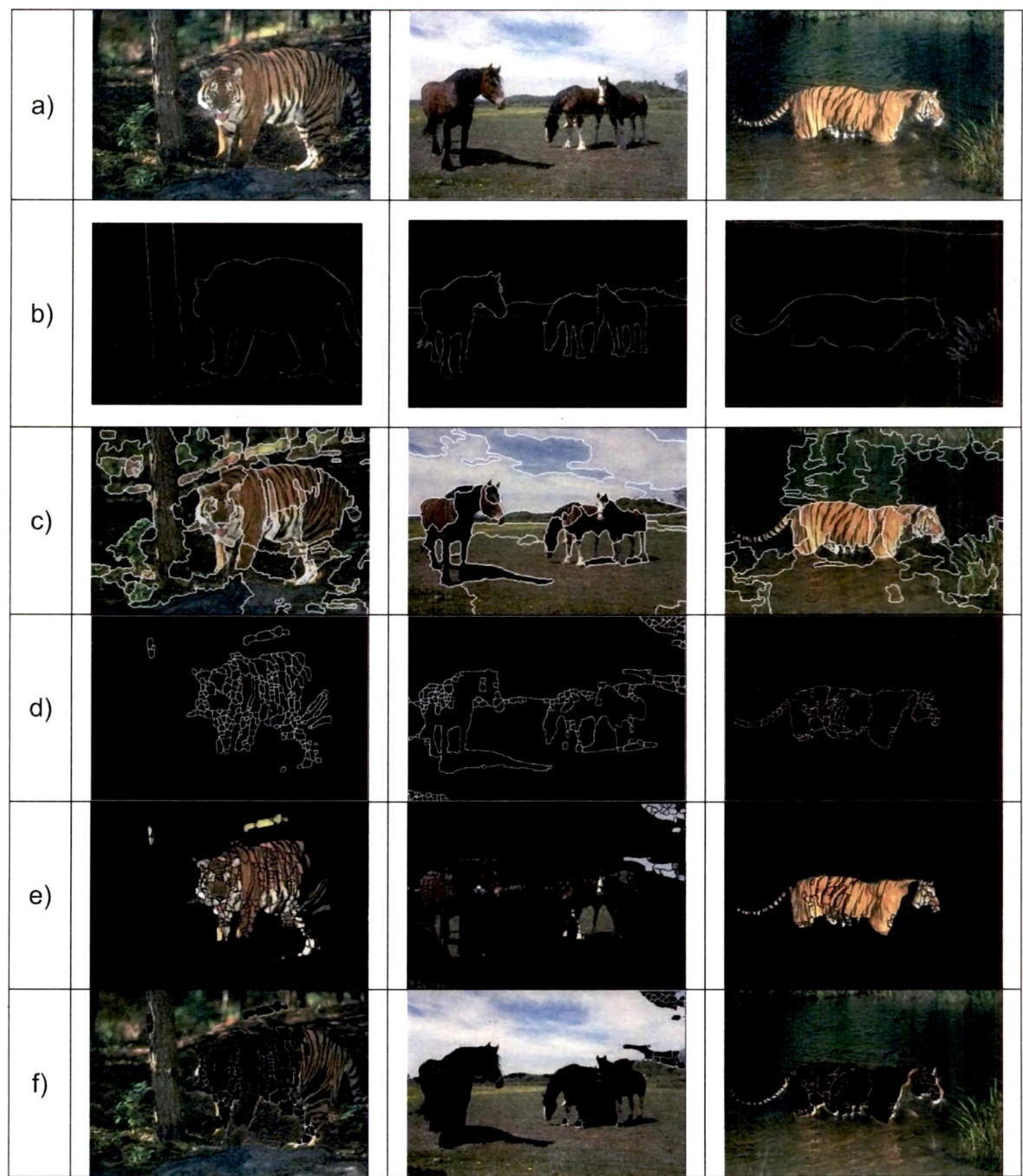


Figure 46. Comparison of segmentation results. (a) Original images BSDB [Fowlkes, on line] [Martin, 2001]. (b) Human segmented Images BSDB [Fowlkes, on line] [Martin, 2001]. (c) Segmentation with JSEG [Deng, on line] [Deng, 2001]. (d) Watershed pixels with Stationary Haar decomposition at level 2. (e) Extracted foreground. (f) Background.

5.4 Discussion

- The proposed method addresses the issue of proper segmentation by enforcing reliable processing of low level cues for avoiding breaks as well as under segmentation by utilizing continuity preserving, well localized visually prominent boundaries for foreground – background separation. The problem of over segmentation is overcome by compositely considering proximity influence and watershed algorithm.
- The method results are tested on variety of images including those of natural images, synthesized images, human faces etc. with diversified textures. The effectiveness of the method is proved for low as well as high resolution images and for size- reduced images by a large factor. The artifacts in watershed regions are significantly reduced, producing less number of small sized regions.
- The foreground extraction results are largely insensitive to illumination variations produced by changes of intensity and spatial locations of the light sources. Thus, the proposed method addresses the issue to the remark **“There seems to be no easy solution”** cited by Deng et al. in [Deng, 2001].
- The watershed artifacts around some portion of boundaries do not affect the foreground boundaries.
- The detected foreground object boundaries are well-localized and well-delineated. The precise processing leading to detection of prominent boundaries & enclosing boundaries of finer prominent details of even small objects is well illustrated in Figure 42.
- The stationary Haar wavelet decomposition at various levels makes the approach suitable for multi-scale hierarchical image segmentation for effective foreground separation and region features extraction. Stationary Haar decomposed image at higher level delineates objects having perceptually superior prominent boundaries. The results of objects revealing for images with poorly defined prominent boundaries are better with exclusion of Haar decomposition step of the method, as illustrated in Figure 35 - Left.

- The results of background separation revealing foreground objects are not affected due to inter- texture and intra-texture variations as presented in Figure 34 to Figure 41.
- For some cases, regions attached to objects of foreground alter the shape of foreground region.
- The extracted region-features and shape features may be utilized for object identification, content based image retrieval, automatic image tagging, and visual scene analysis.
- The proposed method has been tested on a set consisting of about 400 images covering representative images of different categories possessing a wide variety of salient characteristics. The produced results are effective for benchmark database images of Berkeley Segmentation Dataset images [Fowlkes, on line] [Martin, 2001] and SIMPLcity database images [Wang, 2001] [SIMPLcity, on line] and are comparable with human segmented images of BSDb [Fowlkes, on line] [Martin, 2001].
- The quantitative comparisons of the results of proposed method for foreground extraction have been carried out with performance measures $Precision_{fg}$ and $Recall_{fg}$, computed with respect to Ground Truth foreground and presented in Annexure 4 endorses the uniqueness & effectiveness.

5.5 Concluding Remark

Precise processing for detection of prominent boundaries consequently leading to exclusion of background reveals foreground and associated features for image retrieval...

6. Image Retrieval

6.1 Introduction

The chapter deals with image features, characteristics of image databases used and proposed methods for image retrieval. The image retrieval has been carried out on the basis of:

- Color codes of entire image
- Foreground color codes
- Foreground shape correlation
- Combination of foreground color codes and shape correlation
 - With selectable percentage proportion of weight of foreground color codes and foreground shape correlation for composite similarity measure
- Similar face – images containing complex background

The performance of retrieval methods has been evaluated with Precision, Recall, F – measure and P – R curves for various images of different classes and categories. Query responses of some example-images have also been presented. The method and results of face extraction and similar-face-image retrieval are covered lastly in the chapter.

The retrieval of similar images needs to meet two extreme requirements – (i) Retrieve only similar images and (ii) Retrieve as many as possible similar images. The performance indicator for the first requirement is Precision whereas that of the second is Recall. ([Section 3.3](#)). An attempt to improve Precision by either incorporating stringent features or strict image similarity comparison-measures for excluding dissimilar images ends up in exclusion of similar images as well, adversely affecting the Recall. On other hand, an attempt to increase the number of retrieved similar images leading to improvement in Recall by incorporating broader features or relaxed similarity-measures ends up in retrieving more number of similar images along with dissimilar images, adversely affecting Precision. Thus a 'good' CBIR system should retain maximum possible Precision for higher Recall for large image databases consisting of variety of

images. [Section 2.8](#) - Our Observations and [Section 2.9](#) - Our Approaches for image retrieval may be referred for details.

The proposed methods for image retrieval are based on two streams of techniques. The first stream follows broader image descriptors – Color codes and corresponding histograms. Whereas, the second stream follows reliable processing leading to precise detection of prominent boundaries eventually revealing foreground. Image retrieval based on whole image comparison has been carried out with broader image descriptors for retrieving images having similar color code distribution. The other methods are based on foreground of the images. The extracted foreground has been utilized for comparison of objects contained in it. The correlation coefficients have been used as similarity comparison-measure for matching shape of the foregrounds. The other three methodologies combine aforesaid two streams of techniques for image retrieval. The first combinational technique is based on similarity comparison of foreground color codes. The second combinational technique is based on color-codes and shape of the foreground with selectable proportion of weight of shape & color-code in composite similarity-measure. And the third one is the combinational technique applied for application specific CBIR for similar-face image retrieval.

Thus, our approaches for image retrieval techniques exploits reliable processing resulting precise features for better Precision and broader image features for better Recall. *The methods are novel for the extracted features and their use for image retrieval.*

6.2 Image Features

Figure 47 shows the block diagram for extraction of features for the developed CBIR system. Parameters given as input includes selected image name for single image processing or folder name for bulk processing and wavelet decomposition level.

The extraction of image features can be carried out either for one image or for all images stored in a folder for supporting addition of an image into existing image database and to facilitate bulk processing respectively. The color code features of image can be separately extracted. The option of extracting all features take out all features - prominent boundaries based and color code based. The extracted features stored in appropriate data structures are preserved in secondary storage as a file corresponding to the image. The unknown dimension of various extracted features (i.e.

number of contours, number of vertices in contours) and their processing enforced exhaustive applications of programming-skills & utilization of Matlab-features.

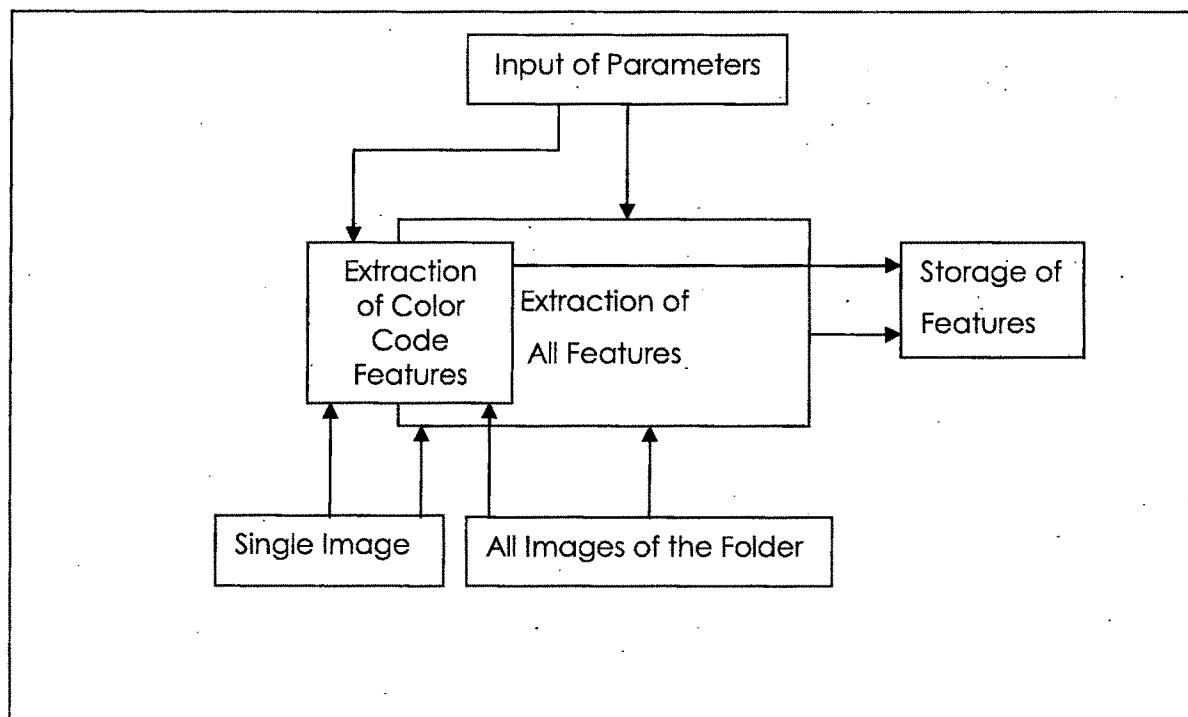


Figure 47. Block diagram - Feature extraction.

The extracted image features listed in Table 5, can be categorized as

- Image attributes – Name (Path), type & dimension.
- Color features - Color codes, Histograms
- Boundaries based features - Regional attributes , edges
- Foreground related attributes – boundaries based and color features for foreground
- Face region features

Table 5. Extracted features.

Sr. No.	Features
1	Path of the image.
2	No of rows of the image.
3	No of column of the image.
4	Image type.
5	Normalized global histogram of the image. For R G and B channels. Not utilized at present.
6	Normalized cumulative global histogram of the image. For R G and B channels. Not utilized at present.
7	Color codes.
8	Normalized histogram of color codes.
9	Normalized cumulative Histogram of color codes. Not utilized at present.
10	Thinned edges.
11	Foreground regions of the image.
12	Background regions of the image.
13	Watershed pixels.
14	Composite prominence measure.
15	Labeled regions.
16	Labeled regions converted to RGB image.
17	Image category. Presently describing 'face' or 'unknown'.
18	Extracted face.
19	Regions corresponding to color codes of whole image.
20	Sorted histogram of color codes of whole image.
21	Regions indices, sorted for region area. Regions are found following prominent boundary based approach.
22	Regional attributes for regions of whole image. Area, Centroid, Bounding box, Extrema, Extent, Solidity (given by Area/ConvexArea), Eccentricity, Convex hull, Minor axis length, Major axis length, orientation.
23	Normalized Un-segmented foreground region of the image.
24	Un-segmented foreground of the image.

Table 5 (Contd.), Extracted features.

Sr. No.	Features
25	Normalized global histogram of foreground of the image. For R G and B channels. Not utilized at present.
26	Normalized cumulative global histogram of foreground of the image. For R G and B channels. Not utilized at present.
27	Color codes for foreground.
28	Normalized histogram of color codes of foreground.
29	Normalized cumulative Histogram of color codes of foreground. Not utilized at present.
30	Regions corresponding to color codes of foreground.
31	Sorted histogram of color codes of foreground.
32	Ratio of two axes of Extrema corresponding to normalized face region. Not utilized at present.
33	Ratio of other two axes of Extrema corresponding to normalized face region. Not utilized at present.
34	Ratio of Minor axis length to Major axis length of the ellipse that has same normalized second central moments as the face region.
35	Orientation of face region. Angle in degree between x-axis and the major axis of the ellipse that has same second moments as the face region.
36	Average of foreground region color code.
37	Ratio of foreground area to image area indicating percentage contribution of foreground region in the image.
38	Two additional extra feature fields Incorporated for future needs.

The color codes are used to describe color attribute of pixels of images. A color code represents a set of colors of RGB color space. Total of 27 color codes are used to represent entire range of RGB color space. The pixels assigned with color codes effectively segments the image by forming regions consisting of pixels with same color - codes. Figure 48 and Figure 49 (a) illustrate the color codes assigned to pixels resulting into segmentation of image. The labeled regions of identical color (same color – codes) are shown with same color in the segmented images. Figure 49 (b) Left is a

separated region of colors corresponding to flowers whereas Figure 49 (b) Right is for the regions corresponding to green leaves.

Refer Annexure 4, Section A-4.4 for details of proposed novel color codes and results of segmentation applied on images of standard databases of BSDb [Fowlkes, on line] [Martin, 2001] and SIMPLicity [Wang, 2001].



Figure 48. Color – code based segmentation.

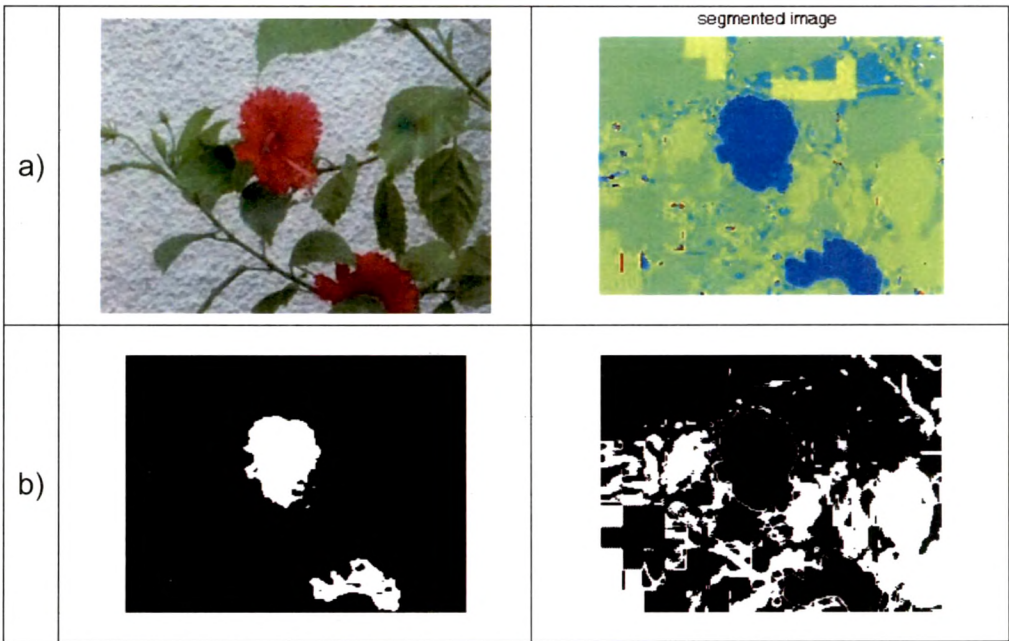










































Figure 49. Color – code based segmentation & regions corresponding to two color-codes.

6.3 SIMPLicity Image Database [Wang, 2001] [SIMPLicity, on line] - Classes & Characteristics

The dataset consists of total 1000 images of 10 classes and 100 images per class. Images are of medium resolutions and reasonable size. The test set has used by many researchers for the purpose of CBIR. Table 6 describes characteristics of image classes.

Table 6. SIMPLIcity Image Database [Wang, 2001] [SIMPLIcity, on line] - Classes and Characteristics

Sr. No.	Class Name	Characteristics	Sample Images	
0	Tribal people	Single or group of tribal-inhabitants in different poses with different backgrounds; Colored faces; typical tribal-dressing;		
				
1	Seashore	Sea, sand, sky with clouds; sea-shore objects; Images cover distant objects;		
				
2	Sculpture	Majority of sculpture images captured as distant objects having sky as background;		
				
3	Bus	Mostly single bus, covering major portion of images; Differently colored various types of buses with different backgrounds.		
				
4	Dinosaur	Different types of dinosaurs with non-textured multi-colored backgrounds;		
				
5	Elephant	Single or multiple elephants in different backgrounds;		
				
6	Flower	Different types and colored flowers covering major portion of images;		
				
7	Horse	Single or multiple horses in different backgrounds;		
				
8	Mountain	Mountains and sky		
				
9	Served Food on Restaurant-table	Typical images of different types of served food on restaurant-tables;		
				

6.4 ALOI Image Database [ALOI, on line] [Geusebroek, 2001]

Amsterdam Library of Object Images (ALOI) provides a collection of color images of one-thousand small objects, recorded for scientific purposes. Over one hundred images per object were captured systematically in controlled conditions for varied viewing angles, illumination angles and illumination colors for the sensory variation in object recordings. The images were captured by one of five lights turned on for 15 different illumination angles, 12 different illumination color configuration and 72 object view point variations. Produced smooth variations in intensity and shadows in different directions and parts of wide variety of objects offer a comprehensive test set for studying segmentation and feature extraction issues. Figure 50 shows few sample-images of an object recorded with such variations.

The test set consisting of some of the images of ALOI [ALOI, on line] [Geusebroek, 2001] database has been used as one of the databases for image retrieval techniques for studying effects of aforesaid variations on foreground shape and foreground color codes.

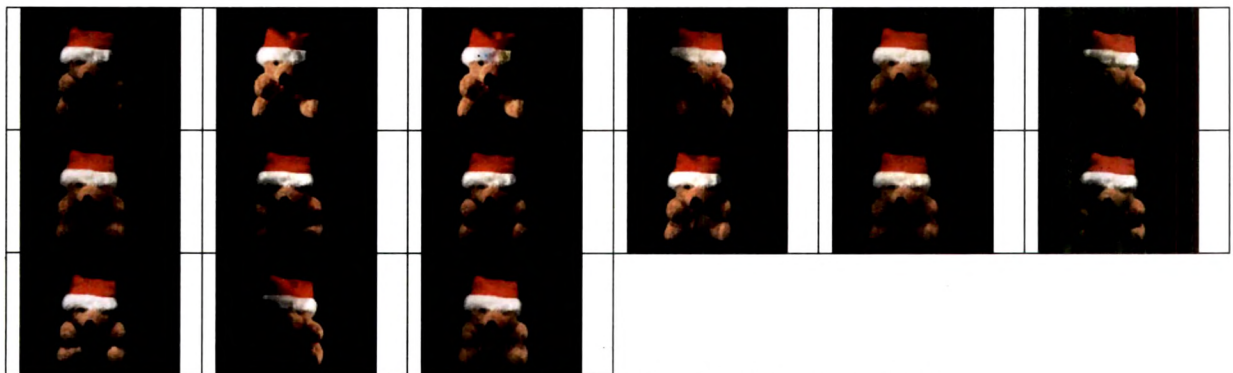


Figure 50. ALOI sample images [ALOI, on line] [Geusebroek, 2001].

6.5 Proposed Techniques

The proposed techniques, based on Color codes of entire image, Foreground color codes, Foreground shape correlation and Combination of foreground color codes & shape correlation follow steps mentioned below for image retrieval. Like all CBIR techniques, incorporated image features and method for computation of similarity measures differentiate methods and their respective performances for retrieval of images. The proposed methods consist of three phases – (i) Input reading (ii) method specific image-feature reading & processing for similarity measures and (iii) output / presentation. The method-specific Step 3 and Step 4 are presented in respective sections of proposed methods. The generic steps for proposed methods are:

Step 1: Read - name of selected query image, name of selected target folder.
and image to be searched for. Perform validations for inputs.

Step 2: Read names of files containing all image features.

Step 3: Read (or extract) required image-features for the query image.

Step 4: For every image-feature-file of target folder,

Read corresponding image-features of the image-feature-file of target folder

Calculate (dis)similarity_index_i for i^{th} image of the database.

Preserve path of data base image, needed for display.

Step 5: Read similarity cut-off.

Step 6: Sort calculated (dis)similarity indices in descending order.

Step 7: Count no of images having better similarity index than the similarity cut-off. Prepare for proper presentation of results.

Step 8: Display all similar images having better similarity index than the similarity cut-off, in order of decreasing similarity (from left to right, row wise).

Algorithm 4. Generic steps for proposed image retrieval methods

The GUI (Annexure 2) based developed CBIR system runs on a stand-alone machine. The database can be expanded by adding images into any folder which can be processed subsequently for feature extraction at once with a mouse click. In absence of indexing mechanism, features are stored in files. A user will be prompted to carry out feature extraction of a query image, if not done earlier with help of the GUI for that single image. The preprocessing of a query image is adopted by considering the high time complexity of algorithms and repetition in the experimentations. For real time deployment of the system, query preprocessing can be eliminated.

Similarity cut-off selection is to be carried out by user with GUI. Lower similarity cut-off signifies higher permitted dissimilarity in the retrieved images.

The image-query response gets presented in a Matlab Figure-window containing a grid of 4 x 4 thumbnail-images along with their storage path.

6.6 Whole Image Color Codes Based CBIR

The proposed method for image retrieval enables user to perform search on the basis of color attributes of entire image. The color code assigned to a pixel is designated for broadly describing the color of pixel. Being a broader descriptor, a color code can accommodate pixel color variations without affecting corresponding color

feature. The proposed technique compares normalized global histogram of color codes constructed for entire query image with that of images of the database for measuring color distribution similarity. The steps involved in reading of image features, their processing and computation of similarity index are shown below. These method specific steps replace corresponding generic Steps 3 & 4 of [Algorithm 4, Section 6.5](#).

The method specific steps are:

Step 3: Read (extract) whole image color code features of given query image.

Step 4: For every image-feature-file of target folder,

Read corresponding whole image color code features of the image-feature-file of target folder.

Calculate (dis)similarity_index_i = $\sum \text{abs}(h_{qj} - h_{ij})$, for $1 < j \leq \text{number of bins}$,

Where,

h_{qj} indicates j^{th} bin of normalized histogram of color codes for the query image.

h_{ij} indicates j^{th} bin of normalized histogram of color codes for i^{th} image of database .

Preserve path of data base image, needed for display.

Algorithm 5. Whole image color codes based image retrieval.

6.6.1 Performance Evaluation

The performance of the method has been tested on image database of SIMPLcity [Wang, 2001] [SIMPLcity, on line] consisting of 1000 images. Exhaustive performance evaluation has been carried out for four classes of database – Bus, Horse, flower and dinosaur. Recall, Precision and F –measure are computed for sample queries of each class of images. The performance measures for different similarity cut-offs have been tabulated for comparison. Average Recall, Average Precision and Average F – measures for the class are tabulated and plotted along with P – R curves for query responses.

6.6.1.1 Query Image Class: Bus

- o The performance evaluation on image database [Wang, 2001] [SIMPLcity, on line] consisting of 1000 images has been shown in Table 7 for
 - 11 varieties of query images [Wang, 2001] [SIMPLcity, on line] for different similarity cut-offs &

- 56 queries
 - The selected query images possess variations in object-poses, number of foreground objects, object-colors, backgrounds and illumination conditions.

Table 7. Precision, Recall & F –measure at different similarity cut-offs. (Whole image color codes). Class - Bus.










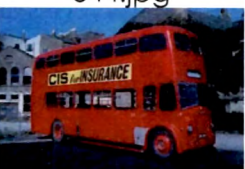

Query Image	Similarity cut-off	Retrieved relevant images - rr	total retrieved images - total	Total relevant images in the database - Total	Recall $r = rr / \text{Total}$	Precision $p = rr / \text{total}$	F measure = $2 / (1/p + 1/r)$
300.jpg 	25	53	82	100	0.53	0.64	0.58
	30	45	65		0.45	0.69	0.54
	40	25	31		0.25	0.81	0.38
	50	13	16		0.13	0.81	0.22
	60	4	6		0.04	0.66	0.08
	70	1	1		0.01	1.00	0.02
310.jpg 	25	55	70	100	0.55	0.79	0.65
	30	46	54		0.46	0.85	0.60
	40	30	33		0.3	0.91	0.45
	50	19	19		0.19	1.00	0.32
	60	7	7		0.07	1.00	0.13
358.jpg 	25	68	99	100	0.68	0.69	0.68
	30	60	75		0.6	0.8	0.69
	40	31	37		0.31	0.84	0.45
	50	17	18		0.17	0.94	0.29
	60	7	7		0.07	1.00	0.13
315.jpg 	25	33	105	100	0.33	0.31	0.32
	30	27	80		0.27	0.34	0.30
	40	13	35		0.13	0.37	0.19
	50	6	14		0.06	0.43	0.11
	60	3	5		0.03	0.6	0.06
319.jpg 	25	32	72	100	0.32	0.44	0.37
	30	24	51		0.24	0.47	0.32
	40	15	26		0.15	0.58	0.24
	50	6	10		0.06	0.6	0.11
	60	2	2		0.02	1.00	0.04
326.jpg 	25	27	184	100	0.27	0.14	0.18
	30	22	140		0.22	0.15	0.18
	40	9	61		0.09	0.15	0.11
	50	5	23		0.05	0.21	0.08
	60	1	7		0.01	0.14	0.02

Table 7 (Contd.). Precision, Recall & F –measure at different similarity cut-offs. (Whole image color codes). Class - Bus.

Query Image	Similarity cut-off	Retrieved relevant images - rr	total retrieved images - total	Total relevant images in the database - Total	Recall $r = rr / \text{Total}$	Precision $p = rr / \text{total}$	F measure = $2 / (1/p + 1/r)$
	25	60	174	100	0.60	0.34	0.43
	30	52	132		0.52	0.40	0.45
	40	35	68		0.35	0.51	0.42
	50	11	18		0.11	0.61	0.19
	60	5	8		0.05	0.63	0.09
	25	62	101	100	0.62	0.61	0.61
	30	50	78		0.50	0.61	0.55
	40	33	46		0.33	0.71	0.45
	50	11	17		0.11	0.65	0.19
	60	8	9		0.08	0.89	0.15
	25	26	187	100	0.26	0.13	0.17
	30	19	145		0.19	0.13	0.15
	40	8	88		0.08	0.09	0.08
	50	5	47		0.05	0.10	0.07
	60	5	15		0.05	0.33	0.09
	25	43	54	100	0.43	0.80	0.56
	30	37	41		0.37	0.90	0.52
	40	23	23		0.23	1.00	0.37
	50	15	15		0.15	1.00	0.26
	60	5	5		0.05	1.00	0.10
	25	48	138	100	0.48	0.35	0.40
	30	40	102		0.40	0.39	0.39
	40	20	46		0.20	0.43	0.27
	50	8	13		0.08	0.62	0.14
	60	4	5		0.04	0.80	0.08

The average Recall and average Precision for the query image class bus for different similarity cut-offs have been tabulated in Table 8. Corresponding P – R curves of sample queries of the table along with average Precision & average Recall are shown in Figure 51. Average Precision, average Recall and average F-measures for different similarity cut-offs for the query images of Table 8 has been plotted in Figure 52.

Table 8. Average Recall, Average Precision & Average F -measure. (Whole image color codes). Class – Bus.

Similarity cut-off	Average Recall	Average Precision	Average F measure = 2 / (1/Avg.p + 1/Avg.r)
25	0.46	0.48	0.47
30	0.38	0.52	0.44
40	0.22	0.58	0.32
50	0.11	0.63	0.18
60	0.05	0.73	0.09
70	0.01	1.00	0.02

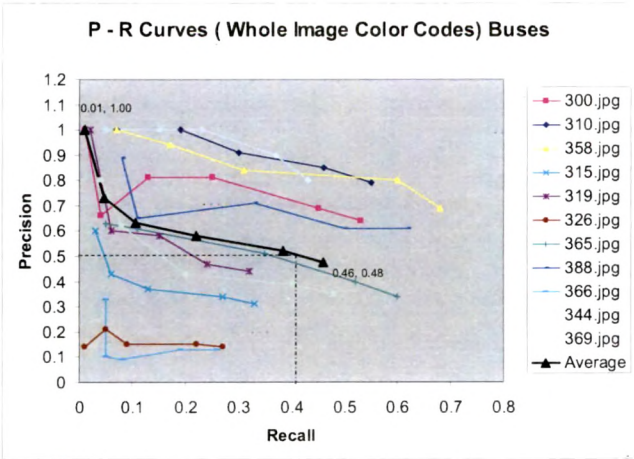


Figure 51. P- R curves (whole image color codes). Class - Bus.

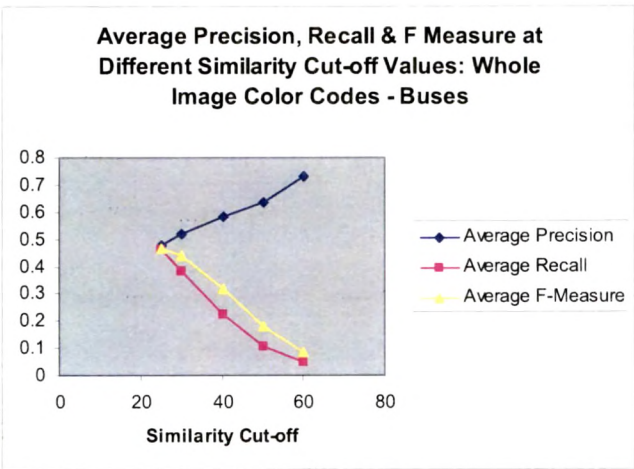


Figure 52. Average Precision, Average Recall, Average F – measures verses Similarity cut-offs. (Whole image color codes). Class – Bus.

Following points are observed:

- The nature of obtained P – R curves matches with the practical P – R curves.
- Stricter similarity cut-off increases the Precision at the cost of Recall.

- Despite vast variations in bus colors, background, poses and illumination conditions, high recall with good precision is achievable for many sample queries.
- For many queries, Precision of 1.0 is achieved.
- The Precision and recall are poor for query images having orange / yellow colored buses as other image-classes contain images of similar color distribution.
- Range of average performance measures for the class
 - 100 % of average Precision for 1 % of average Recall
 - 48 % of average Precision for 46 % of average Recall
 - Giving 52 % of fall in average Precision to raise average Recall by 45 %
- The value of average Recall at average Precision of 0.5 is 0.41 – a reasonably good performance measures.
- Increase in the user selected similarity cut-off indicates higher threshold for the similarity measures for retrieving very similar images. As a result, average Recall falls down due to elimination of images having lesser similarity and average Precision increases as retrieved images are very similar due to higher cut-off resembling Precision – Recall behavior of any practical CBIR system.

6.6.1.2 Query Image Response Examples: Class – Bus

The query responses of a bus image [Wang, 2001] [SIMPLIcity, on line] at two different similarity cut-offs have been shown in Figure 53 and Figure 54.

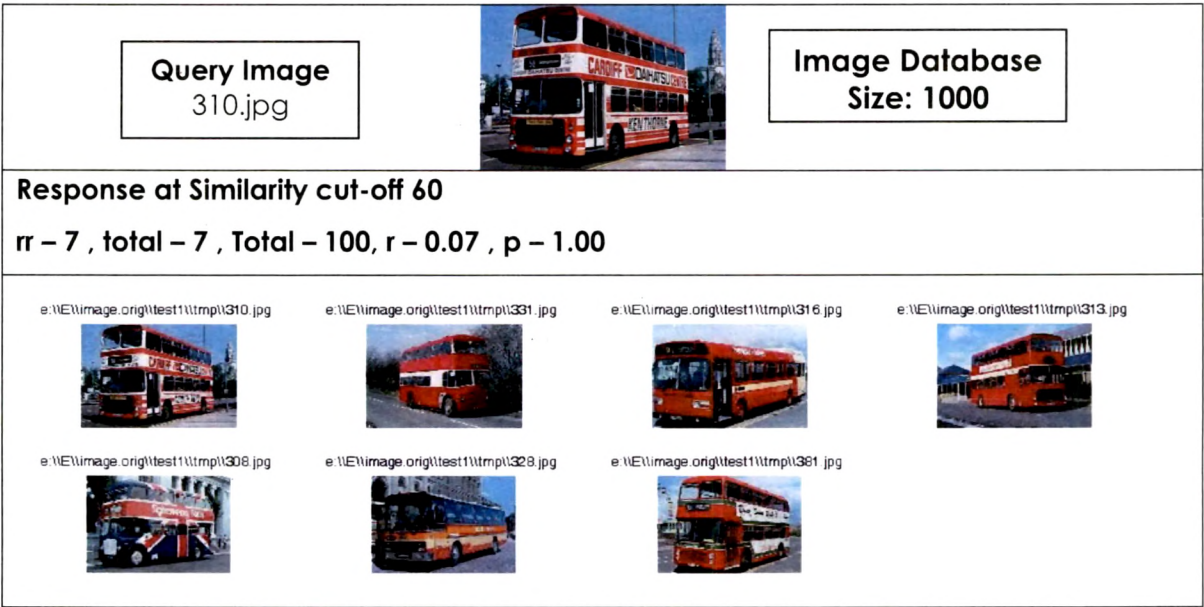


Figure 53. Query response of a bus image at similarity cut-off 60.

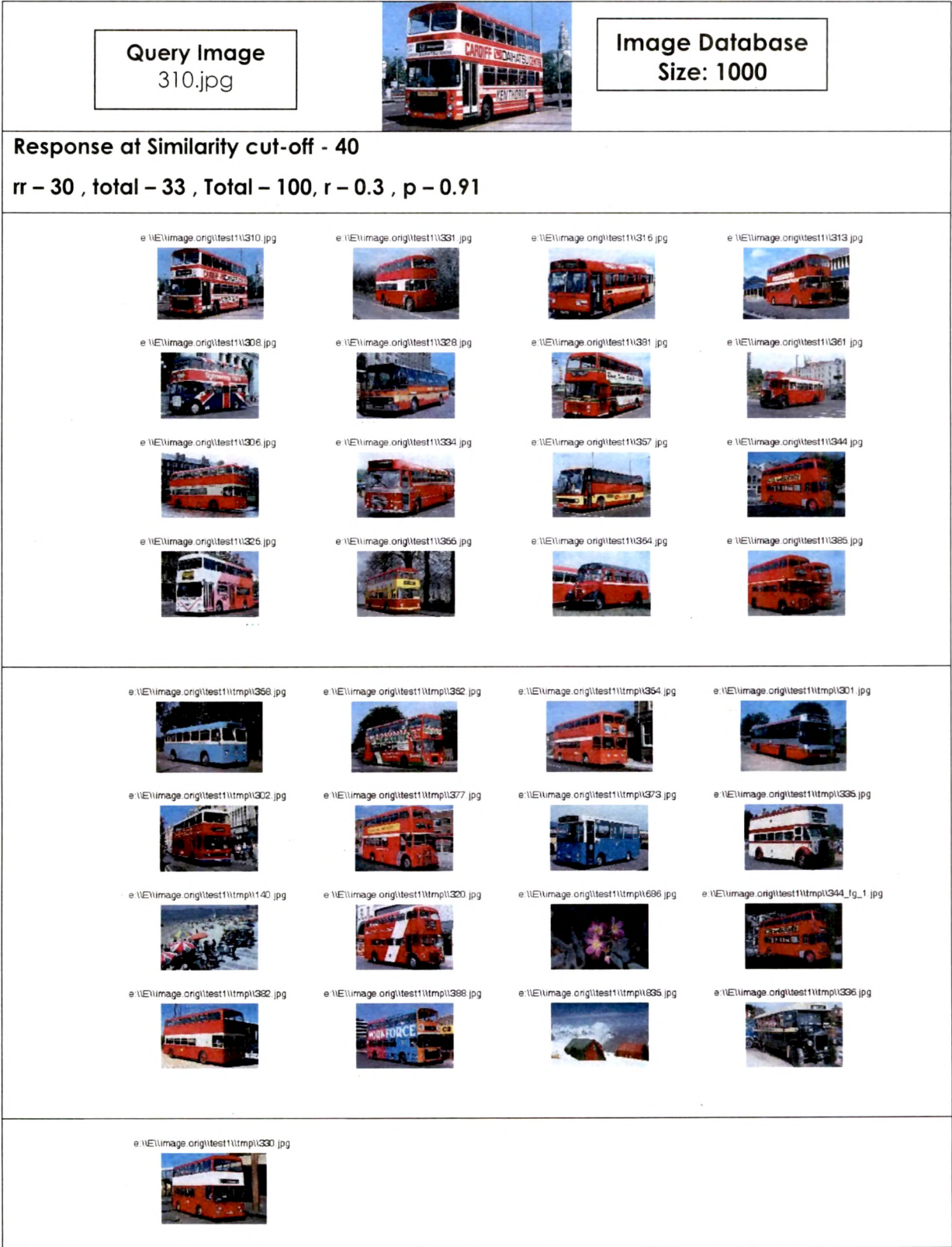


Figure 54. Query response of a bus image at similarity cut-off 40.

6.6.1.3 Query Image Class: Horse

- The performance evaluation on image database [Wang, 2001] [SIMPLcity, on line] consisting of 1000 images has been shown in Table 9 for
 - 6 varieties of query images [Wang, 2001] [SIMPLcity, on line] for different similarity cut-offs &
 - 29 queries
- The selected query images possess variations in object-poses, number of foreground objects, object-colors, backgrounds and illumination conditions.
- The foreground objects constitute relatively lesser percentage-portion of the image compared to image class bus.
- Relatively less variations in the background color distributions among images of the class.

Table 9. Precision, Recall & F –measure at different similarity cut-offs. (Whole image color codes). Class - Horse.





Query Image	Similarity cut-off	Retrieved relevant images - rr	total retrieved images - total	Total relevant images in the database - Total	Recall $r = rr / \text{Total}$	Precision $p = rr / \text{total}$	F measure $e = 2 / (1/p + 1/r)$
700.jpg 	25	53	92	100	0.53	0.57	0.55
	30	51	81		0.51	0.63	0.56
	40	34	56		0.34	0.61	0.44
	50	17	27		0.17	0.63	0.27
	60	7	13		0.07	0.54	0.12
	70	3	5		0.03	0.6	0.06
725.jpg 	25	71	96	100	0.71	0.74	0.72
	30	69	93		0.69	0.74	0.72
	40	61	73		0.61	0.84	0.71
	50	48	54		0.48	0.89	0.62
	60	27	36		0.27	0.75	0.40
	70	13	15		0.13	0.87	0.23
744.jpg 	25	33	53	100	0.33	0.62	0.43
	30	26	42		0.26	0.62	0.37
	40	14	21		0.14	0.67	0.23
	50	7	11		0.07	0.64	0.13

Table 9 (Contd). Precision, Recall & F –measure at different similarity cut-offs. (Whole image color codes).
Class - Horse.

Query Image	Similarity cut-off	Retrieved relevant images - rr	total retrieved images - total	Total relevant images in the database - Total	Recall r = rr / Total	Precision p = rr / total	F measure = 2 / (1/p + 1/r)
757.jpg 	25	70	78	100	0.7	0.90	0.79
	30	68	73		0.68	0.93	0.79
	40	54	57		0.54	0.95	0.69
	50	41	41		0.41	1	0.58
	60	23	23		0.23	1	0.37
701.jpg 	25	81	95	100	0.81	0.85	0.83
	30	76	86		0.76	0.88	0.82
	40	65	70		0.65	0.93	0.76
	50	54	55		0.54	0.98	0.70
	60	33	34		0.33	0.97	0.49
	70	14	14		0.14	1.00	0.25
707.jpg 	25	78	91	100	0.78	0.86	0.82
	30	73	83		0.73	1.00	0.84
	40	59	66		0.59	0.89	0.71
	50	46	49		0.46	0.94	0.62
	60	37	37		0.37	1.00	0.54
	70	23	23		0.23	1.00	0.37

The average Recall and average Precision for the query image class horse for different similarity cut-offs has been tabulated in Table 10. The P – R curves for sample queries of the table and corresponding average Precision & average Recall are shown in Figure 55. Average Precision, average Recall and average F-measures for different similarity cut-offs for the query images of Table 10 have been plotted in Figure 56.

Table 10. Average Recall, Average Precision & Average F - measure. (Whole image color codes). Class – Horse.

Similarity cut-off	Average Recall	Average Precision	Average F measure = 2 / (1/Avg.p + 1/Avg.r)
25	0.55	0.76	0.64
30	0.52	0.80	0.63
40	0.41	0.81	0.55
50	0.30	0.85	0.45
60	0.25	0.85	0.39
70	0.13	0.87	0.23

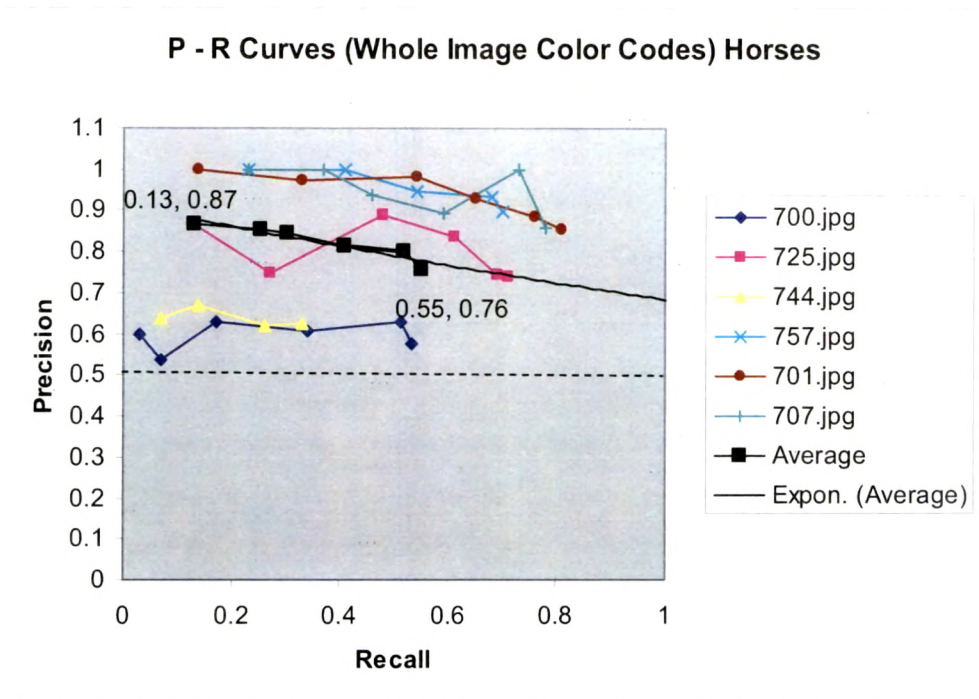


Figure 55. P- R curves (whole image color codes). Class- Horse.

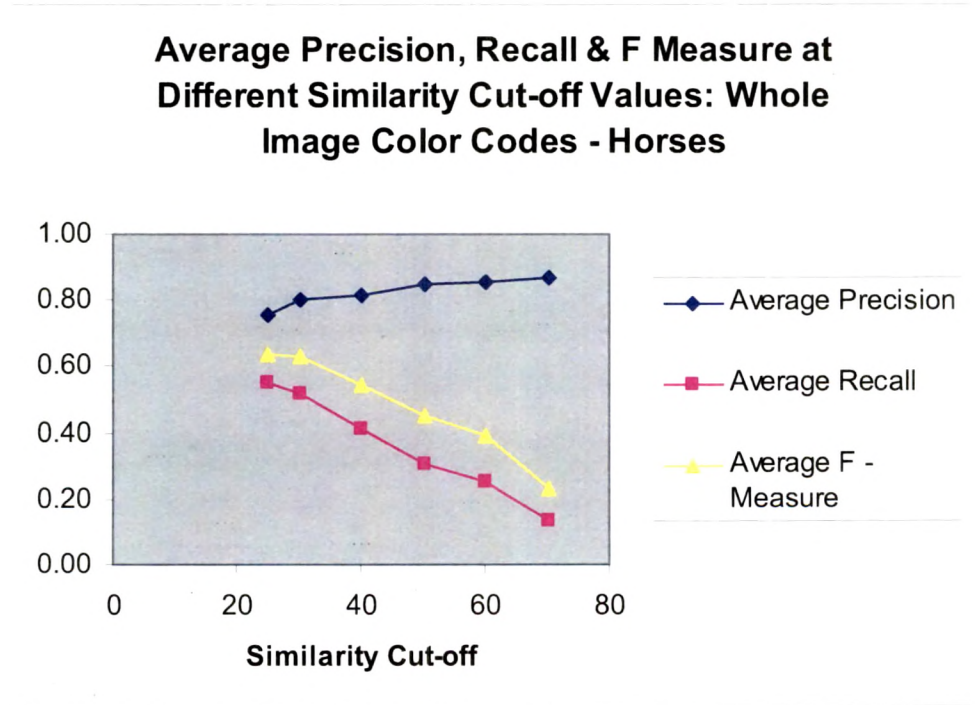


Figure 56. Average Precision, Average Recall, Average F – measures verses Similarity cut-offs. (Whole image color codes). Class – Horse.

Following points are observed:

- The nature of obtained P – R curves is close to ideal P – R curves for some of the query images.
- For the variations in colors & poses of foreground objects, background and illumination conditions, high recall with good precision is achievable for many sample queries.
- Stricter similarity cut-off increases the Precision at the cost of Recall.
- For many queries, Precision of 1.0 is achieved.
- For many queries, Recall greater than of 0.7 is achieved.
- The uniqueness of the background colors combined with foreground colors ends up with color distributions not common in the images of other classes giving very good performance measures.
- Range of average performance measures for the class
 - 87 % of average Precision for 13 % of average Recall
 - 76 % of average Precision for 55 % of average Recall
 - Giving only 11 % of fall in average Precision to raise average Recall by 42%.
- The exponentially extended trend line is well above Precision = (0.5) line and does not intersect average Recall till its maximum possible value, implies more than 50% of average Precision for all average Recall values indicating exceptionally good performance measures.

6.6.1.4 Query Image Response Examples: Class - Horse

The query response of a horse image [Wang, 2001] [SIMPLcity, on line] at similarity cut-off 25 has been shown in Figure 57. The Recall of 71% with Precision of 73% at similarity cut-off of 25 is remarkable, giving 0.72 as F-measure. The query response of another image consisting of two horses has been shown in Figure 58 giving 100 % precision with Recall of 14%.

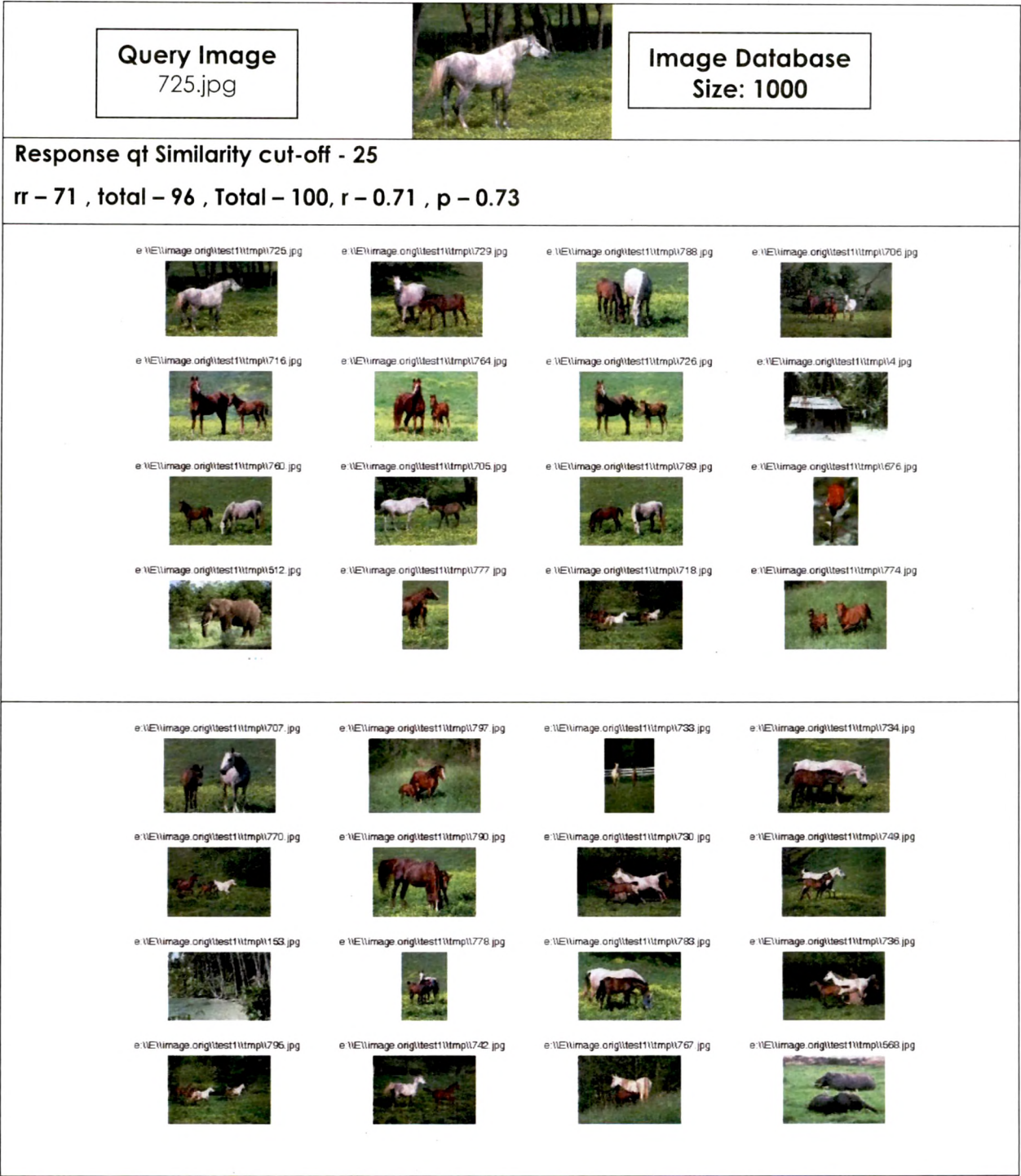


Figure 57. Query response of a horse image at similarity cut-off 25.



Figure 57 (Contd.). Query response of a horse image at similarity cut-off 25.

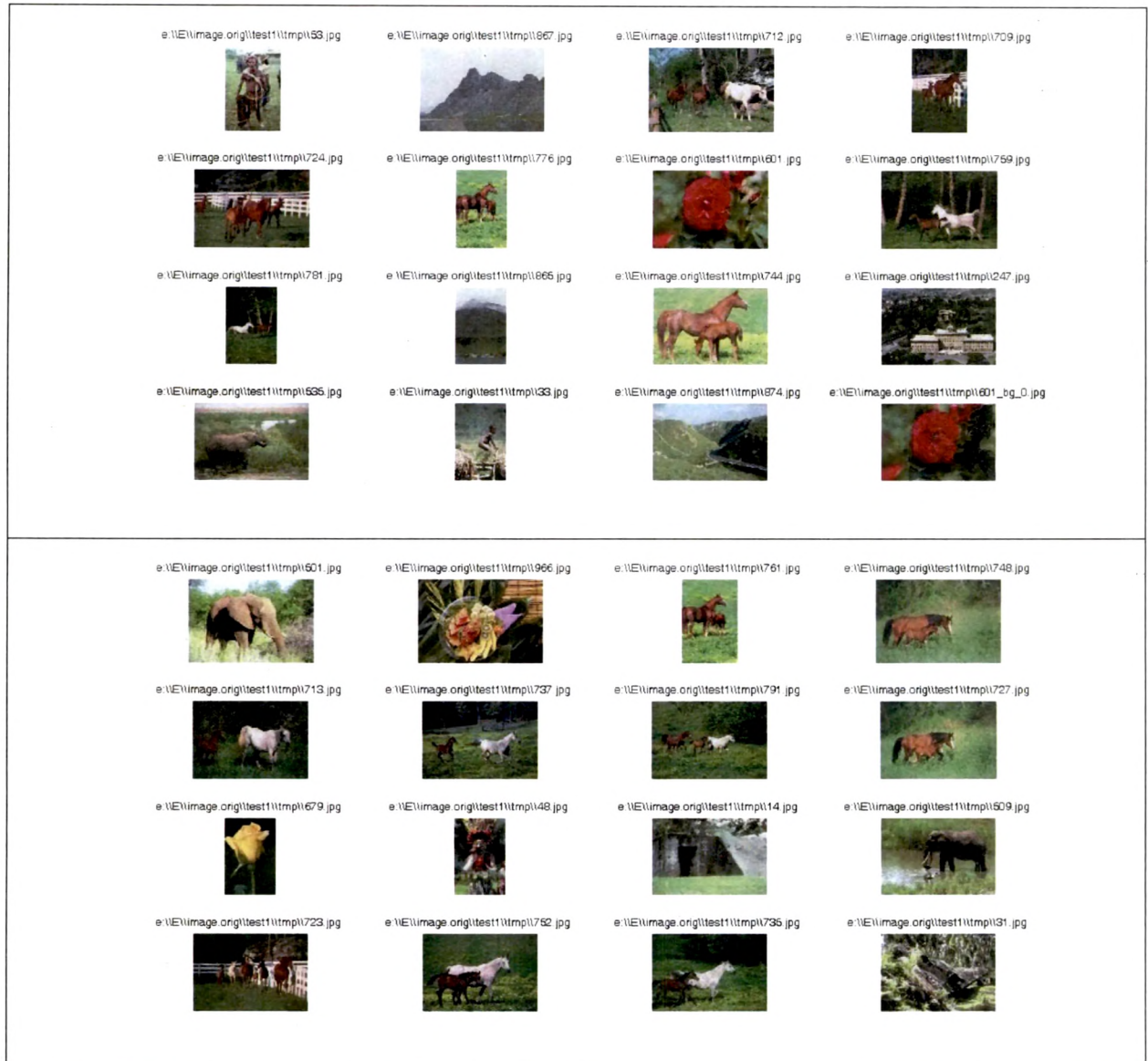


Figure 57 (Contd.). Query response of a horse image at similarity cut-off 25.

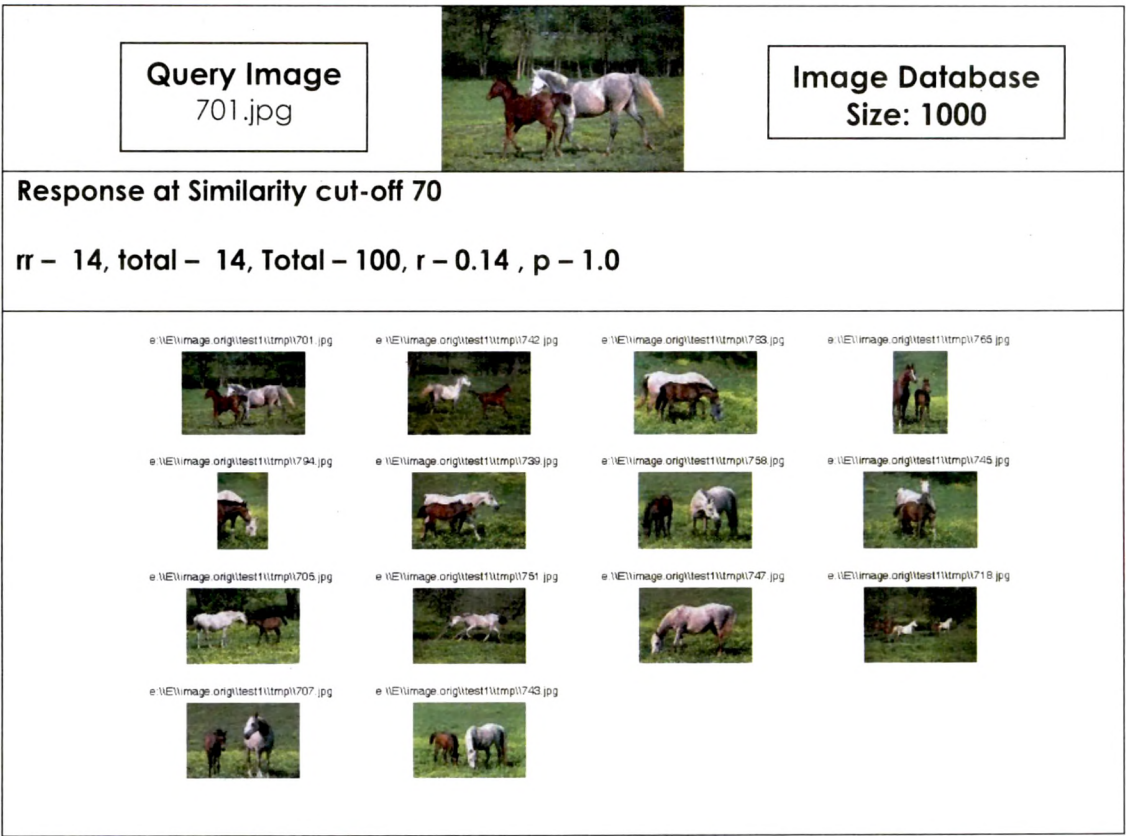


Figure 58. Query response of another horse image at similarity cut-off 70.

6.6.1.5 Query Image Class - Flower

- The performance evaluation on image database [Wang, 2001] [SIMPLIcity, on line] consisting of 1000 images has been shown in Table 11 for
 - 10 varieties of query images [Wang, 2001] [SIMPLIcity, on line] for different similarity cut-offs
 - 44 queries
- The selected query images possess variations in object-poses, number of foreground objects, object-colors, backgrounds and illumination conditions.
- The foreground objects generally constitute significant portion of the image.

Table 11. Precision, Recall & F –measure at different similarity cut-offs. (Whole image color codes). Class - Flower.




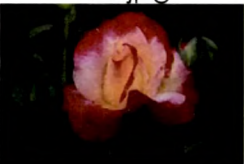


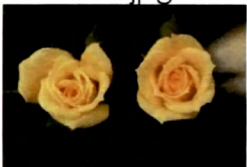



Query Image	Similarity cut-off	Retrieved relevant images - rr	total retrieved images - total	Total relevant images in the database - Total	Recall $r = rr / \text{Total}$	Precision $p = rr / \text{total}$	F measure $= 2 / (1/p + 1/r)$
602.jpg 	25	21	28	57	0.37	0.75	0.49
	30	16	21		0.28	0.76	0.41
	40	13	14		0.23	0.93	0.37
	50	9	9		0.16	1	0.27
	60	3	3		0.05	1	0.10
606.jpg 	25	34	43	57	0.60	0.79	0.68
	30	29	31		0.51	0.94	0.66
	40	17	18		0.31	0.94	0.45
	50	9	9		0.16	1	0.27
	60	2	2		0.03	1	0.07
644.jpg 				24			
	40	14	20		0.58	0.7	0.64
	50	13	15		0.54	0.87	0.67
	60	5	5		0.21	1	0.34
655.jpg 				57			
	40	17	21		0.30	0.81	0.44
	50	8	8		0.14	1	0.25
	60	3	3		0.05	1	0.10
656.jpg 	25	4	6	9	0.44	0.67	0.53
	30	4	4		0.44	1	0.62
	40	1	1		0.11	1	0.20
	50	1	1		0.11	1	0.20
	60	1	1		0.11	1	0.20

Table 11 (Contd.). Precision, Recall & F –measure at different similarity cut-offs. (Whole image color codes). Class - Flower.

Query Image	Similarity cut-off	Retrieved relevant images - rr	total retrieved images - total	Total relevant images in the database - Total	Recall r = rr / Total	Precision p = rr / total	F measure = 2 / (1/p + 1/r)
696.jpg 	25	4	4	57	0.07	1	0.13
	30	3	3		0.05	1	0.10
	40	3	3		0.05	1	0.10
	50	2	2		0.03	1	0.07
	60	1	1		0.02	1	0.03
675.jpg 	40	14	18	24	0.58	0.78	0.67
	50	6	6		0.25	1	0.40
	60	2	2		0.08	1	0.15
682.jpg 	25	9	24	57	0.15	0.38	0.22
	30	5	18		0.09	0.28	0.13
	40	4	11		0.07	0.36	0.12
	50	2	7		0.03	0.29	0.06
	60	1	2		0.01	0.50	0.03
618.jpg 	25	30	35	57	0.53	0.86	0.65
	30	26	28		0.46	0.93	0.61
	40	16	16		0.28	1.00	0.44
	50	7	7		0.12	1.00	0.22
	60	3	3		0.05	1.00	0.10
621.jpg 	25	51	56	57	0.89	0.91	0.90
	30	28	31		0.49	0.90	0.64
	40	18	20		0.32	0.90	0.47
	50	8	8		0.14	1.00	0.25
	60	5	5		0.09	1.00	0.16

The average Recall and average Precision for the query image class flower for different similarity cut-offs has been tabulated in Table 12. The P – R curves for sample queries of Table 11 and corresponding average Precision & average Recall are shown in Figure 59. Average Precision, average Recall and average F-measure for different similarity cut-offs for the query images of Table 12 have been plotted in Figure 60.

Table 12. Average Recall, Average Precision & Average F - measure. (Whole image color codes). Class – Flower.

Similarity cut-off	Average Recall	Average Precision	Average F measure = $2 / (1/\text{Avg.p} + 1/\text{Avg.r})$
25	0.44	0.76	0.55
30	0.33	0.83	0.47
40	0.24	0.84	0.37
50	0.17	0.92	0.29
60	0.07	0.90	0.13

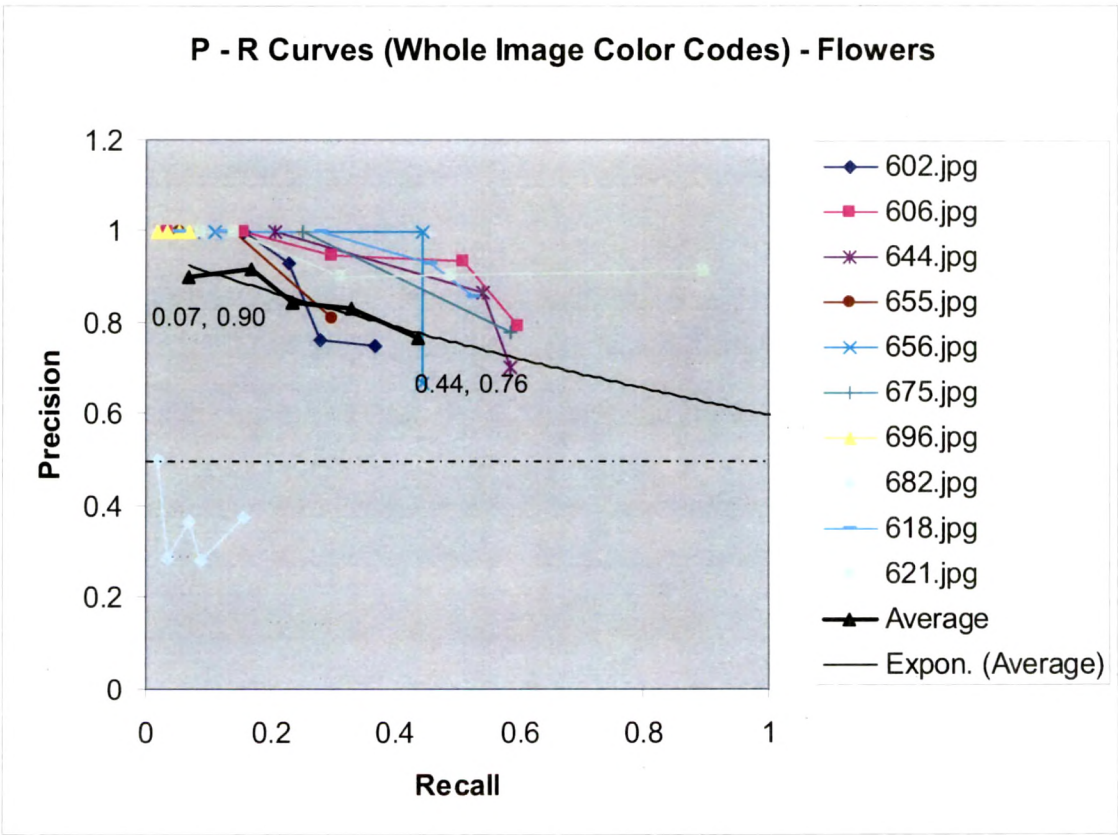


Figure 59. P- R curves (whole image color codes). Class- Flower.

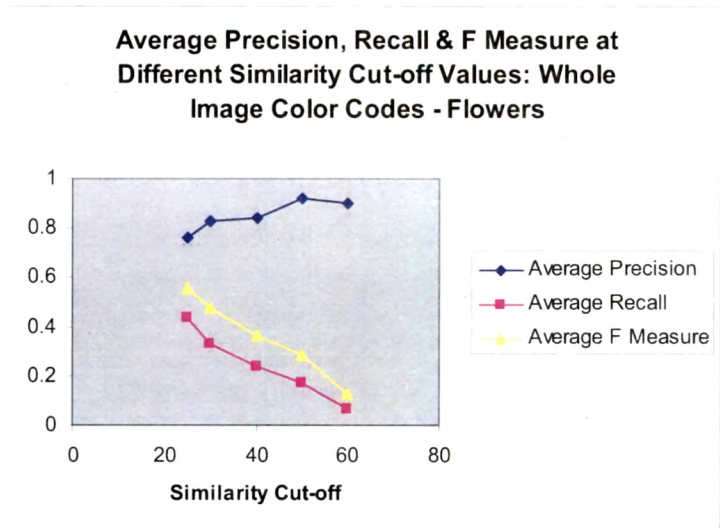


Figure 60. Average Precision, Average Recall, Average F – measures versus Similarity cut-offs. (Whole image color codes). Class – Flower.

Following points are observed:

- The nature of obtained P – R curves is close to ideal P – R curve for some of the query images and similar to practical P – R curves for majority of query images.
- For the variations in colors & poses of foreground objects, background and illumination conditions, high recall with good precision is achieved for many sample queries.
- Stricter similarity cut-off increases the Precision at the cost of Recall.
- For many queries, Precision of 1.0 is achieved.
- Poor Precision and Recall values for image 682.jpg are because of blurred (filtered) background constituting major portion of the image.
- At lower cut-offs, images of colored human faces and served restaurant food also gets retrieved because of similar visual cues (Figure 61).
- Range of average performance measures for the class
 - 90 % of average Precision for average 7 % of Recall
 - 76 % of average Precision for 44 % of average Recall
 - Giving only 14 % of fall in average Precision to raise average Recall by 37%.
- The exponentially extended trend line is well above Precision = (0.5) line and not intersecting till average Recall value of 1, implies average precision above 50% for all average Recall values – good performance measures.

6.6.1.6 Query Image Response Example: Class – Flower

The query response of a flower image [Wang, 2001] [SIMPLcity, on line] having typical foreground colors is shown for similarity cut-off of 40 in Figure 61. The Recall is calculated for total of 57 red / pink colored flower images of the database.

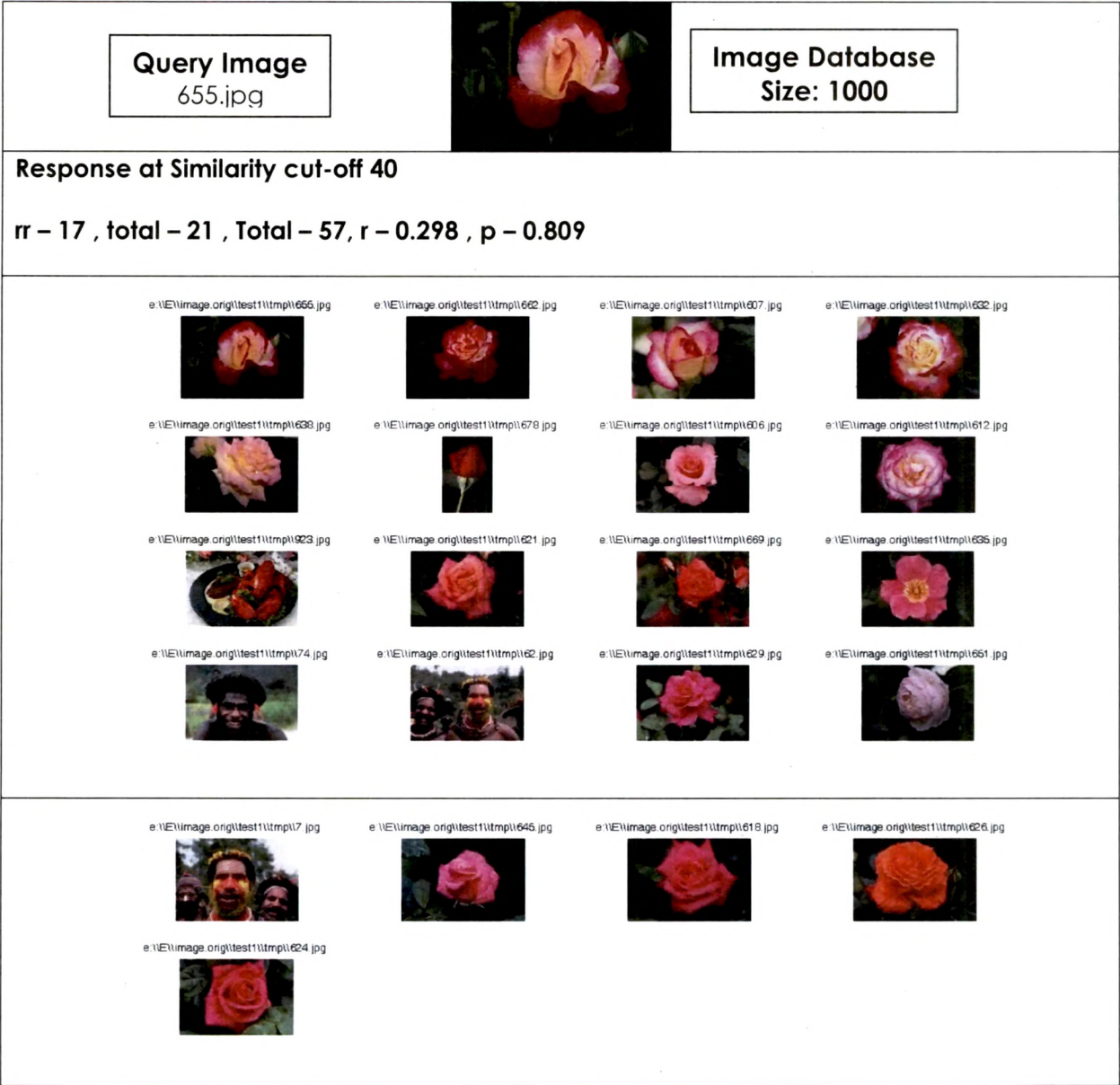


Figure 61. Query response of a flower image at similarity cut-off 40.

6.6.1.7 Query Image Class – Dinosaur

- The performance evaluation on image database [Wang, 2001] [SIMPLcity, on line] consisting of 1000 images has been shown in Table 13 for
 - 6 varieties of query images [Wang, 2001] [SIMPLcity, on line] for different similarity cut-offs &
 - 32 queries
- The selected query images possess variations in object-poses, object-colors.
- The backgrounds are simple, multi-color toned & non-textured.
- The foreground objects constitute relatively lesser percentage-portion of the image compared to image class bus.

Table 13. Precision, Recall & F –measure at different similarity cut-offs. (Whole image color codes). Class – Dinosaur.




Query Image	Similarity cut-off	Retrieved relevant images - rr	total retrieved images - total	Total relevant images in the database - Total	Recall $r = rr / \text{Total}$	Precision $p = rr / \text{total}$	F measure = $2 / (1/p + 1/r)$
406.jpg 	25	42	53	100	0.42	0.79	0.55
	30	35	46		0.35	0.76	0.48
	40	28	34		0.28	0.82	0.42
	50	20	24		0.2	0.83	0.32
	60	12	14		0.12	0.86	0.21
	70	7	7		0.07	1	0.13
408.jpg 	25	52	79	100	0.52	0.66	0.58
	30	48	72		0.48	0.67	0.56
	40	28	40		0.28	0.7	0.40
	50	20	25		0.2	0.8	0.32
	60	8	8		0.08	1	0.15
	70	2	2		0.02	1	0.04
415.jpg 	25	13	21	100	0.13	0.62	0.21
	30	11	13		0.11	0.85	0.19
	40	6	7		0.06	0.86	0.11
	50	4	4		0.04	1	0.08
	60	2	2		0.02	1	0.04

Table 13 (Contd.). Precision, Recall & F –measure at different similarity cut-offs. (Whole image color codes). Class – Dinosaur.




Query Image	Similarity cut-off	Retrieved relevant images - rr	total retrieved images - total	Total relevant images in the database - Total	Recall $r = rr / \text{Total}$	Precision $p = rr / \text{total}$	F measure = $2 / (1/p + 1/r)$
	25	26	26	100	0.26	1	0.41
	30	22	22		0.22	1	0.36
	40	14	14		0.14	1	0.25
	50	9	9		0.09	1	0.17
	60	5	5		0.05	1	0.10
	25	47	75	100	0.47	0.63	0.54
	30	44	61		0.44	0.72	0.55
	40	30	35		0.3	0.86	0.44
	50	14	16		0.14	0.88	0.24
	60	6	7		0.06	0.86	0.11
	25	49	59	100	0.49	0.83	0.62
	30	44	48		0.44	0.92	0.59
	40	28	28		0.28	1	0.44
	50	10	10		0.1	1	0.18
	60	3	3		0.03	1	0.06

Table 14 lists the average Recall and average Precision for the query image class dinosaur for different similarity cut-offs. The P – R curves for sample queries of the table and corresponding average Precision & average Recall are shown in Figure 62. Average Precision, average Recall and average F-measures for different similarity cut-offs for the query images of the Table 14 have been plotted in Figure 63.

Table 14. Average Recall, Average Precision & Average F - measure. (Whole image color codes). Class – Dinosaur.

Similarity cut-off	Average Recall	Average Precision	Average F measure = $2 / (1/\text{Avg.p} + 1/\text{Avg.r})$
25	0.33	0.75	0.46
30	0.34	0.82	0.48
40	0.22	0.87	0.36
50	0.11	0.92	0.20
60	0.08	0.95	0.15
70	0.05	1	0.09

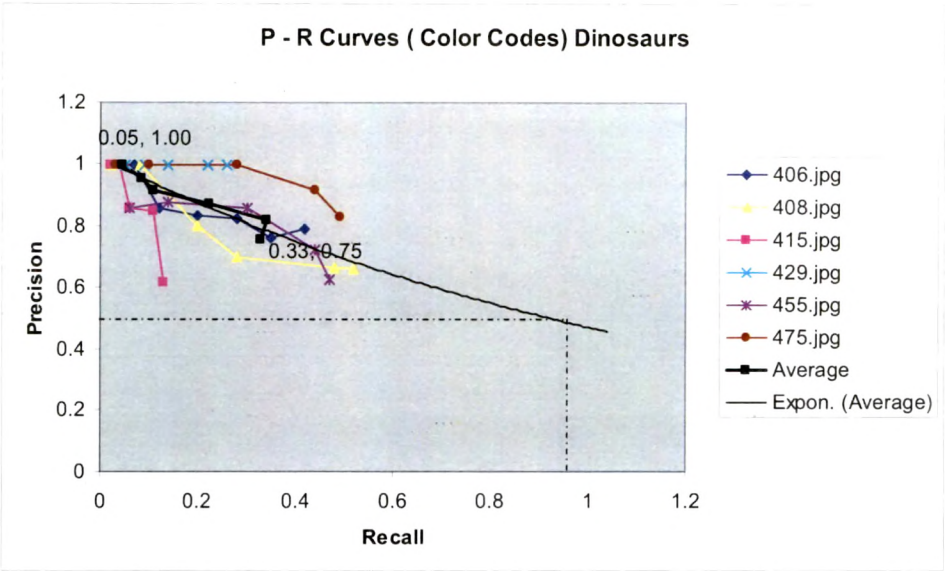


Figure 62. . P- R curves (whole image color codes). Class- Dinosaur.

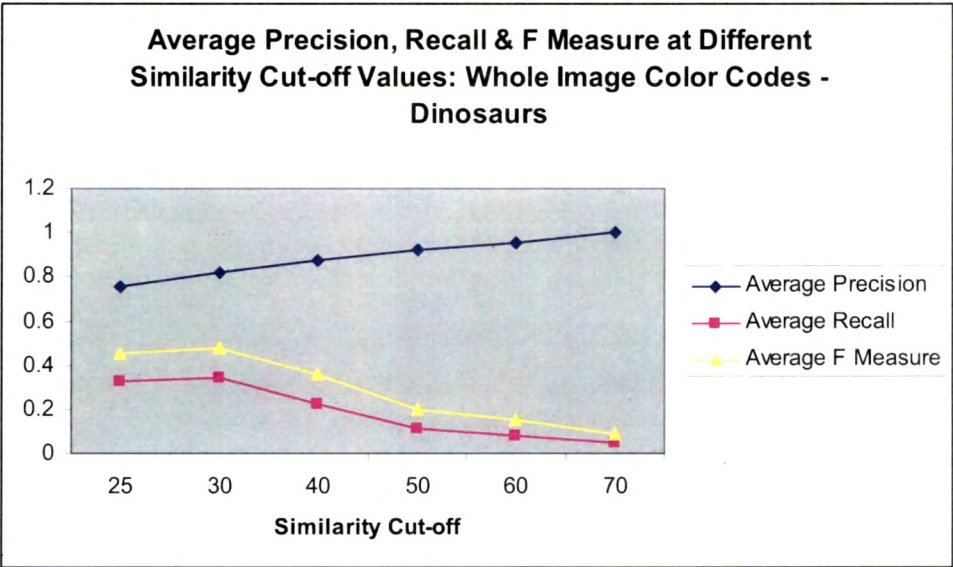


Figure 63. Average Precision, Average Recall, Average F – measures verses Similarity cut-offs. (Whole image color codes). Class – Dinosaur.

Following points are observed:

- The nature of obtained P – R curves is close to ideal P – R curve for some of the query images and similar to practical P – R curves for majority of query images.
- For the variations in colors & poses of foreground objects high recall with good precision is achieved for many sample queries.
- Stricter similarity cut-off increases the Precision at the cost of Recall.

- For many queries, Precision of 1.0 is achieved.
- Range of average performance measures for the class
 - 100 % of average Precision for 5 % of average Recall
 - 75 % of average Precision for 33 % of average Recall
 - Giving 25 % of fall in Precision to raise Recall by 28 %.
- The exponentially extended trend line intersects average Precision = (0.5) line at average Recall at value 0.95 (approx.) implies good performance measures.

6.6.1.8 Query Image Response Example: Class - Dinosaur

The query response of a dinosaur image [Wang, 2001] [SIMPLcity, on line] is shown for similarity cut-off of 40 in Figure 64.

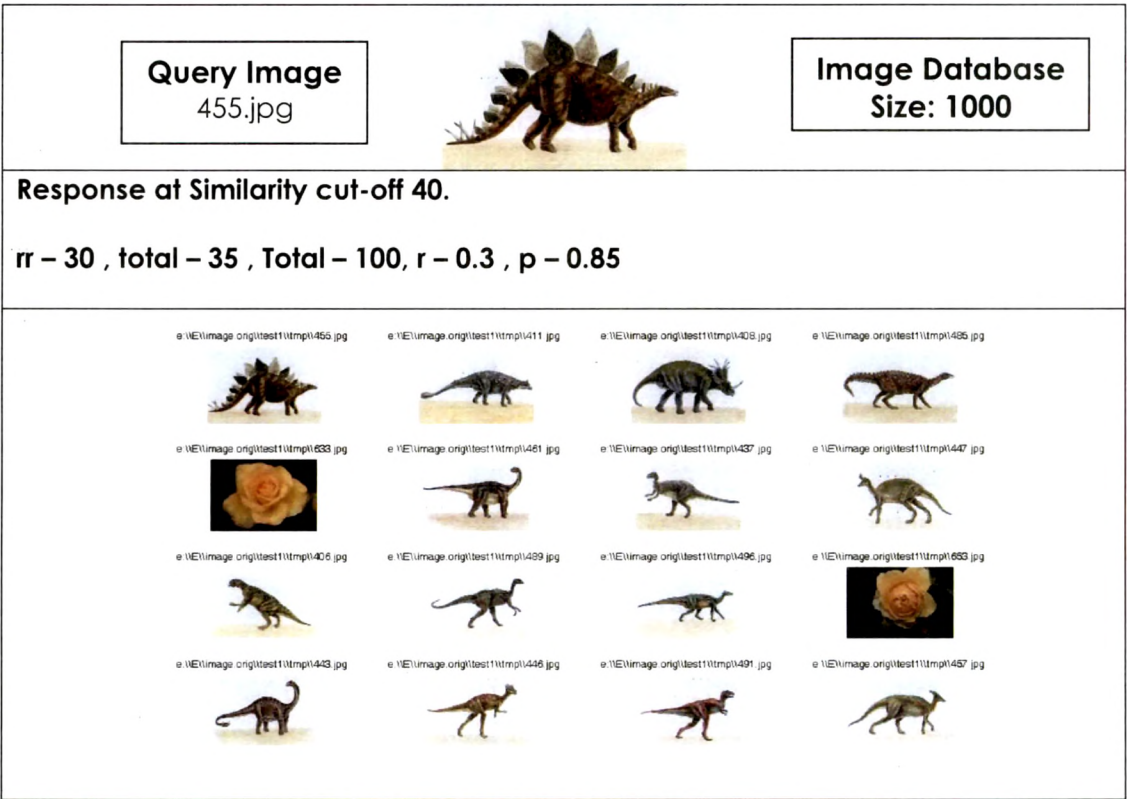


Figure 64. Query response of a dinosaur image at similarity cut-off 40.

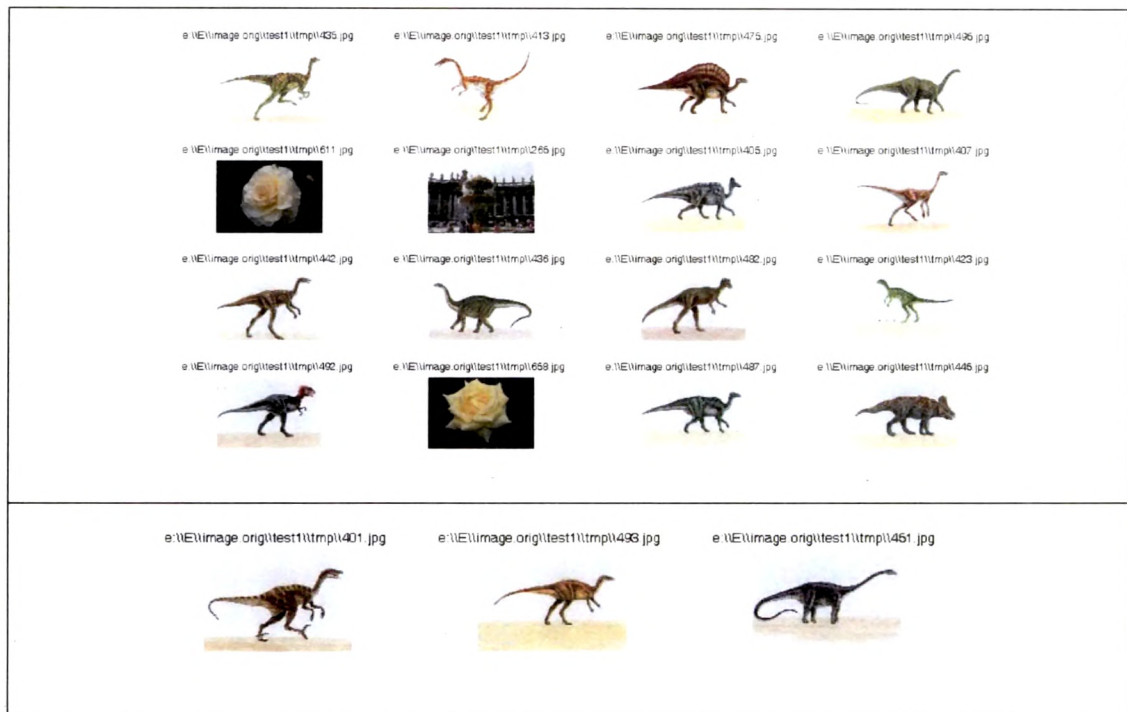


Figure 64 (Contd.). Query response of a dinosaur image at similarity cut-off 40.

6.6.1.9 Query Image Response Examples: Other Classes

The query responses shown in Figure 65 & Figure 66 can well illustrate the issue of subjectivity involved in the intention of user and image content description. Is user intending to retrieve blue skied images or images of blue sky with white clouds? Or, is he aiming to get seashore images or images containing water? The answer will determine the number of relevant images retrieved and hence the Precision, Recall & finally the performance of the system. The Figure 67 to Figure 70 show the query responses of images of other classes of SIMPLICity [Wang, 2001] [SIMPLICity, on line].

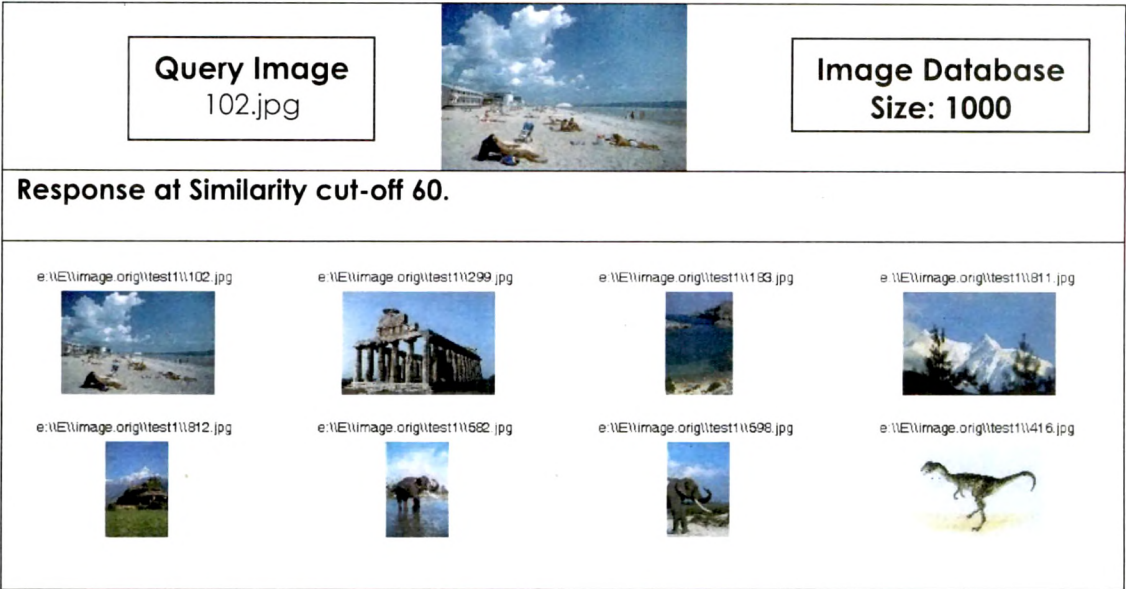


Figure 65. Query response of a sea-shore image at similarity cut-off 60.

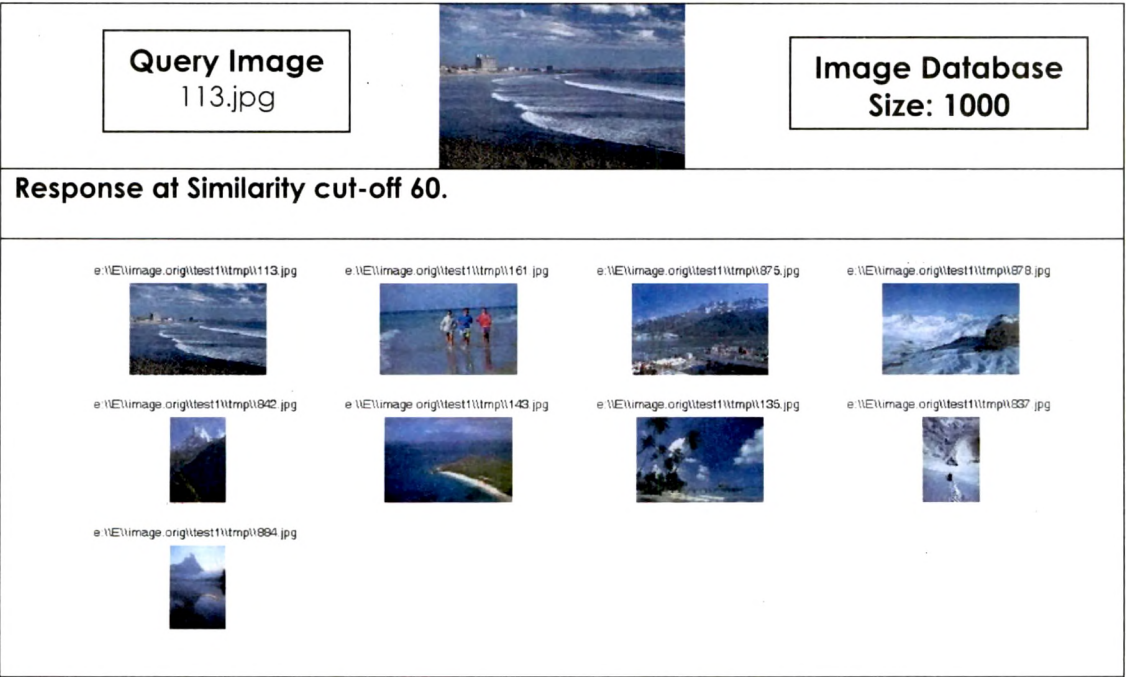


Figure 66. Query response of another sea-shore image at similarity cut-off 60.

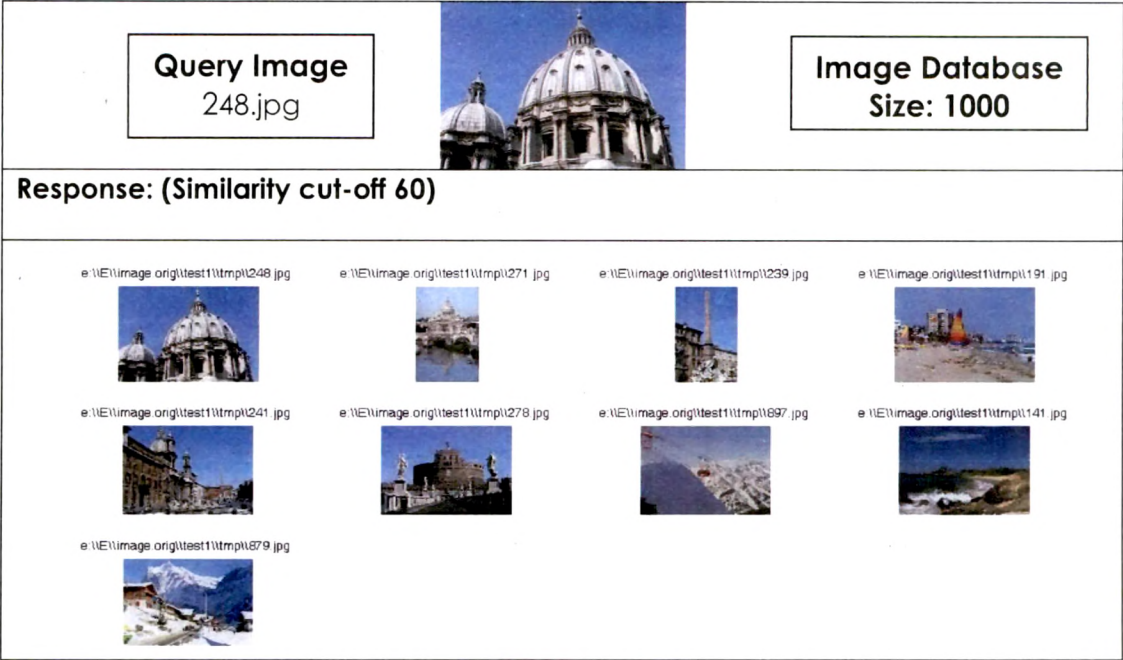


Figure 67. Query response of a sculpture image at similarity cut-off 60.

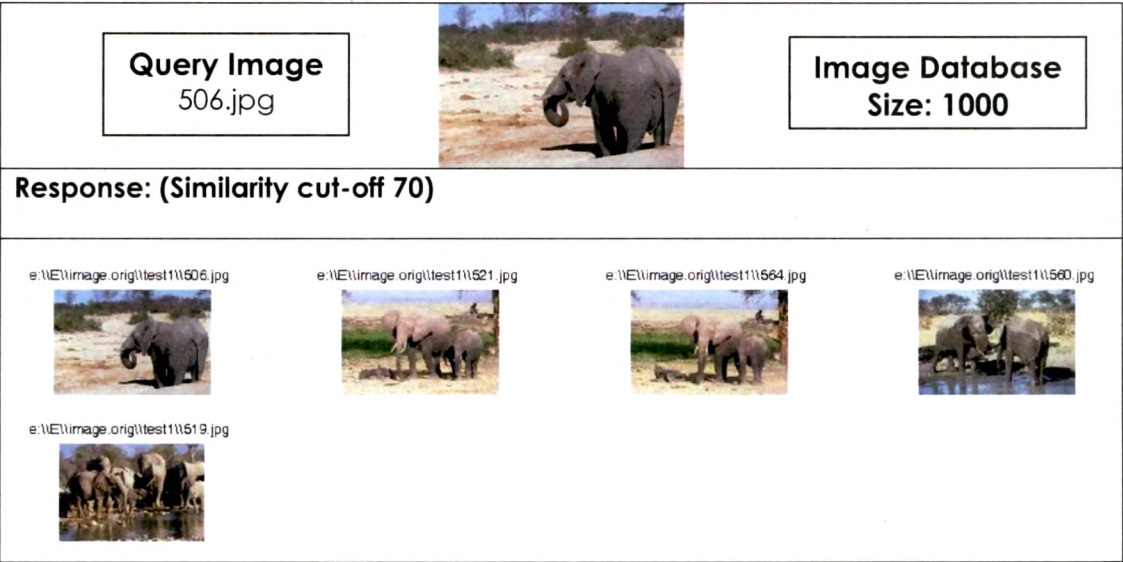


Figure 68. Query response of an elephant image at similarity cut-off 70.

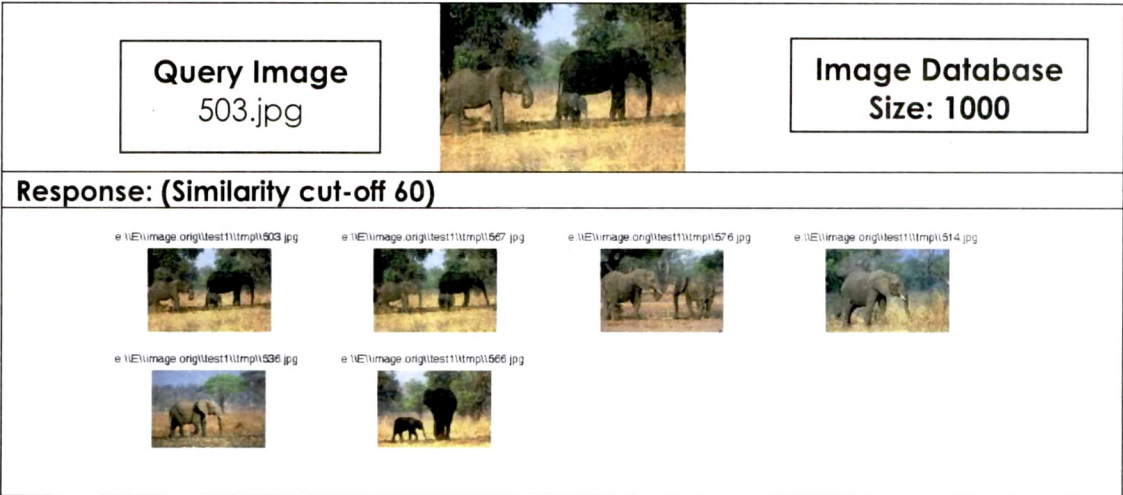


Figure 69. Query response of another elephant image at similarity cut-off 60.



Figure 70. Query response of served food image at similarity cut-off 60.

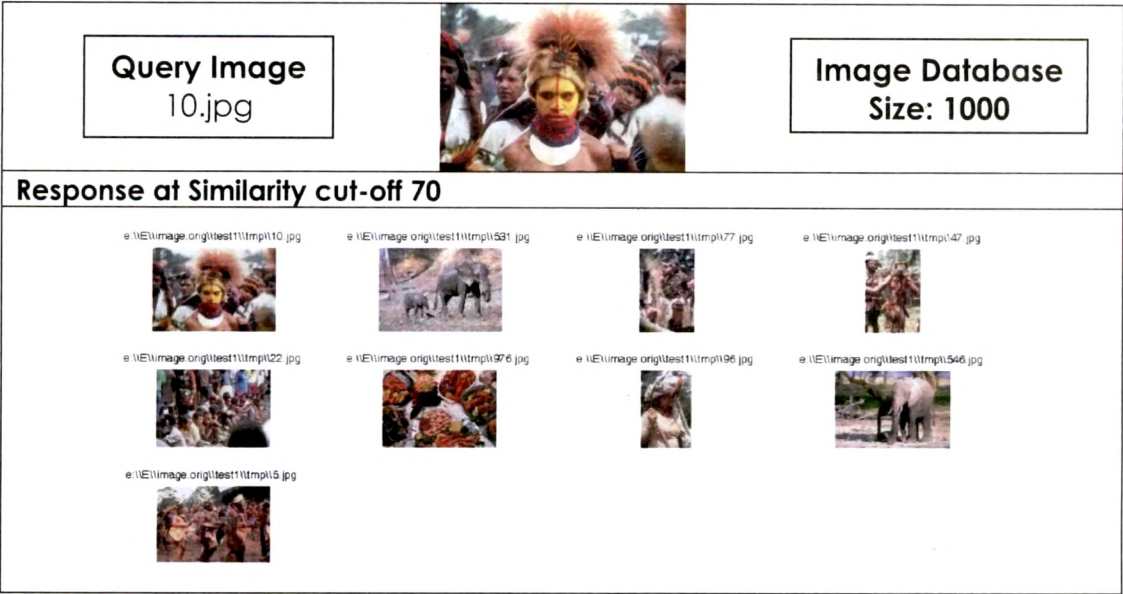


Figure 71. Query response of an image of a tribal man with color painted on face at similarity cut-off 60.

The average Recall, average Precision and average F-measures for all test queries have been tabulated in Table 15 and class wise P – R curves and method average P – R curves are shown in Figure 72. The average Recall, average Precision and average F-measures for all test queries for the method have been plotted in Figure 73.

Table 15. Average Recall, Average Precision & Average F - measure. (Whole image color codes). All queries for the method.

Similarity cut-off	Average Recall	Average Precision	Average F measure = 2 / (1/Avg.p + 1/Avg.r)
25	0.44	0.69	0.54
30	0.39	0.74	0.51
40	0.27	0.78	0.40
50	0.17	0.83	0.29
60	0.11	0.86	0.20
70	0.06	0.96	0.12

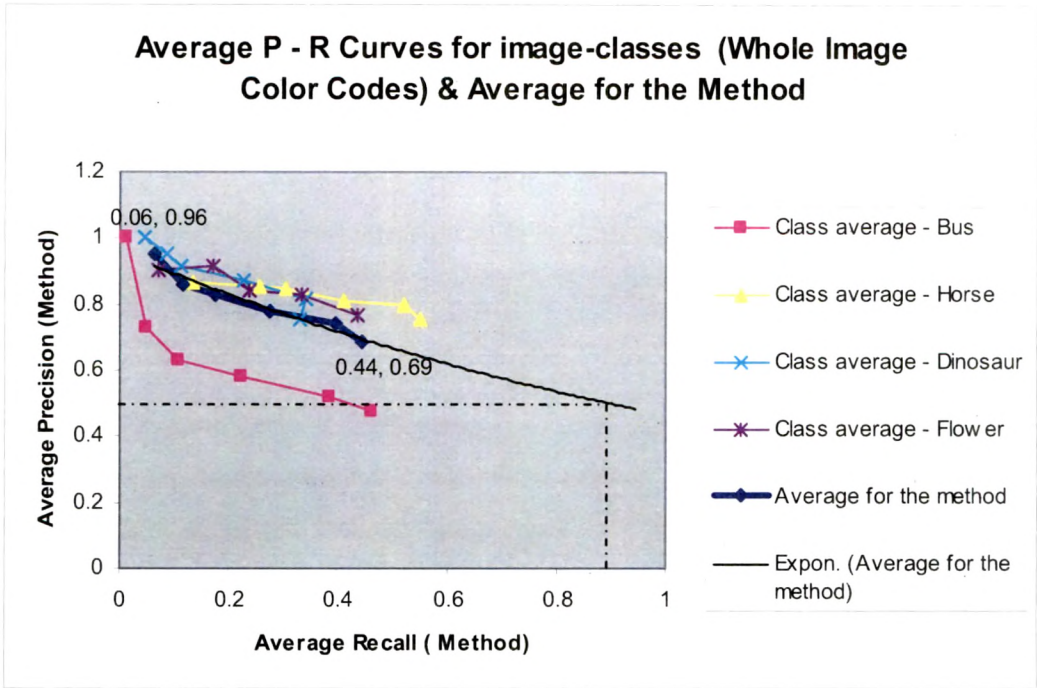


Figure 72. P- R curves (whole image color codes). All queries for the method.

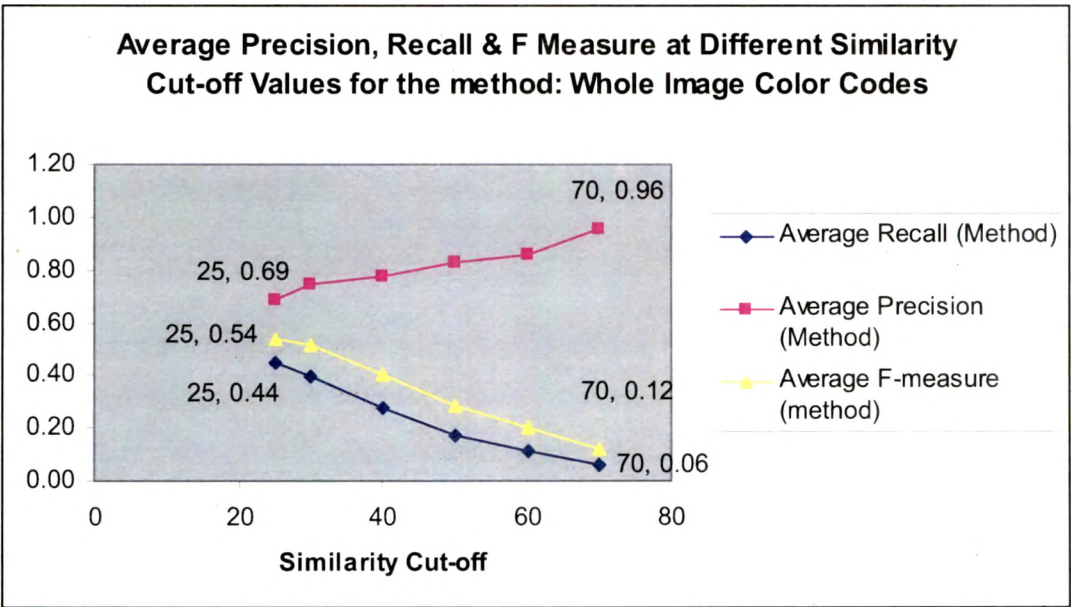


Figure 73. Average Precision, Average Recall, Average F – measures verses Similarity cut-offs. (Whole image color codes). All queries for the method.

6.6.2 Discussion

- The performance has been evaluated on 1000 images of standard data base [Wang, 2001] [SIMPLicity, on line] consisting of 10 classes of images for total 161 queries with different similarity cut-offs for 33 query images of 4 different classes.
- The method is robust to illumination, pose and view point variations as it is based on broader color descriptors.
- The feature extraction and retrieval methods require lesser computations compared to boundary detection based proposed methods.
- The broader descriptors are characterized to yield higher Recall. So are the color codes.
- The method works well even on images of poor quality and low resolutions.
- The ranking of nearly similar images is high.
- The method is not sensitive to image scale and rotation.
- The Precision measures obtained for majority of the queries of all four classes are significantly high with good Recall.
- Range of average performance measures for all queries of all classes
 - 96 % of average Precision for 6 % of average Recall
 - 69 % of average Precision for 44 % of average Recall
 - Giving 27 % of fall in average Precision to raise average Recall by 38%.
- The exponentially extended trend line intersects average Precision = (0.5) line at average Recall at value 0.90 (approx.) implies good performance measures.
- The results with proposed method are better than many reported in literature.

6.7 Foreground Color Codes Based CBIR

Foreground based image retrieval enables user to search images on the basis of objects contained in the image. The exclusion of background narrows down subjectivity induced diversity about the image content. Precisely detected foreground encompassing prominent boundaries yielding foreground region attributes and color codes of the foreground are used as combined features for image retrieval. The normalized histogram constructed for foreground region is compared with that of image of image database. The algorithm is applied on color codes of foreground for

measuring color distribution similarity of foreground regions of images under considerations. Following method specific steps replace corresponding generic Steps 3 & 4 of [Algorithm 4, Section 6.5](#) :

Step 3: Read (or extract) foreground color code features of given query image.

Step 4: For every image-feature-file of target folder,

Read corresponding foreground color code features of the image-feature-file of target folder.

Calculate (dis)similarity_index $i = \sum \text{abs}(h_{qi} - h_{ij})$, for $1 < j \leq \text{number of Bins}$, Where,

h_{qi} indicates j^{th} bin of normalized histogram of color codes for the query image

h_{ij} indicates j^{th} bin of normalized histogram of color codes for i^{th} image of database

Store path of data base image, needed for display.

Algorithm 6. Foreground color codes based image retrieval.

6.7.1 Performance Evaluation

The performance of the method has been tested on image database of SIMPLcity [Wang, 2001] [SIMPLcity, on line] consisting of 371 images. Exhaustive performance evaluation has been carried out for two classes of database – Bus and flower. Recall, Precision and F –measure are computed for sample queries of each class of images for different similarity cut-offs. Average Recall, Average Precision and Average F – measures for the class are tabulated to analyze performance of the method for given class of images. P – R curves for query responses along with Average Precision and Average Recall are plotted for performance analysis. Average Recall, Average Precision and Average F – measures are also plotted for different similarity cut-off.

6.7.1.1 Query Image Class: Bus

- The performance evaluation on image database [Wang, 2001] [SIMPLcity, on line] consisting of 371 images has been shown in Table 16 for
 - 11 varieties of query images [Wang, 2001] [SIMPLcity, on line] for different similarity cut-offs &
 - 57 queries

- o The set of selected query images is same as the set used for image retrieval using whole image color codes.

Table 16. Precision, Recall & F –measure at different similarity cut-offs. (Foreground color codes). Class - Bus.

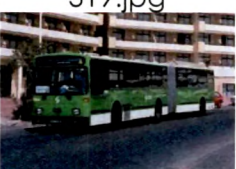
Query Image	Similarity cut-off	Retrieved relevant images - rr	total retrieved images - total	Total relevant images in the database - Total	Recall $r = rr / \text{Total}$	Precision $p = rr / \text{total}$	F measure = $2 / (1/p + 1/r)$
	25	64	74	100	0.64	0.86	0.73
	30	56	60		0.56	0.93	0.70
	40	36	38		0.36	0.95	0.52
	50	24	24		0.24	1.00	0.39
	60	10	10		0.1	1.00	0.18
	70	4	4		0.04	1.00	0.08
	25	59	61	100	0.59	0.97	0.73
	30	52	53		0.52	0.98	0.68
	40	39	39		0.39	1.00	0.56
	50	25	25		0.25	1.00	0.40
	60	13	13		0.13	1.00	0.23
	70	2	2		0.02	1.00	0.04
	25	21	22	100	0.21	0.95	0.34
	30	19	20		0.19	0.95	0.32
	40	7	8		0.07	0.88	0.13
	50	3	3		0.03	1.00	0.06
	60	2	2		0.02	1.00	0.04
	25	41	47	100	0.41	0.87	0.56
	30	32	37		0.32	0.86	0.47
	40	25	25		0.25	1.00	0.40
	50	12	12		0.12	1.00	0.21
	60	6	6		0.06	1.00	0.11
	25	8	47	100	0.08	0.17	0.11
	30	6	38		0.06	0.16	0.09
	40	3	22		0.03	0.14	0.05
	50	3	18		0.03	0.17	0.05
	60	1	5		0.01	0.20	0.02

Table 16 (Contd.). Precision, Recall & F –measure at different similarity cut-offs. (Foreground color codes). Class - Bus.

Query Image	Similarity cut-off	Retrieved relevant images - rr	total retrieved images - total	Total relevant images in the database - Total	Recall $r = rr / \text{Total}$	Precision $p = rr / \text{total}$	F measure= $2 / (1/p + 1/r)$
<div>326.jpg</div> 	25	26	59	100	0.26	0.44	0.33
	30	18	64		0.18	0.28	0.22
	40	8	24		0.08	0.33	0.13
	50	3	5		0.03	0.60	0.06
	60	2	2		0.02	1.00	0.04
<div>365.jpg</div> 	25	62	73	100	0.62	0.85	0.72
	30	56	61		0.56	0.92	0.70
	40	33	34		0.33	0.97	0.49
	50	11	11		0.11	1.00	0.20
	60	2	2		0.02	1.00	0.04
<div>388.jpg</div> 	25	74	78	100	0.74	0.95	0.83
	30	65	67		0.65	0.97	0.78
	40	39	40		0.39	0.98	0.56
	50	21	22		0.21	0.95	0.34
	60	7	7		0.07	1.00	0.13
<div>366.jpg</div> 	25	15	57	100	0.15	0.26	0.19
	30	10	39		0.1	0.26	0.14
	40	8	28		0.08	0.29	0.13
	50	5	12		0.05	0.42	0.09
	60	2	2		0.02	1.00	0.04
<div>344.jpg</div> 	25	30	32	100	0.3	0.94	0.45
	30	21	23		0.21	0.91	0.34
	40	10	11		0.1	0.91	0.18
	50	3	3		0.03	1.00	0.06
	60	2	2		0.02	1.00	0.04
<div>369.jpg</div> 	25	77	82	100	0.77	0.94	0.85
	30	69	71		0.69	0.97	0.81
	40	44	44		0.44	1.00	0.61
	50	16	16		0.16	1.00	0.28
	60	4	4		0.04	1.00	0.08

Table 17. Average Recall, Average Precision & Average F -measure. (Foreground color codes). Class – Bus.

Similarity cut-off	Average Recall	Average Precision	Average F measure = 2 / (1/Avg.p + 1/Avg.r)
25	0.40	0.75	0.52
30	0.34	0.75	0.46
40	0.21	0.77	0.33
50	0.11	0.83	0.19
60	0.04	0.93	0.08
70	0.03	1.00	0.06

The average Precision and average Recall for class Bus have been tabulated in Table 17 and plotted in Figure 74 along with P – R curves for individual bus image query responses. The average Recall, average Precision and average F-measures with respect to similarity cut-off for all test queries of the class have been presented in Figure 75.

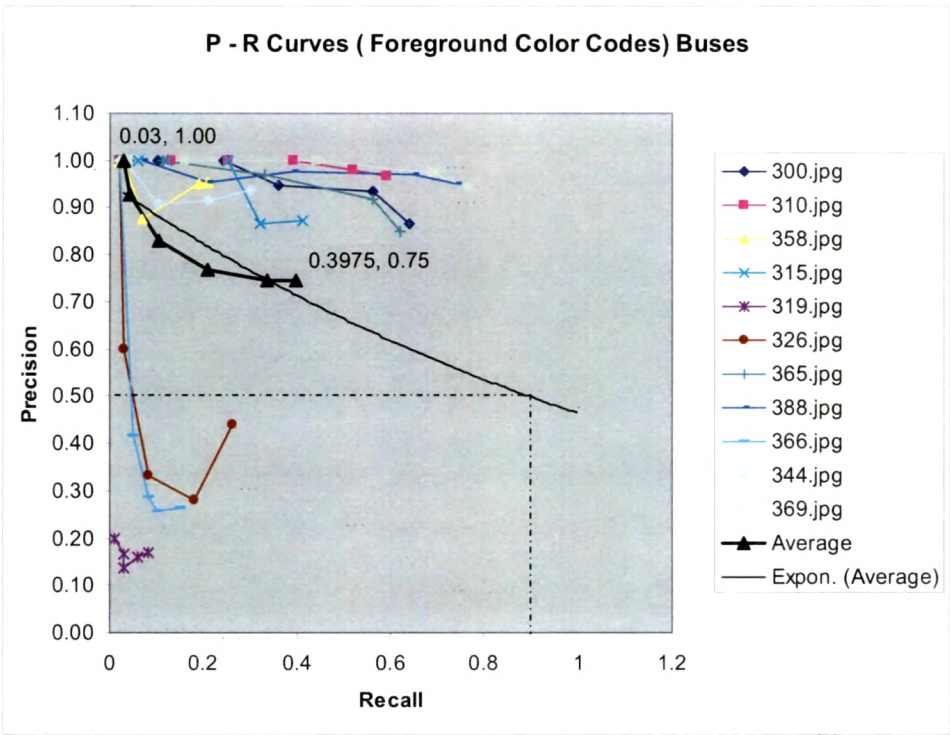


Figure 74. P- R curves (Foreground color codes). Class - Bus.

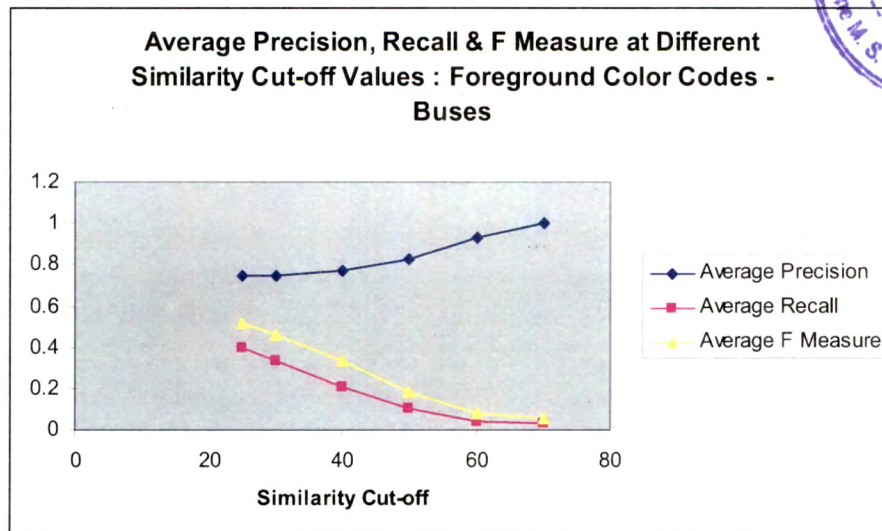


Figure 75. P- R curves (Foreground color codes). Class - Bus.

Following points are observed:

- The nature of obtained P – R curves matches with the practical P – R curves.
- Stricter similarity cut-off increases the Precision at the cost of Recall.
- Despite vast variations in bus colors, poses and illumination conditions, high recall with good precision is achievable for many sample queries.
- For all but one queries, Precision of 1.0 is achieved.
- The Precision and recall measures have been improved for all but one (319.jpg) query images.
- Improvement in precision is contributed by two factors – only foreground region based comparison and reduced image database size.
- Range of average performance measures for the class
 - 100 % of average Precision for 3 % of average Recall
 - 75 % of average Precision for 40 % of average Recall
 - Giving 25 % of fall in average Precision to raise average Recall by 37 %
- The exponentially extended trend line intersects average Precision = (0.5) line at average Recall at value 0.9 (approx.) implies quite good performance measures for images of the class.

6.7.1.2 Query Response Example: Class - Bus

The query response of a bus image [Wang, 2001] [SIMPLIcity, on line] at similarity cut-off of 25 is shown in Figure 76. The Recall of 59 % with 97 % of Precision is to be noted.

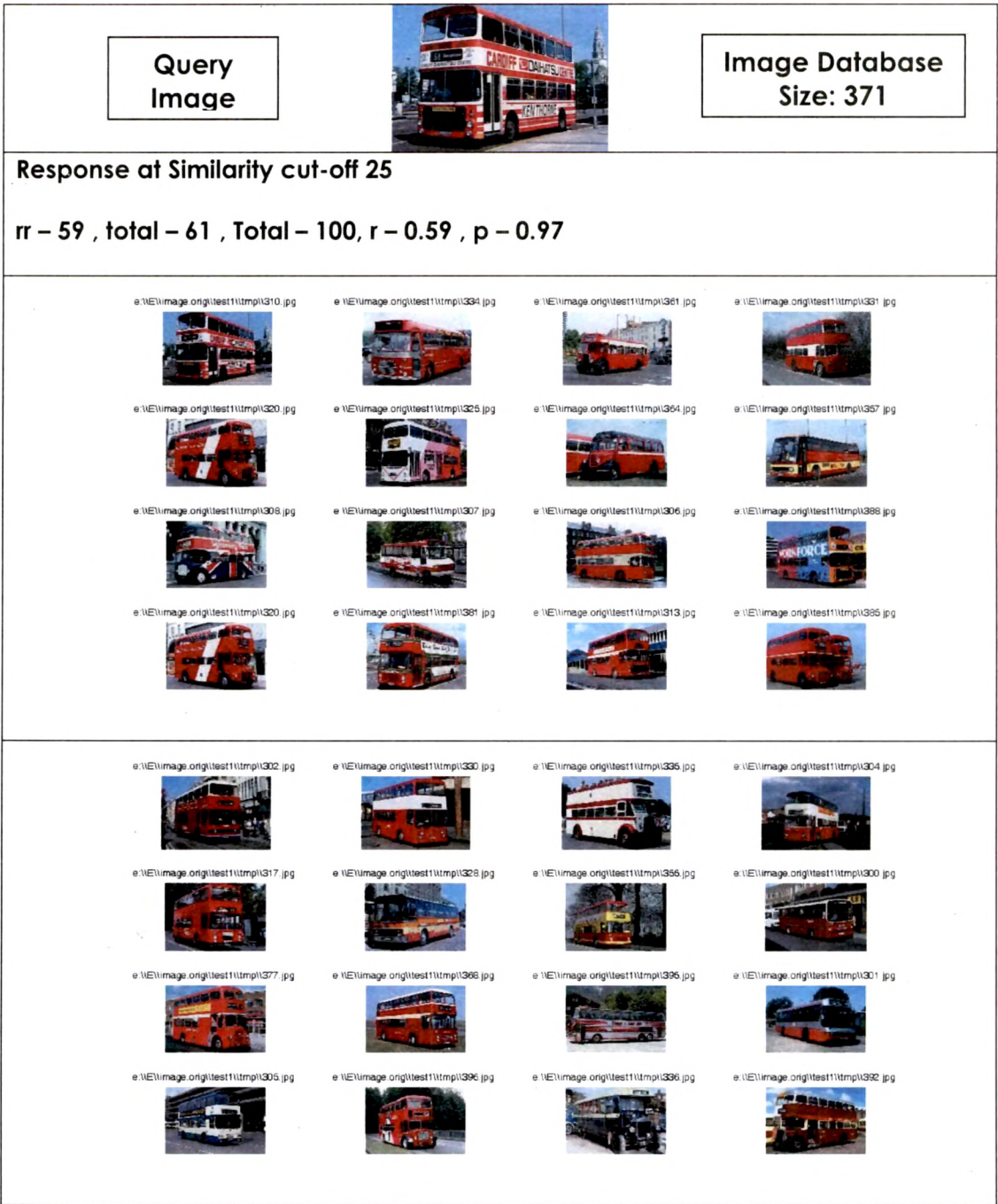


Figure 76. Query response of a bus image at similarity cut-off 25. (FGCC)

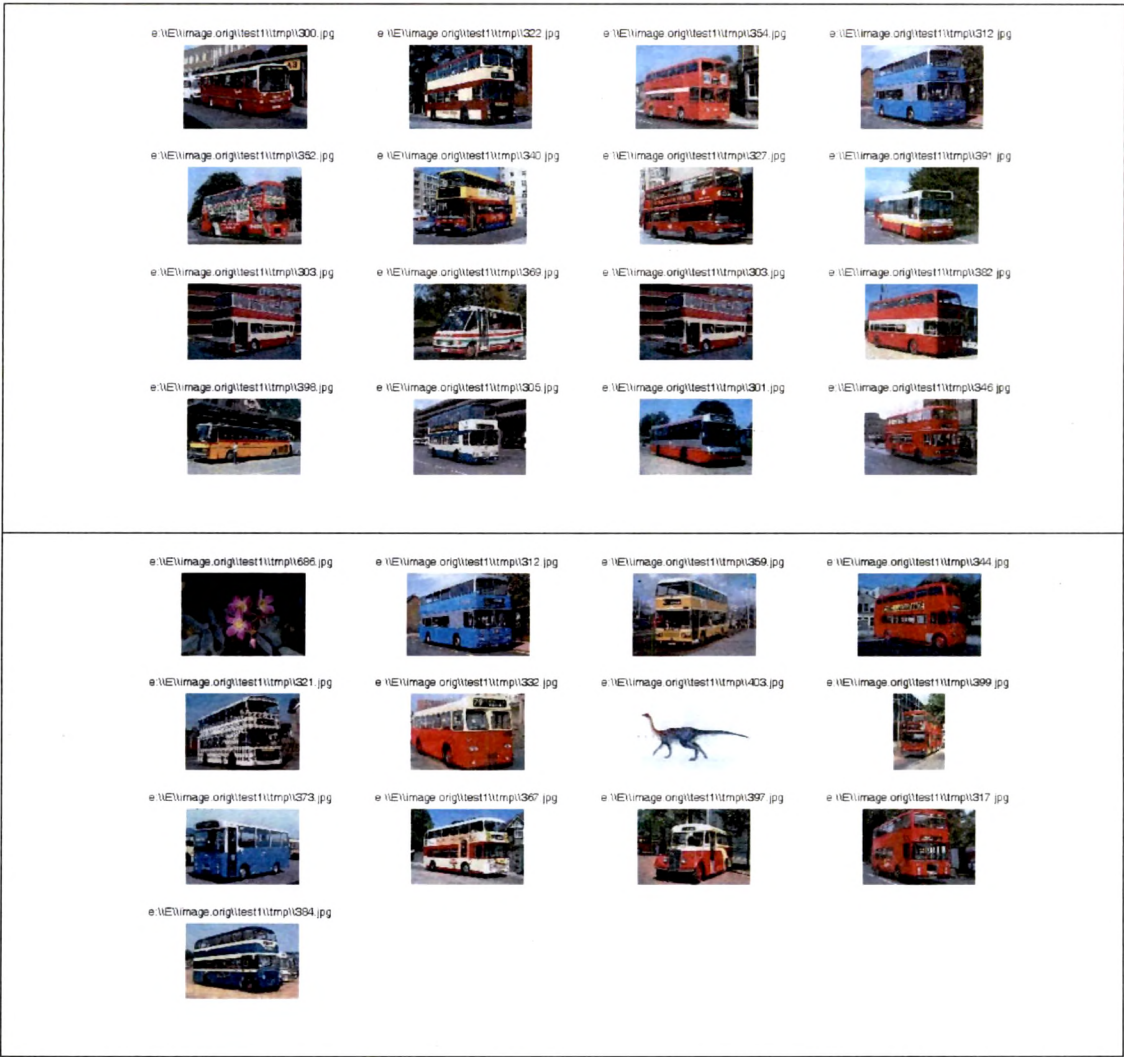


Figure 76 (Contd.). Query response of a bus image at similarity cut-off 25. (FGCC)

6.7.1.3 Query Image Class – Flower

- The performance evaluation on image database [Wang, 2001] [SIMPLcity, on line] consisting of 371 images has been shown in Table 18 for
 - 10 varieties of query images [Wang, 2001] [SIMPLcity, on line] for different similarity cut-offs &
 - 51 queries
- The set of selected query images consists of 8 images used in the set for image retrieval using whole image color codes. Two images used in the first set have been replaced because of their inferior foreground extraction.

Table 18. Precision, Recall & F –measure at different similarity cut-offs. (Foreground color codes). Class - Flower.



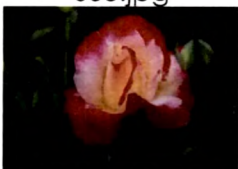



Query Image	Similarity cut-off	Retrieved relevant images - rr	total retrieved images - total	Total relevant images in the database - Total	Recall $r = rr / \text{Total}$	Precision $p = rr / \text{total}$	F measure= $2/(1/p + 1/r)$
	25	33	34	57	0.58	0.97	0.73
	30	31	31		0.54	1.00	0.70
	40	27	27		0.47	1.00	0.64
	50	20	20		0.35	1.00	0.52
	60	11	11		0.19	1.00	0.32
	70	9	9		0.16	1.00	0.27
	25	15	45	24	0.63	0.33	0.43
	30	15	37		0.63	0.41	0.50
	40	13	30		0.54	0.43	0.48
	50	11	26		0.46	0.42	0.44
	60	10	18		0.42	0.56	0.48
	25	17	50	57	0.30	0.34	0.32
	30	15	35		0.26	0.43	0.33
	40	10	17		0.18	0.59	0.27
	50	5	8		0.09	0.63	0.15
	60	4	4		0.07	1.00	0.13
	25	1	19	9	0.11	0.05	0.07
	30	1	11		0.11	0.09	0.10
	40	1	5		0.11	0.20	0.14
	50	1	2		0.11	0.50	0.18
	60	1	1		0.11	1.00	0.20
	25	12	13	57	0.21	0.92	0.34
	30	5	5		0.09	1.00	0.16
	40	4	4		0.07	1.00	0.13
	50	2	2		0.03	1.00	0.07
	60	2	2		0.03	1.00	0.07
	25	17	64	24	0.71	0.27	0.39
	30	17	56		0.71	0.30	0.42
	40	16	42		0.67	0.38	0.48
	50	16	34		0.67	0.47	0.55
	60	13	28		0.54	0.46	0.50

Table 18 (Contd.). Precision, Recall & F –measure at different similarity cut-offs. (Foreground color codes). Class - Flower.





Query Image	Similarity cut-off	Retrieved relevant images - rr	total retrieved images - total	Total relevant images in the database - Total	Recall $r = rr / \text{Total}$	Precision $p = rr / \text{total}$	F measure= $2/(1/p + 1/r)$
<div>682.jpg</div> 	25	32	34	57	0.56	0.94	0.70
	30	30	31		0.53	0.97	0.68
	40	27	28		0.47	0.96	0.63
	50	20	20		0.35	1.00	0.52
	60	13	13		0.23	1.00	0.37
<div>618.jpg</div> 	25	26	27	57	0.46	0.96	0.62
	30	25	26		0.44	0.96	0.60
	40	20	21		0.35	0.95	0.51
	50	18	18		0.32	1.00	0.48
	60	15	15		0.26	1.00	0.42
<div>621.jpg</div> 	25	27	28	57	0.47	0.96	0.63
	30	24	25		0.42	0.96	0.59
	40	20	21		0.35	0.95	0.51
	50	18	18		0.32	1.00	0.48
	60	15	15		0.26	1.00	0.42
<div>640.jpg</div> 	25	5	5	10	0.5	1.00	0.67
	30	4	4		0.4	1.00	0.57
	40	3	3		0.3	1.00	0.46
	50	2	2		0.2	1.00	0.33
	60	1	1		0.1	1.00	0.18

Table 19. Average Recall, Average Precision & Average F - measure. (Foreground color codes). Class – Flower.

Similarity cut-off	Average Recall	Average Precision	Average F measure = $2 / (1/\text{Avg.p} + 1/\text{Avg.r})$
25	0.45	0.68	0.54
30	0.41	0.71	0.52
40	0.35	0.75	0.48
50	0.29	0.80	0.43
60	0.22	0.90	0.35
70	0.16	1.00	0.28

The average Precision and average Recall for class Flower have been tabulated in Table 19 and plotted in Figure 77 along with P – R curves for individual query responses. The average Recall, average Precision and average F-measures for all test queries of the class have been presented in Figure 78.

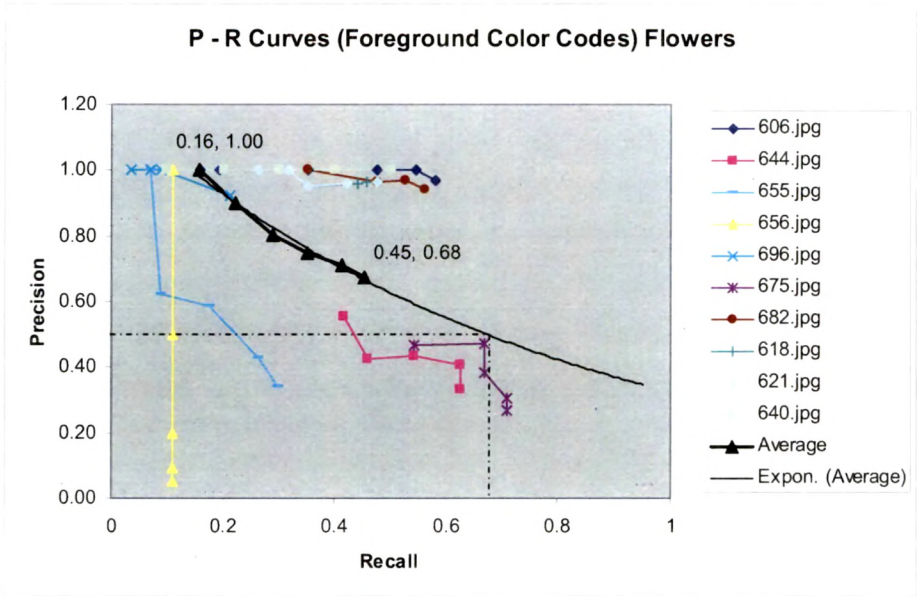


Figure 77. P- R curves (Foreground color codes). Class- Flower.

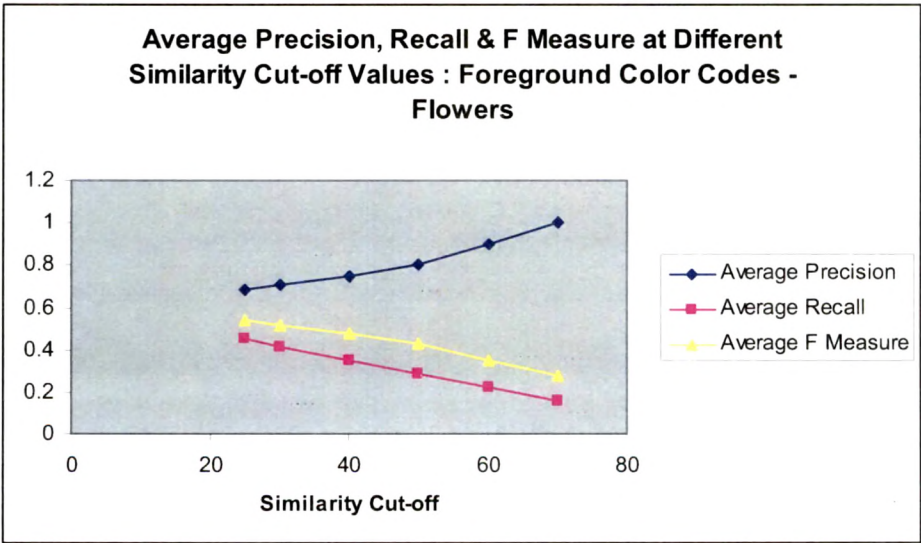


Figure 78. Avg. Precision, Avg. Recall, Avg. F – measures verses Similarity cut-offs. (Foreground color codes). Class – Flower.

Following points are observed:

- The nature of obtained P – R curves matches with the practical P – R curves.
- Stricter similarity cut-off increases the Precision at the cost of Recall.
- Despite vast variations in foreground colors and illumination conditions, high recall with good precision is achievable for many sample queries.
- For many queries, Precision of 1.0 is achieved for reasonable Recall.
- Range of average performance measures for the class
 - 100 % of average Precision for 16 % of average Recall
 - 68 % of average Precision for 45% of average Recall
 - Giving 32 % of fall in average Precision to raise average Recall by 29 %
- The exponentially extended trend line intersects average Precision = (0.5) line at average Recall at value 0.67 (approx.) implies good performance measures for images of the class.

6.7.1.4 Query Response Example: Class - Flower

The query responses of two different flower images [Wang, 2001] [SIMPLcity, on line] at lowest similarity cut-off of 25 are shown in Figure 79 & Figure 80 respectively. The selected query images are of typical foreground colors.



Figure 79. Query response of a flower image at similarity cut-off 25. (FGCC)

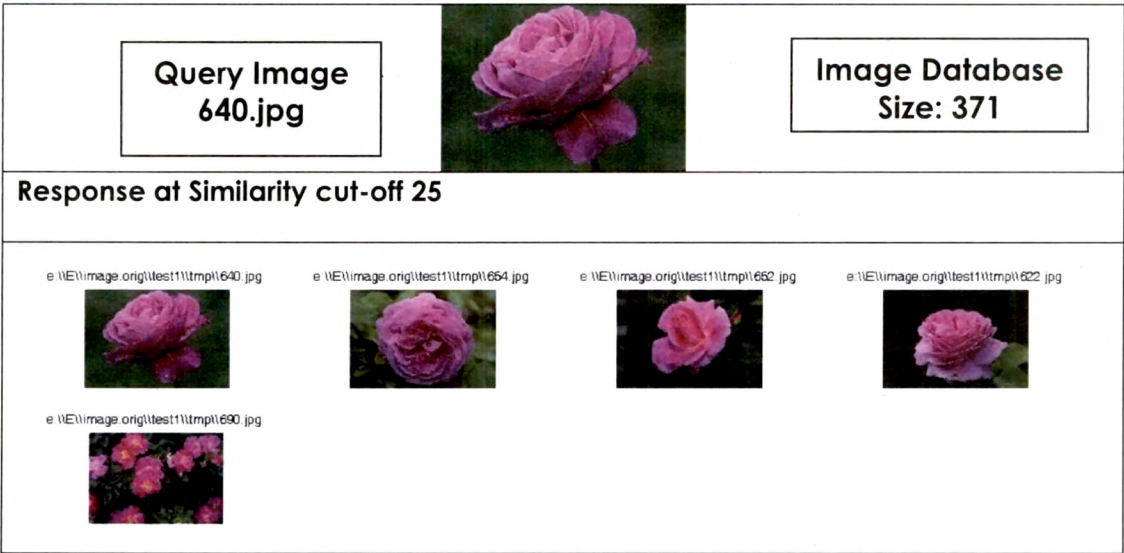


Figure 80. Query response of another flower image at similarity cut-off 25. (FGCC)

6.7.2 Discussion

- The performance has been evaluated on 371 images of standard data base [Wang, 2001] [SIMPLcity, on line] consisting of all images of two classes and some images from other classes for total 115 queries with different similarity cut-offs for 21 query images of 2 different classes.
- The method is not suitable for images containing very small foreground objects and objects touching to image boundaries which may not be encompassed by prominent boundaries.
- The extracted background excludes background and related features from comparison enabling user to perform search based on objects contained in the image.
- The method is robust to illumination and less sensitive to pose and view point variations as it is based on broader color descriptors of the extracted foreground.
- The feature extraction and retrieval methods require significant computations.
- The performance of the method is not sensitive to regions attached to foreground objects, because, for given image, such regions can be made to constitute small percentage of total extracted foreground by performing feature extraction at higher wavelet level. The low resolution and poor

quality of images affect foreground extraction and hence performance of the method.

- The Precision measures obtained for majority of the queries are significantly high with good Recall.
- The exponentially extended trend line intersects average Precision = (0.5) line of both classes at average Recall at values 0.9 & 0.67 (approx.) respectively, imply good performance measures.

6.8 Foreground Shape Correlation Based CBIR

The normalized unsegmented foreground region has been utilized as the feature for the image comparison. The method specific steps, replacing Step 3 & Step 4 of algorithm specified in [Algorithm 4, Section 6.5](#) are:

Step 3: Read (or extract) normalized unsegmented foreground region features for the query image. Call it R_q .

Step 4: For every image-feature-file of target folder,

Read normalized unsegmented foreground region features of the image-feature-file of target folder. Let us call it R_d .

Calculate correlation coefficients of R_q & R_d .

Find the significant correlation coefficient.

Calculate (dis)similarity_index_i = $100 - \text{abs}(\text{significant correlation coefficient of } R_q \text{ \& } R_d) * 100$

Algorithm 7. Foreground shape correlation based image retrieval.

Unsegmented foreground regions have been obtained by excluding background regions found in step 8 of Algorithm 3.

The query response for the method is shown in Figure 82. Note that the performance & results with 0 % weight of foreground color code attributes in the composite similarity constraint of next method – Foreground Color Codes & Shape Correlation corresponds to this method of image retrieval.

6.8.2 Performance Evaluation

The performance of the method has been tested on image database of SIMPLicity [Wang, 2001] [SIMPLicity, on line] consisting of 371 images. The Precision, Recall and F – measures are shown in Table 19 for 10 images of class Flower and total of 20 queries. Corresponding P – R curves are plotted in Figure 81.

Table 20. Precision, Recall & F –measure at different similarity cut-offs. (Foreground shape correlation). Class - Flower.



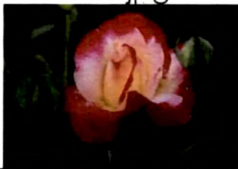



Query Image	Similarity cut-off	Retrieved relevant images - rr	total retrieved images - total	Total relevant images in the database - Total	Recall $r = rr / \text{Total}$	Precision $p = rr / \text{total}$	F measure= $2/(1/p + 1/r)$
606.jpg 	60	39	41	100	0.39	0.95	0.55
	70	4	4		0.04	1.00	0.08
644.jpg 	60	45	47	100	0.45	0.96	0.61
	70	23	23		0.23	1.00	0.37
655.jpg 	60	10	10	100	0.1	1.00	0.18
	70	1	1		0.01	1.00	0.02
656.jpg 	60	34	36	100	0.34	0.94	0.50
	70	7	7		0.07	1.00	0.13
696.jpg 	60	1	1	100	0.01	1.00	0.02
	70	1	1		0.01	1.00	0.02
675.jpg 	60	2	2	100	0.02	1.00	0.04
	70	1	1		0.01	1.00	0.02

Table 20 (Contd.). Precision, Recall & F –measure at different similarity cut-offs. (Foreground shape correlation).
Class - Flower.

Query Image	Similarity cut-off	Retrieved relevant images - rr	total retrieved images - total	Total relevant images in the database - Total	Recall $r = rr / \text{Total}$	Precision $p = rr / \text{total}$	F measure= $2/(1/p + 1/r)$
682.jpg 	60	4	5	100	0.04	0.80	0.08
	70	3	3		0.03	1.00	0.06
618.jpg 	60	42	45	100	0.42	0.93	0.58
	70	19	19		0.19	1.00	0.32
621.jpg 	60	47	50	100	0.47	0.94	0.63
	70	26	26		0.26	1.00	0.41
640.jpg 	60	6	7	100	0.06	0.86	0.11
	70	2	2		0.02	1.00	0.04

Table 21. Average Recall, Average Precision & Average F - measure. (Foreground Shape correlation). Class – Flower.

Similarity cut-off	Average Recall	Average Precision	Average F measure = $2 / (1/\text{Avg.p} + 1/\text{Avg.r})$
60	0.23	0.94	0.37
70	0.09	1.00	0.16

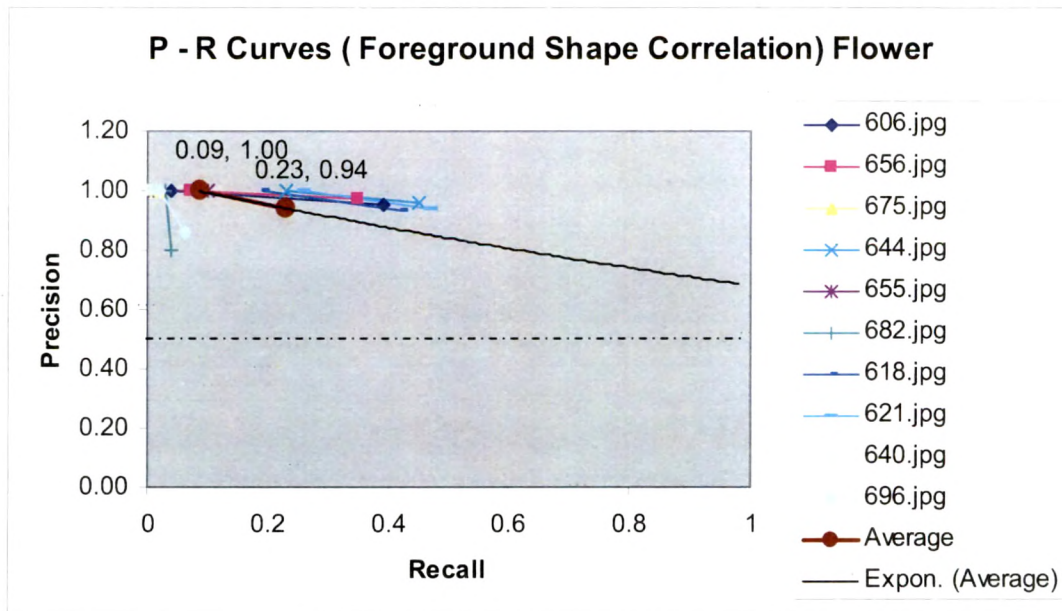


Figure 81. P- R curves (Foreground shape correlation). Class- Flower.

Following points are observed:

- The P – R curves obtained are close to ideal P – R curves.
- Stricter similarity cut-off increases the Precision at the cost of Recall.
- Despite vast variations in foreground colors and illumination conditions, good Recall with good Precision is achieved for many sample queries.
- For all queries, Precision of 1.0 is achieved for reasonable Recall.
- Range of average performance measures for the class
 - 100 % of average Precision for 9 % of average Recall
 - 94 % of average Precision for 23% of average Recall
 - Giving only 6 % of fall in average Precision to raise average Recall by 14 %
- The exponentially extended trend line is well above average Precision = (0.5) line implies good performance measures for images of the class.

6.8.2.1 Query Response Example

Figure 82 shows query response of a flower image with similarity cut-off 60 giving 34 flower images based on shape comparison. It should be noted that for same query image, whole image color code based approach retrieves maximum of 4 flower images whereas foreground color code based approach retrieves only the query image even for lowest similarity cut-off. (Table 11 and Table 18 respectively.)

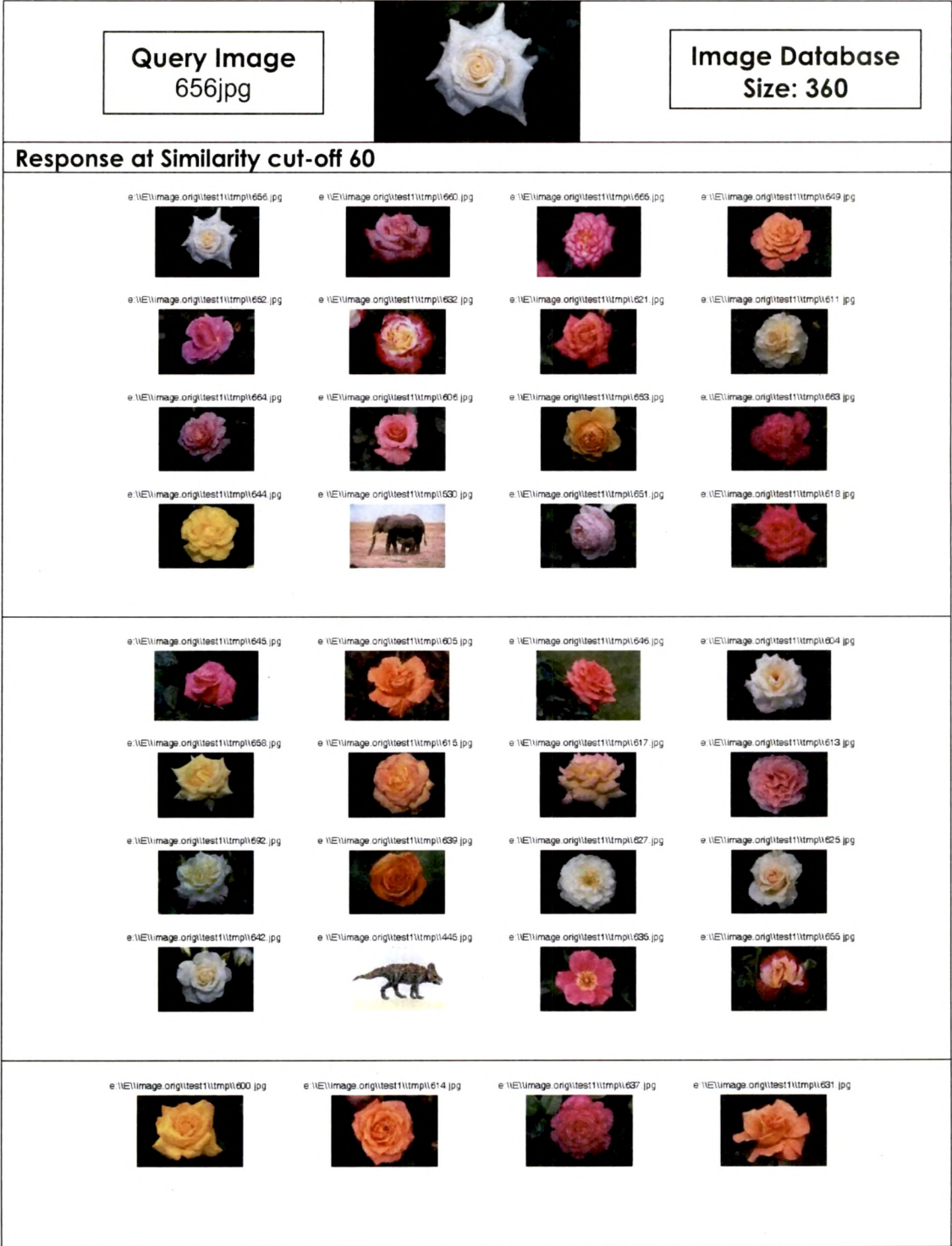


Figure 82. Query response of a flower image at similarity cut-off 60. (FG shape correlation)

6.8.3 Discussion

- The method is very sensitive to the shape of the foreground. Foreground object shape altering regions affect the performance of the method very adversely.
- The method gives very good results for images where foreground is not containing attached unwanted regions. E.g. images of ALOI database, images of class flower and class bus of SIMPLcity [Wang, 2001] [SIMPLcity, on line] database. The method may not perform equally well for the images like those of BSDb [Fowlkes, on line] [Martin, 2001].
- Relaxed similarity cut-off ends up in poor Precision. Recommended similarity cut-off is above 60% for better performance.
- Shape matching being a stricter constraint, relatively higher Precision and lower Recall are observed for the method.
- The shape correlation technique has been applied for face region matching for the purpose of similar-face image retrieval.

6.9 Foreground Color Codes & Foreground Shape Based CBIR

The proposed method compositely exploits two foreground features – shape and color codes. The weight proportion of these two features in the similarity measures is selectable by the user. Thus, the composite similarity measure signifies the proportionate emphasis of user's search criteria. The normalized unsegmented foreground region and foreground color codes have been utilized as the features for the image comparison. The method specific steps, replacing Step 3 & Step 4 of [Algorithm 4, Section 6.5](#) are:

Step 3: Read (or extract) foreground color code features of given query image.

Read (or extract) normalized unsegmented foreground region features for the query image. Call it R_q .

Step 4: For every image-feature-file of target folder,

Read corresponding foreground color code features of the image-feature-file of target folder.

Calculate (dis)similarity_index $i = \sum \text{abs}(h_{qi} - h_{ij})$, for $1 < j \leq \text{number of bins}$, Where,

h_{qj} indicates j^{th} bin of normalized histogram of color codes for the query image

h_{ij} indicates j^{th} bin of normalized histogram of color codes for i^{th} image of database

Read normalized unsegmented foreground region features of the image-feature-file of target folder. Let us call it Rd_i .

Calculate correlation coefficients of Rq & Rd_i .

Find the significant correlation coefficient.

Calculate $(dis)similarity_index1_i = 100 - abs(\text{significant correlation coefficient of } Rq \text{ \& } Rd_i) * 100$

Read **Foreground_Colorcode_weight**

Calculate composite $(dis)similarity$ index as

$(dis)similarity_index_i = (\text{Foreground_Colorcode_weight} * (dis)similarity_index) + ((1.0 - \text{Foreground_Colorcode_weight}) * (dis)similarity_index1)$

Algorithm 8. Foreground color codes & foreground shape based image retrieval.

6.9.1 Performance Evaluation

The performance of the method for different combinations of weights of foreground color code and foreground shape correlation in composite similarity index is tabulated for an image (455.jpg [Wang, 2001] [SIMPLicity, on line]) of Dinosaur class is shown in Table 22. The respective P – R curves are presented in Figure 83.

Table 22. Precision, Recall & F –measure for different proportionate weights at different similarity cut-offs. (Foreground Color codes & foreground shape correlation).

Similarity cut-off for 455.jpg	% Weight of FG CC in Similarity Index	Retrieved relevant images - rr	total retrieved images - total	Recall $r = rr / \text{Total}$	Precision $p = rr / \text{total}$	F measure
Total relevant images in the database, Total = 100						
50	0	29	82	0.29	0.35	0.32
60		7	17	0.07	0.41	0.12
70		1	1	0.01	1.00	0.02
50	10	29	46	0.29	0.63	0.40
60		3	5	0.03	0.60	0.06
70		1	1	0.01	1.00	0.02
50	20	26	38	0.26	0.68	0.38
60		3	3	0.03	1.00	0.06
70		1	1	0.01	1.00	0.02

Table 22 (Contd.). Precision, Recall & F –measure for different proportionate weights at different similarity cut-offs.
(Foreground Color codes & foreground shape correlation).

Similarity cut-off for 455.jpg	% Weight of FG CC in Similarity Index	Retrieved relevant images - rr	total retrieved images - total	Recall $r = rr / \text{Total}$	Precision $p = rr / \text{total}$	F measure
Total relevant images in the database, Total = 100						
50	30	25	34	0.25	0.74	0.37
60		6	7	0.06	0.86	0.11
70		1	1	0.01	1.00	0.02
50	40	26	34	0.26	0.76	0.39
60		8	9	0.08	0.89	0.15
70		1	1	0.01	1.00	0.02
50	50	27	36	0.27	0.75	0.40
60		12	15	0.12	0.80	0.21
70		2	2	0.02	1.00	0.04
50	80	29	35	0.29	0.83	0.43
60		23	28	0.23	0.82	0.36
70		12	14	0.12	0.86	0.21
50	100	30	38	0.3	0.789474	0.43
60		21	26	0.21	0.807692	0.33
70		16	20	0.16	0.8	0.27

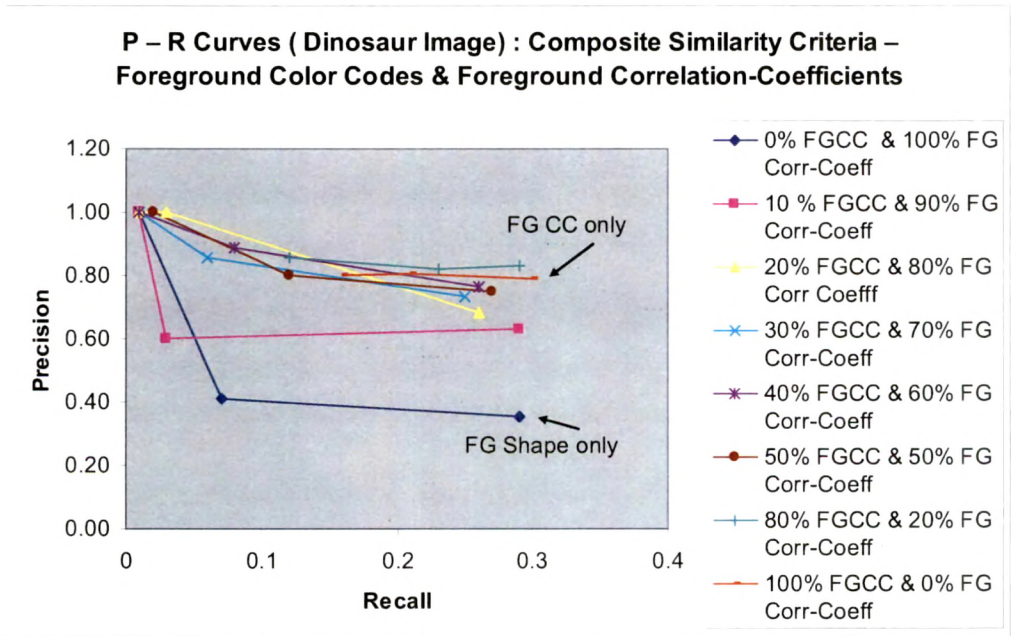




Figure 83. P- R curves for different proportionate weights of Foreground color codes & foreground shape correlation.

The suitable proportion of weight for best retrieval performance is image specific. The user is given a choice of selecting the weight and altering the proportion if required for successive retrievals. The Precision and Recall for two images of flower given as queries with 70% - 30% & 30%-70% weight proportion (Foreground CC & Foreground shape correlation) has been computed in Table 23. The high Precision is noteworthy.

Table 23. Precision, Recall & F –measure for two different proportionate weights at different similarity cut-offs. (Foreground Color codes & foreground shape correlation).

Query Image	Similarity cut-off	% Weight of FG CC in Similarity Index	Retrieved relevant images - rr	total retrieved images - total	Recall $r = rr / \text{Total}$	Precision $p = rr / \text{total}$	F measure $= 2 / (1 / p + 1 / r)$
Total relevant images in the database, Total = 100							
 606.jpg	40	70	28	28	0.28	1.00	0.44
	50		16	16	0.16	1.00	0.28
	60		14	14	0.14	1.00	0.25
	70		3	3	0.03	1.00	0.06
	40	30	35	35	0.35	1.00	0.52
	50		18	18	0.18	1.00	0.31
	60		12	12	0.12	1.00	0.21
	70		4	4	0.04	1.00	0.08
 644.jpg	40	70	19	30	0.19	0.63	0.29
	50		12	19	0.12	0.63	0.20
	60		10	11	0.1	0.91	0.18
	70		7	7	0.07	1.00	0.13
	40	30	18	24	0.18	0.75	0.29
	50		14	14	0.14	1.00	0.25
	60		11	11	0.11	1.00	0.20
	70		6	6	0.06	1.00	0.11

6.9.1.1 Query Response Examples:

Figure 84 and Figure 85 show respective retrieval results for a flower image with 20 % and 10 % weight of foreground color codes in the composite similarity index. The reduction in weight of foreground color codes and corresponding increase in the weight of foreground shape correlation results into retrieval of more flower images, not necessarily while flower images. Image retrieval with same query image for foreground color code method (100 % weight of foreground color codes), only the query image gets retrieved (Table 11). The reduction of similarity cut-off results into retrieval of more images as shown in Figure 86 & Figure 87.

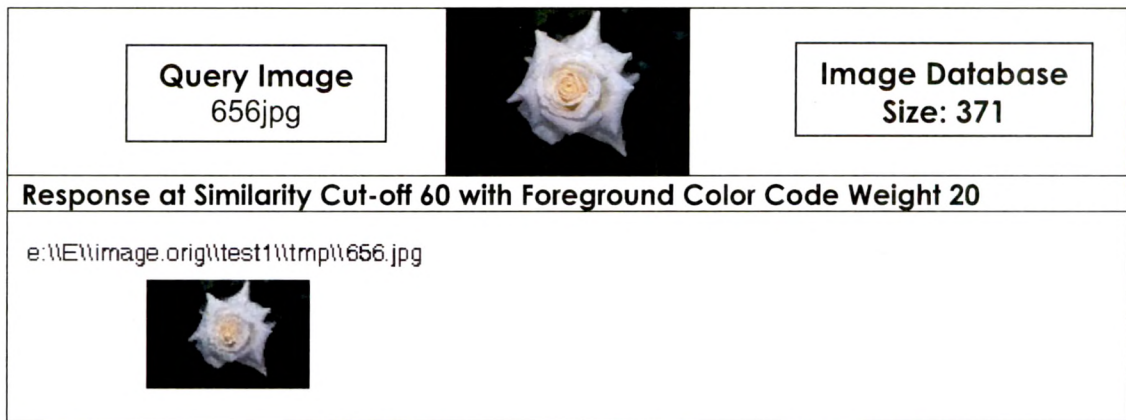


Figure 84. Query response of a flower image at similarity cut-off 60 with FG CC weight 20. (FGCC & FG shape correlation)

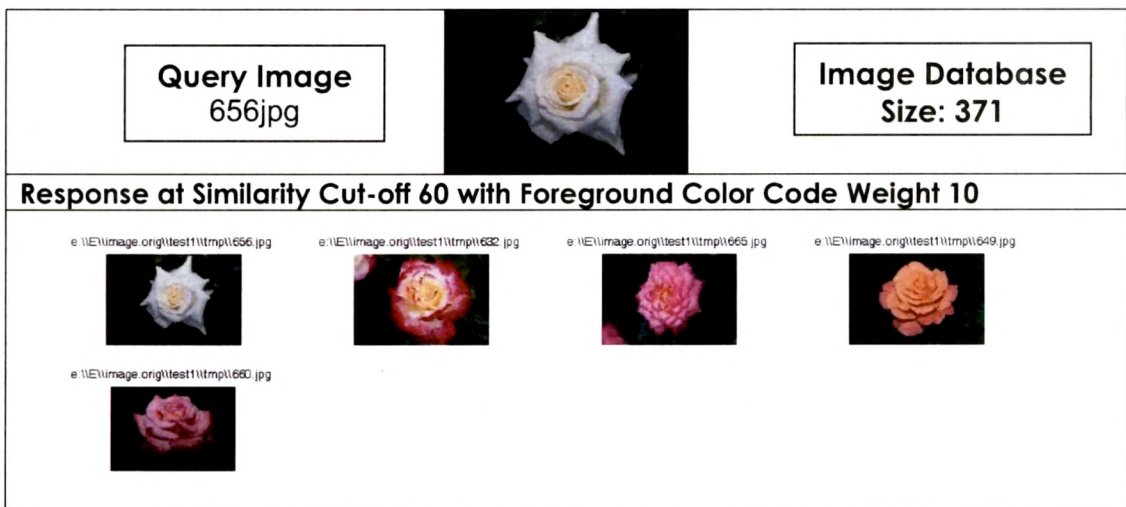


Figure 85. Query response of same flower image at similarity cut-off 60 with FG CC weight 10. (FGCC & FG shape correlation)

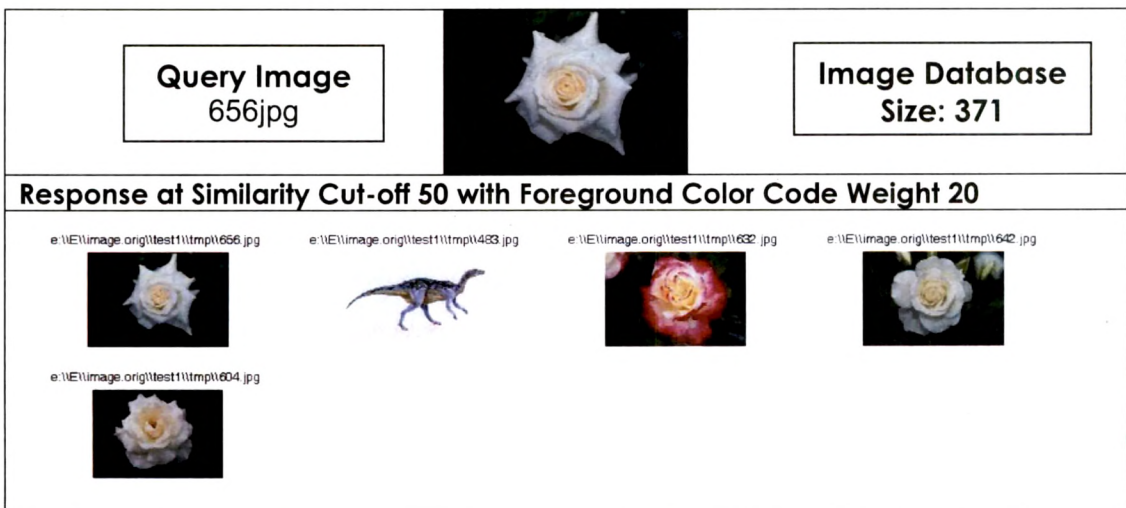


Figure 86. Query response of same flower image at similarity cut-off 50 with FG CC weight 20. (FGCC & FG shape correlation)

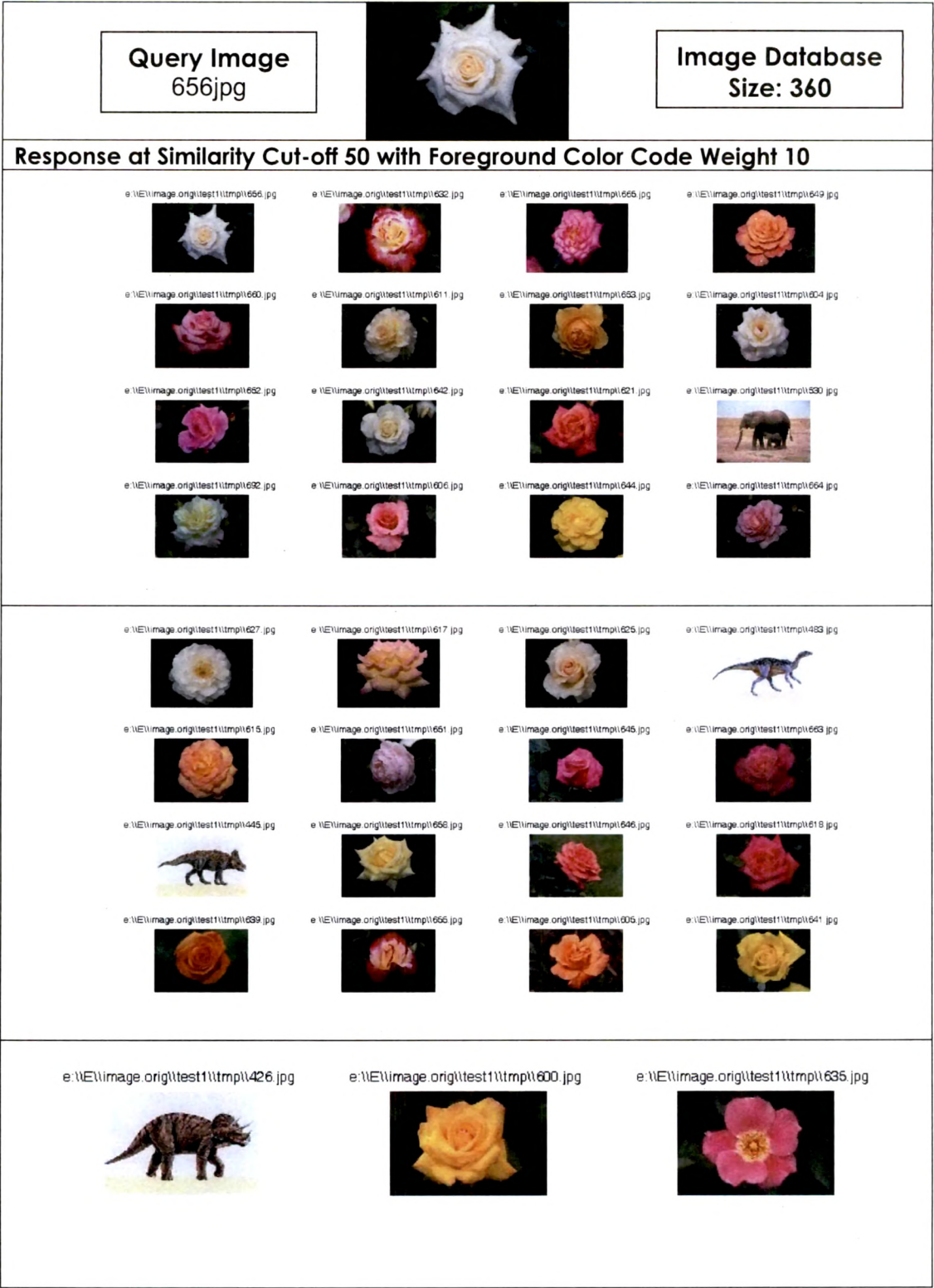


Figure 87. Query response of same flower image at similarity cut-off 50 with FG CC weight 10. (FGCC & FG shape correlation)

6.9.2 Discussion

- The method exploits advantages of all previously proposed methods. The foreground based approach eliminates unwanted, major contributing background and related features enabling object based search with additional constraints of foreground shape and colors.
- The selectable proportion of weight of foreground color codes and foreground shape ends up in good performance of the system.

6.10 Comparisons - Query Responses of Various Algorithms

The section provides comparison of query responses of various proposed method of image retrieval for same query images. The first example is for comparison of response of various methods for a flower image of SIMPLcity [Wang, 2001] [SIMPLcity, on line] image database whereas the second example is for the ALOI image data base [ALOI, on line] [Geusebroek, 2001]. Typical characteristics of the databases are described in [Section 6.4](#).

6.10.1 Example 1 - SIMPLcity image database [Wang, 2001] [SIMPLcity, on line]

Figure 88 to Figure 91 are the respective query responses of proposed four methods for same query-image with same similarity cut-off of 60. The query response with whole image color codes based retrieval gives 2 similar flower images shown in Figure 88. The response of same query image with foreground based color code approach yields 15 similar – red flower images as shown in Figure 89. The response with foreground shape correlation method is shown in Figure 90 indicates retrieval of 39 flower images (not only red) and 2 non-flower images. The foreground color codes and foreground shape based approach with 30% weight to foreground color method yields response shown in Figure 91 giving 12 similarly shaped red flower images.

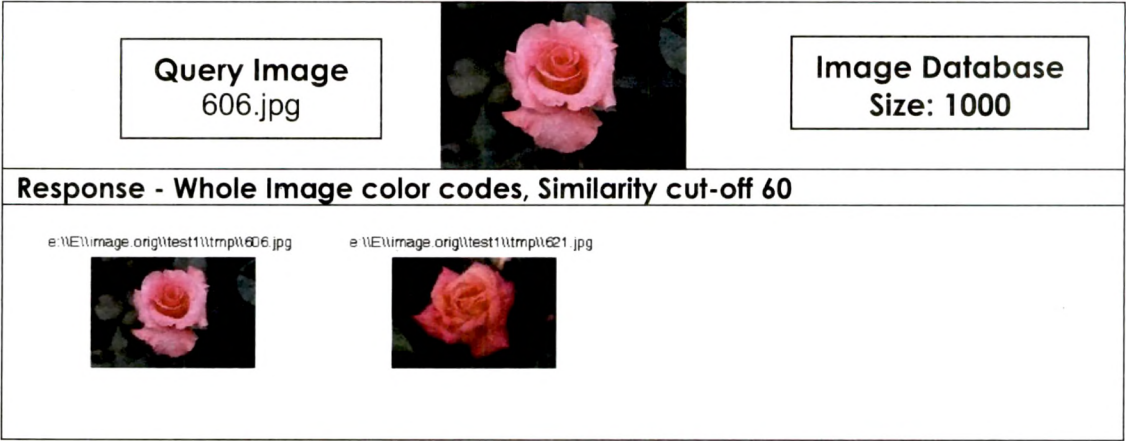


Figure 88. Query response of a flower image at similarity cut-off 60. (Whole image color codes).



Figure 89. Query response of same flower image at similarity cut-off 60. (Foreground color codes).



Figure 90. Query response of same flower image at similarity cut-off 60. (Foreground shape correlation).

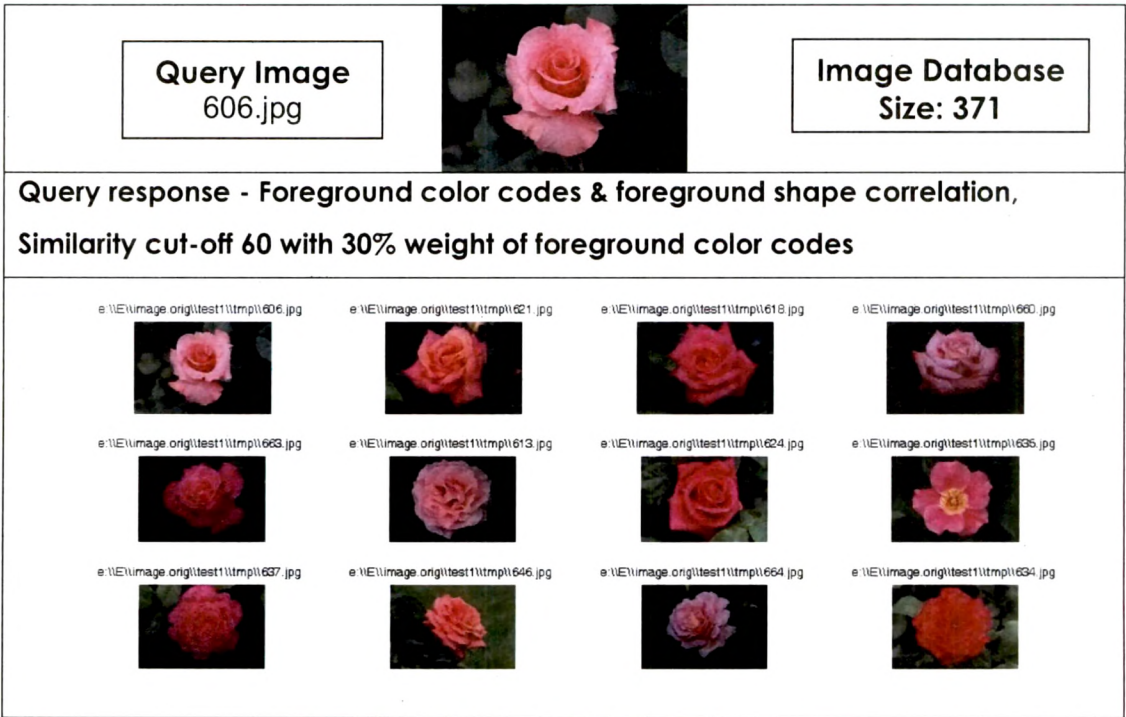


Figure 91. Query response of same flower image at similarity cut-off 60 with 30% weight of FGCC. (Foreground color codes & foreground shape correlation).

6.10.2 Example 2 - ALOI image database [ALOI, on line] [Geusebroek, 2001]

The effect of illumination changes and object view point variations on image retrieval has been illustrated with following query responses of various methods for same query image – a toy duck for similarity cut-off of 70. The major portion of the background in the image causes poor Precision for whole image color code method as shown in Figure 92. The foreground color code based approach improves the Precision by giving yellow / white colored toy images as the response as shown in Figure 93. The foreground shape based method for retrieval gives good Recall and Precision with duck toys ranked higher, as shown in Figure 94. The foreground based composite approach with color codes and shape correlation with 30% weight of foreground color codes give the best performance as shown in Figure 95.

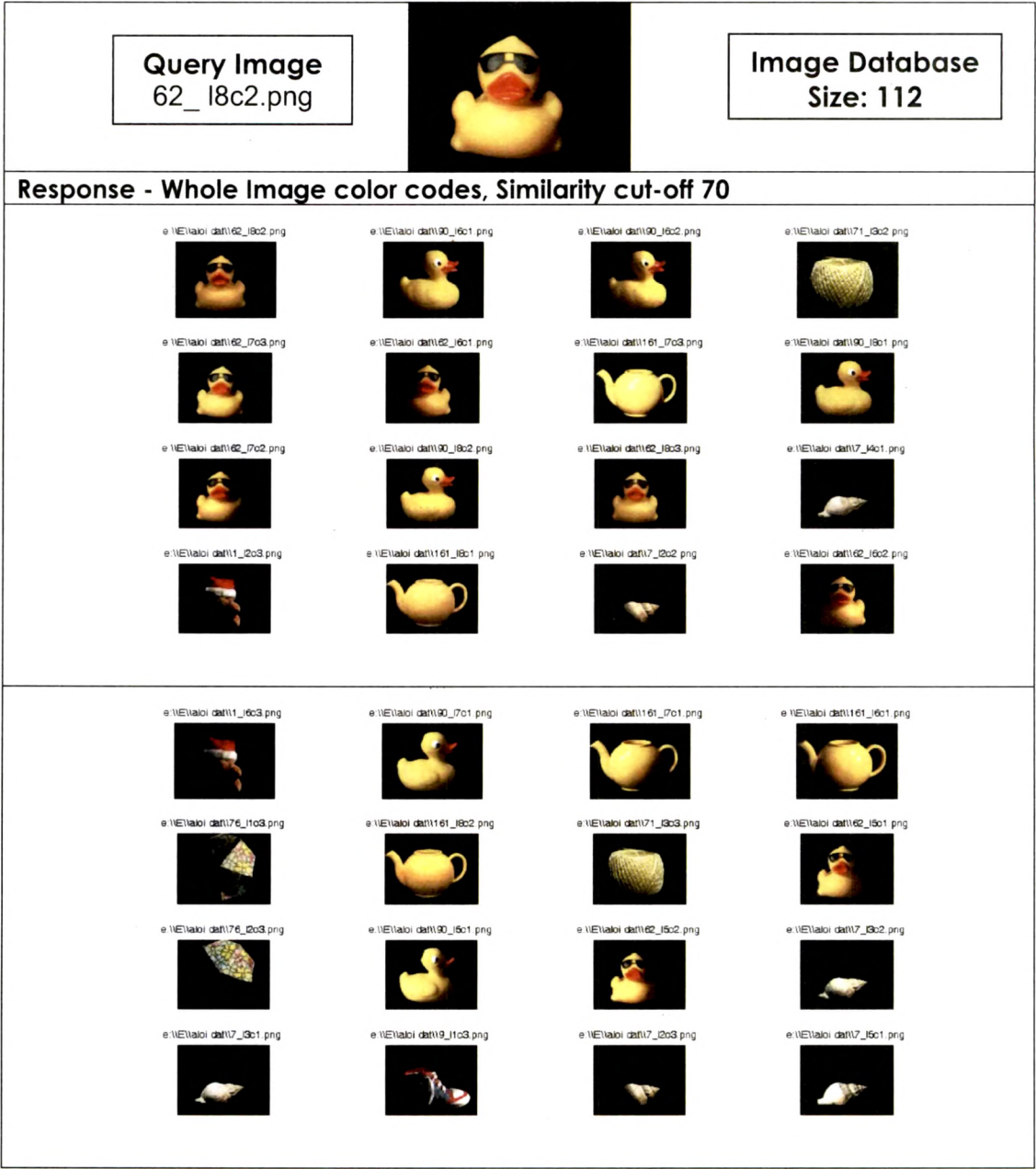


Figure 92. Query response of an ALOI image at similarity cut-off 70. (Whole image color codes).

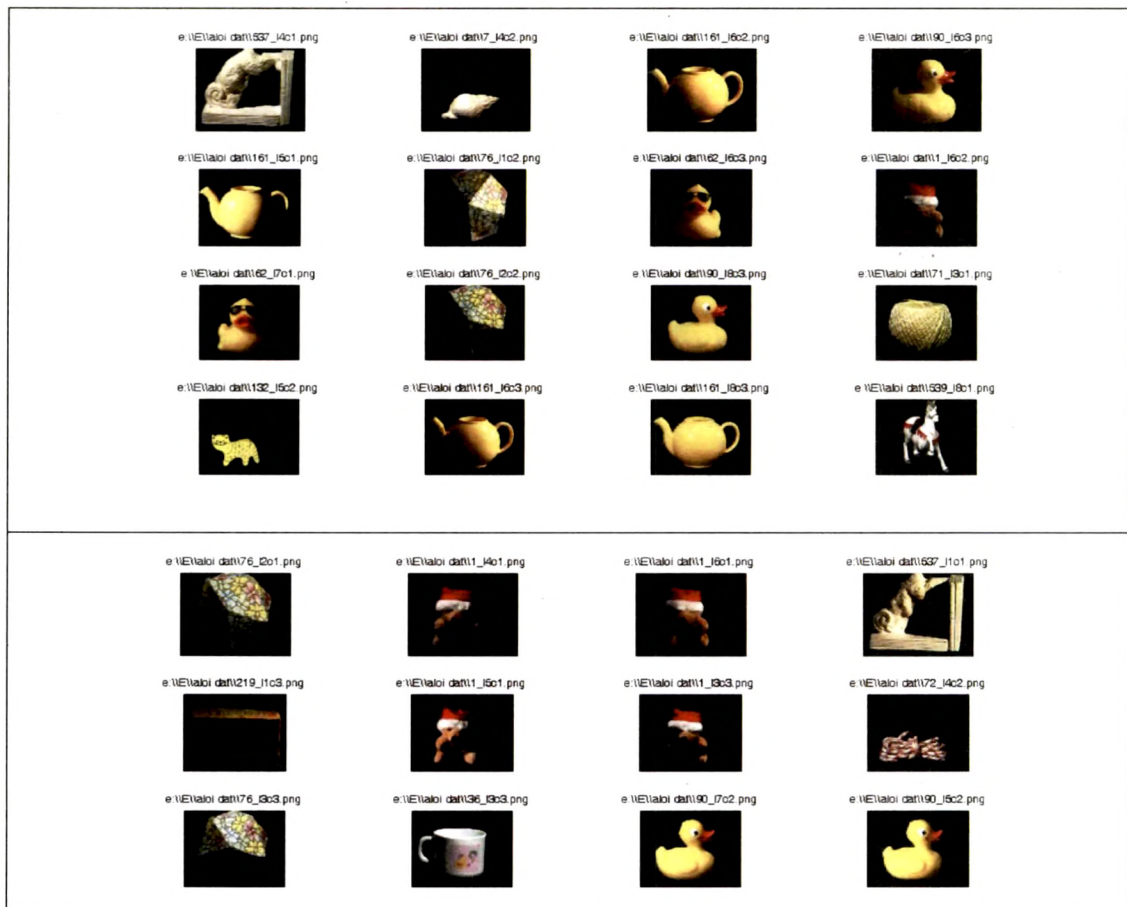


Figure 92 (Contd.). Query response of an ALOI image at similarity cut-off 70. (Whole image color codes).

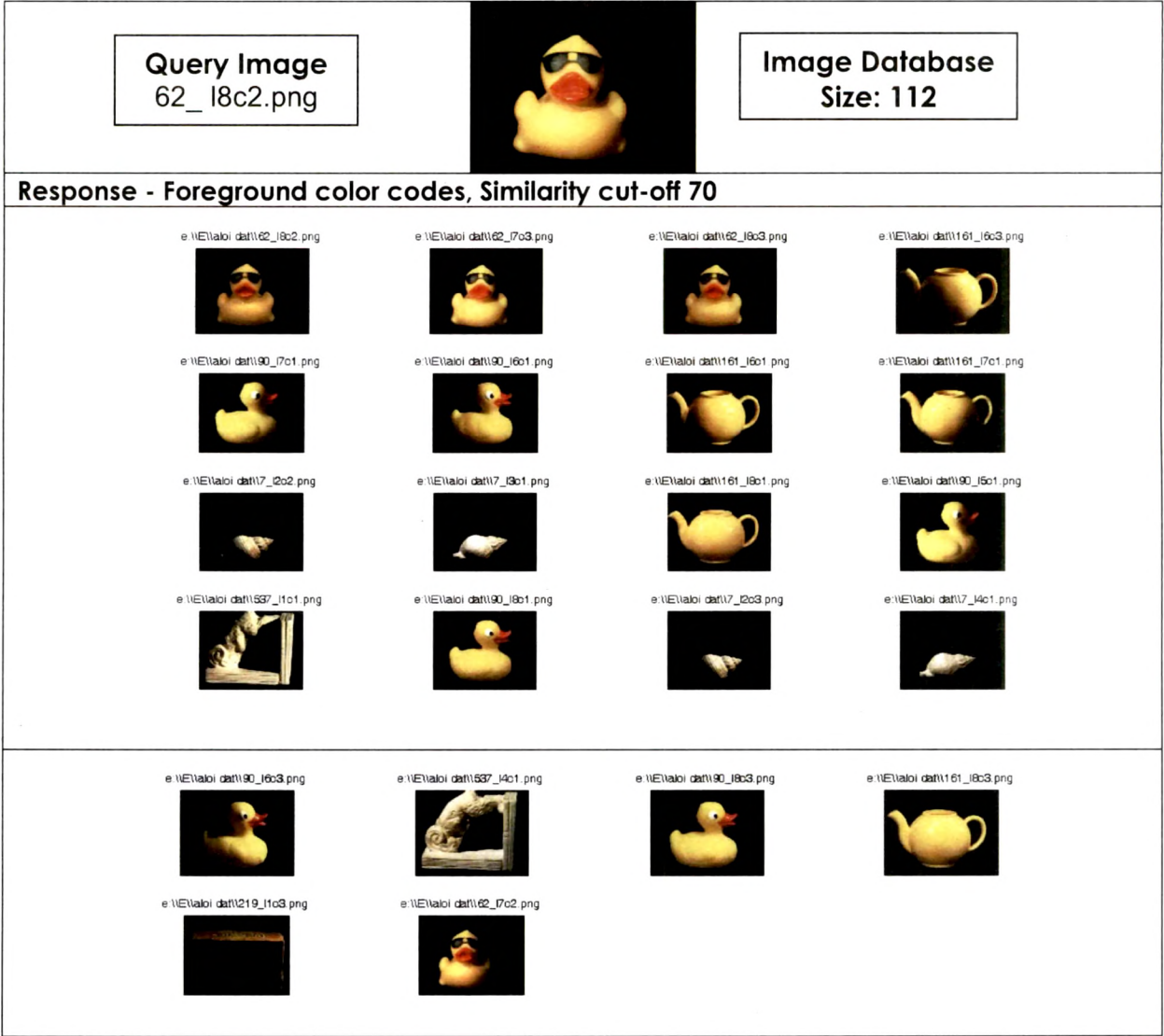


Figure 93. Query response of same ALOI image at similarity cut-off 70. (Foreground color codes).

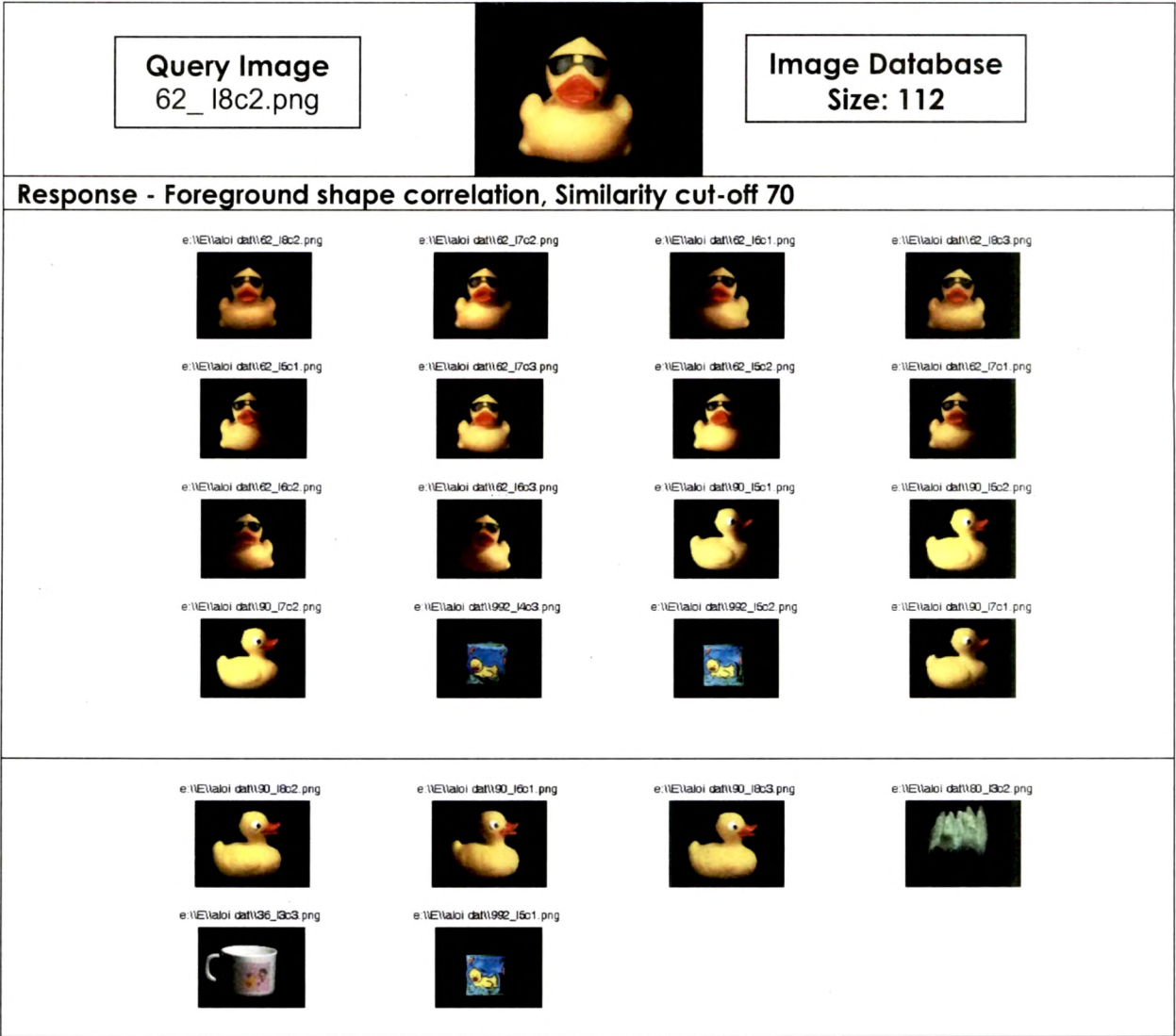


Figure 94. Query response of same ALOI image at similarity cut-off 70. (Foreground shape correlation).

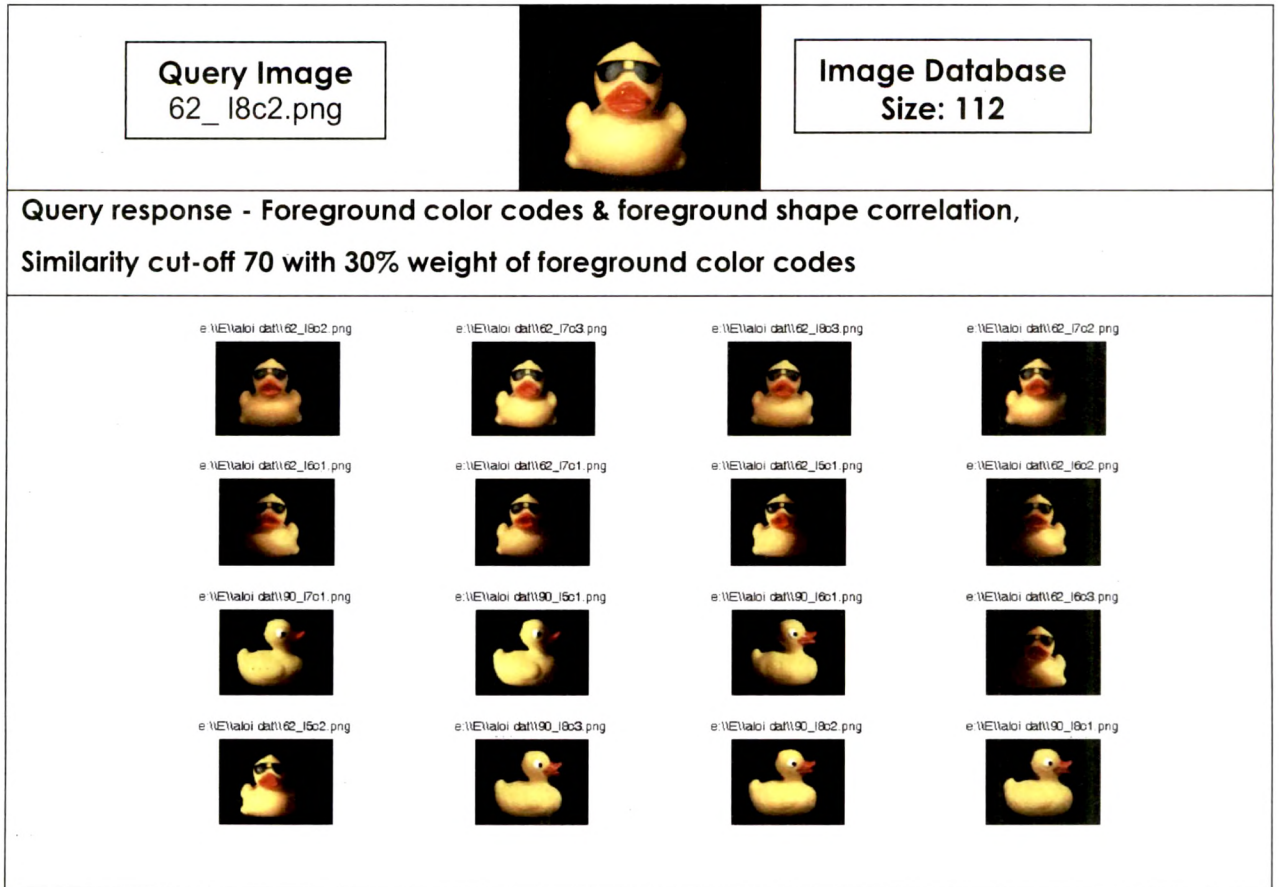


Figure 95. Query response of same ALOI image at similarity cut-off 70 with 30% weight of FGCC. (Foreground color codes & foreground shape correlation).

6.10.3 Discussion

- The whole image color code based approach is well suitable for finding images having similar foreground-background color code distribution. High Recall values are possible because of broader color descriptors. The method faces conventional limitations because of not considering shapes or any other regional features.
- The foreground region based approaches enable foreground object based image search by considering foreground color codes and foreground shape correlation. The foreground color code based approach permits retrieval of foreground shape variant similar images. The foreground shape correlation has been a stricter comparison and very sensitive to extracted foreground shape.
- The composite similarity measure of foreground color codes & foreground shape performs well for majority of query images.

6.11 Application Specific CBIR - Similar Face Image Retrieval

This section is an application of proposed methods for application-specific CBIR to retrieve similar face images from images containing complex background. The feature extraction phase consists of face region extraction technique based on prominent boundaries detection based foreground separation. The similarity measure for face regions is the shape correlation comparison as described in [Section 6.8](#). The high success ratio of precise face region extraction for complex backgrounds and illumination variations has been exploited for shape correlation based similarity comparison for image retrieval.

6.11.1 Face Extraction from Images Containing Complex Background

The section proposes novel method for human frontal face extraction from color images characterized by uncontrolled illumination conditions for image capturing and complex backgrounds. The method incorporates stationary Haar wavelet transform & proximity influence for prominent boundaries detection and watershed transform, proximity influence & morphological operations to separate foreground / background along with region and color attributes for human face extraction. The method exploits redundancy by coalescing local color cues of all color channels to emphasize reliable processing to precisely detect the human face by avoiding under-segmentation and reducing over-segmentation & artifacts. The method has been tested on face-image collection of standard database and on images captured by an amateur photographer for various complex backgrounds having diversified textures, varied illumination conditions and multiple background objects. The presented results show the effectiveness of the method for frontal face extraction, proving it suitable as an input to applications like digital album catalogue, content based image retrieval, face recognition and facial expression recognition.

First well-thought out algorithm for image segmentation would be a watershed algorithm [Beucher, 1979]. The watershed algorithm has inherent characteristic of finding local minima – catchments basins producing over-segmentation of regions and introducing artifacts. Hence, watershed transform cannot be applied directly on images having textures, texture or smooth color tone variations e.g. natural images. Any filtering-preprocessing before applying watershed algorithm results into loss of information introducing either leaks in the boundaries or spurious boundaries leading to improper segmentation. Hence, other techniques required to be combined with

watershed algorithm to overcome short-comings of watershed algorithm exploiting advantages of the same. So has been done in the proposed method.

The face extraction is a process of isolating face region from all other regions. The performance of the face extraction is challenged by characteristics like illumination variations, shadows, skin-colored other regions, face-shaped other regions, multiple objects, diversified indoor & outdoor background textures, wide range of face skin colors and hair colors, different hair styles, different face orientations, different image resolutions along with aforesaid image segmentation issues as image segmentation being the inevitable first step of the process.

The issue of extraction of human face from images captured in uncontrolled illumination conditions having complex & non-uniform background has been addressed in the proposed method by performing proper segmentation followed by foreground / background separation and face region extraction. The image characteristics - illumination variations and diversity & variations in the background textures impose challenges at the segmentation and face extraction phase. The proposed method enforces reliable processing of local color cues of all color and gray channels for forming continuity preserving prominent boundaries incorporating Stationary Haar wavelet at various levels. The prominent boundaries, proximity influence and watershed transforms are compositely used for revealing foreground from the image. This foreground may consist of human face, hair and attached regions due to complex & non-uniform background. Various region attributes along with color attributes are used to extract the face. The method overcomes issues of under-segmentation by precise processing of low level cues generating well localized, delineated leak-free boundaries which are further categorized as prominent boundaries encompassing visually prominent regions in the image. The method minimizes over-segmentation & artifacts producing proper segmentation needed for face extraction in majority of the cases containing illumination variations. The incorporation of Stationary Haar decomposition at various levels makes method suitable for hierarchical framework. The performance of the method has been evaluated on face images of standard dataset Caltech 101 [Caltech, on line] [Fei-Fei, 2004].

6.11.1.1 The Method

The proposed method exploits redundancy by finally coalescing local cues of R, G, B and Gray color channels for prominent boundaries detection and for formation of composite watershed regions utilized for face extraction.

The region attributes considered for determining the face region are defined in the method regionprops of Matlab 14 as follows: Orientation - The angle (in degree) between the x-axis and the major axis of the ellipse that has the same second-moments as the region. Extent – The portion of the pixels in the bounding box that are also in the region. Eccentricity – Measured for the ellipse that has the same second-moments as the region. It is the ratio of the distance the foci of the ellipse and its major axis length.

The steps involved for the face extraction are:

Step 1: Apply foreground extraction Step 1 to Step 10 of [Algorithm 3, Section 5.2](#).

Step 2: For all regions of foreground,

Exclude small regions.

For remaining regions,

Mark a region as face region if orientation > 70, Eccentricity > 0.3, extent < 0.95, axis_ratio (Minor axis length / Major axis length) > 0.4 and if region contains skin color. Mark image category as face.

Step 3: Map face region on the image to extract face image.

Algorithm 9. Face extraction from images containing complex background.

The thresholds are empirically determined to reflect shape attributes of face.

6.11.1.2 Results

The method has been tested on test set consists of 115 face-images of Caltech 101 face dataset [Caltech, on line] [Fei-Fei, 2004] comprising of faces of 15 persons and other high resolution images captured by an amateur photographer. The test set images are selected to cover various illumination variations and backgrounds. The images containing mustache and bearded face are excluded from the test set. The Caltech dataset [Caltech, on line] [Fei-Fei, 2004] found most appropriate as the data set because it contains color face images of medium size and reasonable resolution. The multiple images of persons have been captured at various indoor-outdoor places with various complex backgrounds under different illumination conditions generating shadows and illumination variations. Figure 96 illustrates results of various phases of face extraction incorporating two different levels of wavelet decomposition – level 1 and level 2. The prominent boundaries detected for Gray

channel has been shown in Figure 96 (b). Various regions of segmented color channels have been shown in Figure 96 (c) to Figure 96 (f). The composite regions of segmented image are shown in Figure 96 (g). The watershed pixels constituting region isolating boundaries have been shown in Figure 96 (h). Figure 96 (i) and Figure 96 (j) corresponds to separated background and foreground respectively. The foreground is marked as black in the background image and background is marked as black in the foreground image. The extracted faces are finally shown in Figure 96 (k). In general, reduction of image dimension by a large factor may adversely affect the segmentation performance due to interpolated pixel color values. The unaffected result of face extraction for a size reduced high resolution image has been shown in Figure 97. Figure 98 illustrates unsuccessful face extraction due to improper segmentation because of extreme illumination variations on the face. The extracted faces of some of the images of test set have been shown in Figure 99. The typically picked up images contain distinctiveness like non-face skin color regions, other face-shaped regions, skin-colored hair, off-centered face, multiple background objects and most importantly illumination variations due to different lightning conditions at indoor-outdoor locations. Table 24 depicts the effectiveness and robustness of the method for successfully extracting faces for 82.6 % of the images of the test set.

6.11.1.3 Discussion

- The proposed method is novel for prominent boundaries, proximity influence, Stationary Haar Wavelet, used for generation of composite watershed regions for foreground separation that is combined with face-region & face-color attributes for frontal face extraction.
- The well localized and delineated continuity preserving prominent boundaries detected by precise and reliable processing of low level color cues of all color and gray channels form the basis of high (82.6 %) successful face extraction ratio for 115 test images of Caltech 101 [Caltech, on line] [Fei-Fei, 2004] dataset.
- The extraction of face has been tested on various images containing performance affecting characteristics like illumination variations, shadows, skin-colored other regions, face-shaped other regions, multiple objects, diversified indoor & outdoor complex-background textures, wide range of

- face skin colors and hair colors, different hair styles, different face orientations etc. The stationary Haar wavelet decomposition at various levels, prominent boundaries & proximity influence avoids under-segmentation and reduces over-segmentation & artifacts – inherent characteristics of watershed transform.
- The method results are not affected due to interpolation operation involved in a size reduction of a high resolution image by a large factor.
 - As shape and color of the face region is largely altered due to mustache & bearded, the method fails for face extraction of such cases. Similarly, a dark face segmenting shadow also produces ill results of face extraction.
 - Still, precisely extracted faces with high performance ratio for variety of images proves the suitability and versatility of the method for applications like digital album catalogue, content based image retrieval, face recognition and facial expression recognition.

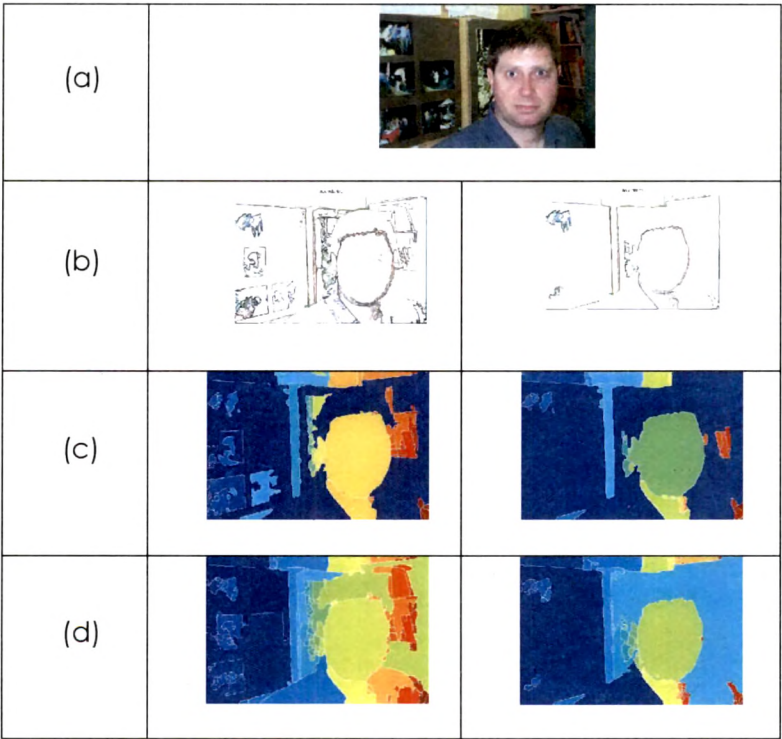


Figure 96. Various steps of face extraction. Left - stationary Haar wavelet decomposition at level 1. Right - stationary Haar wavelet decomposition at level 2. (a) Original image [Caltech, on line] [Fei-Fei, 2004]. (b) Detected prominent boundaries of gray channel. (c) Segmented regions of red color channel. (d) Segmented regions of green color channel.

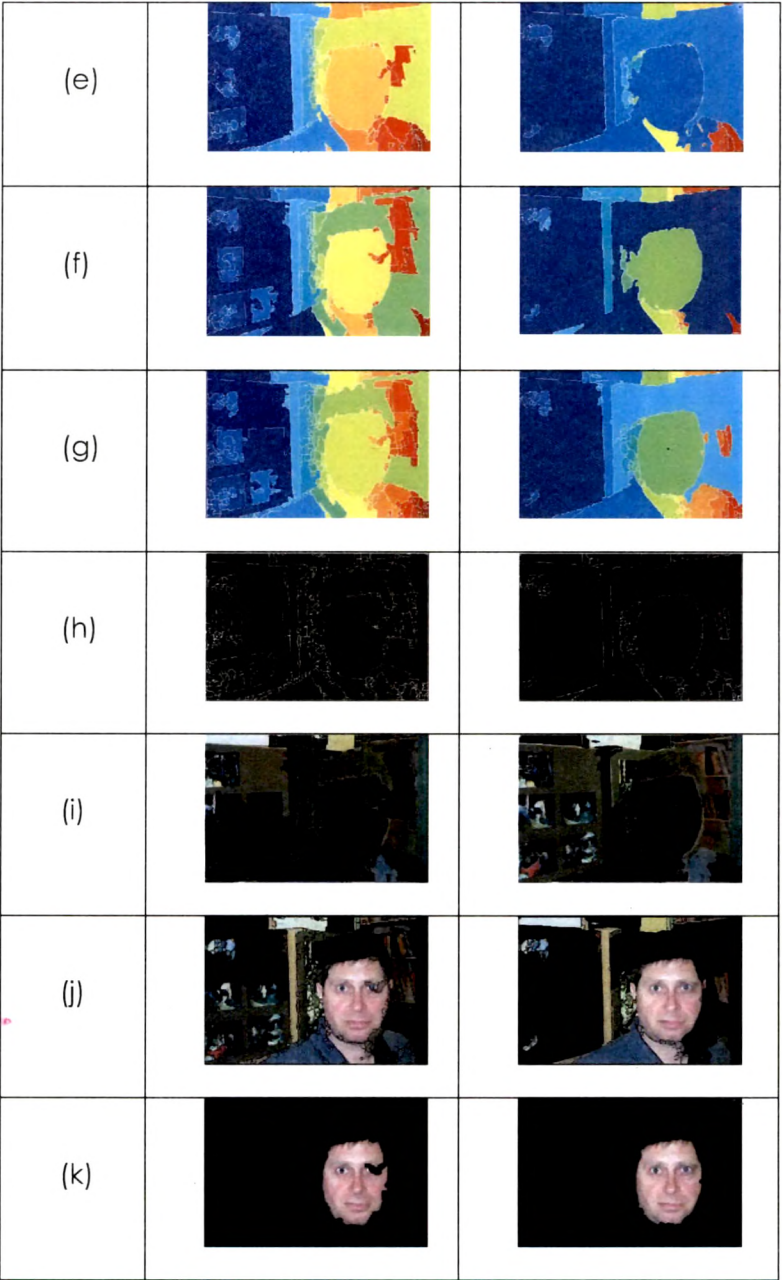


Figure 96 (Contd.). Various steps of face extraction. (e) Segmented regions of blue color channel. (f) Segmented regions of gray color channel. (g) Composite segmented regions. (h) Corresponding watershed pixels of (g). (i) Background. (j) Foreground. (k) Extracted face.

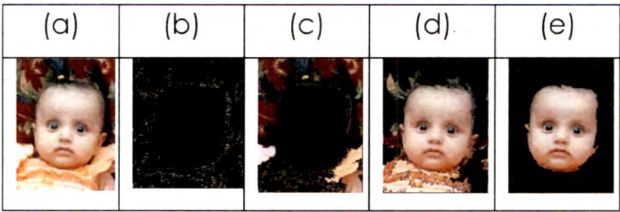


Figure 97. Face extraction in high resolution image reduced to 1/8th of the original size. (a) Original image. (b) Watershed pixels. (c) Background. (d) Foreground. (e) Extracted face.

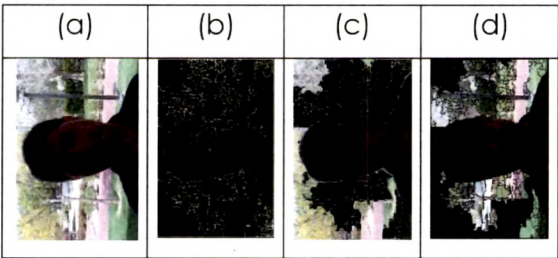


Figure 98. Example of unsuccessful face extraction. (a) Original image [Caltech, on line] [Fei-Fei, 2004]. (b) Watershed pixels. (c) Background. (d) Foreground.

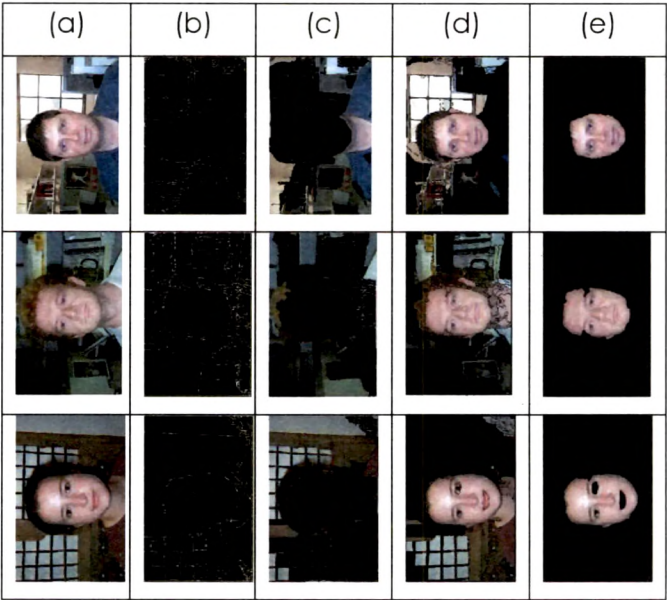















Figure 99. Face extractions of various images with complex background and non-uniform illuminations. (a) Original images [Caltech, on line] [Fei-Fei, 2004]. (b) Watershed pixels. (c) Backgrounds. (d) Foregrounds. (e) Extracted face.



Figure 99 (Contd.). Face extractions of various images with complex background and non-uniform illuminations. (a) Original images [Caltech, on line] [Fei-Fei, 2004]. (b) Watershed pixels. (c) Backgrounds. (d) Foregrounds. (e) Extracted face.

Table 24. Performance evaluation of face extraction method for various images [Caltech, on line] [Fei-Fei, 2004].

Person id	Sample Images	No. of Successful face extraction / Total images of the person in the test set	Successful face extraction %
Person 1		17 / 21	80.9
Person 2		7 / 11	63.6
Person 3		4 / 4	100
Person 4		11 / 14	78.5
Person 5		14 / 16	87.5
Person 6		5 / 7	71.4
Person 7		11 / 13	84.6
Person 8		4 / 4	100
Person 9		4 / 4	100
Person 10		3 / 5	60
Person 11		4 / 4	100
Person 12		5 / 6	83.3
Person 13		4 / 4	100
Other		2 / 2	100
Total		95 / 115	82.6



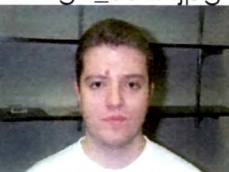


6.11.2 Similar Face Image Retrieval

The method described in [Algorithm 7, Section 6.8](#) is applied on precisely extracted face region for comparison of shape correlation coefficients to retrieve similar face images.

6.11.2.1 Performance Evaluation

The performance measure Precision & Recall have been tabulated in Table 25 for 5 sample images of Caltech database.

Table 25. Precision, Recall & F – measure at different similarity cut-offs. Similar face-image retrieval.

Query Images Caltech [Caltech, on line] [Fei-Fei, 2004]	Similarity cut-off	Retrieved relevant images - rr	total retrieved images - total	Total relevant images in the database - Total	Recall $r = rr /$ Total	Precision $p = rr /$ total	F - measure
Image_0001.jpg 	50	17	60	21	0.81	0.28	0.42
	60	15	25		0.71	0.6	0.65
	70	6	8		0.29	0.75	0.42
	80	2	2		0.10	1	0.18
Image_0007.jpg 	50	18	71	21	0.86	0.25	0.39
	60	16	54		0.76	0.30	0.43
	70	10	24		0.48	0.42	0.45
	80	7	7		0.33	1	0.50
Image_0047.jpg 	50	8	59	14	0.57	0.14	0.22
	60	8	45		0.57	0.18	0.27
	70	7	19		0.5	0.37	0.43
	80	3	3		0.21	1	0.35
Image_0079.jpg 	50	5	21	16	0.31	0.24	0.27
	60	5	6		0.31	0.83	0.45
	70	5	5		0.31	1	0.47
	80	4	4		0.25	1	0.40
Image_0122.jpg 	60	4	25	13	0.31	0.16	0.21
	70	3	4		0.23	0.75	0.35
	80	1	1		0.08	1	0.15

The average Precision and average Recall at different similarity cut-offs have been shown in Table 26. The corresponding P – R curves are plotted in Figure 100.

Table 26. Average Recall, average Precision & average F – measure at different similarity cut-off. (Similar face image retrieval).

Similarity cut-off	Average Recall	Average Precision	Average F-measure
50	0.51	0.23	0.32
60	0.44	0.6	0.51
70	0.36	0.75	0.49
80	0.19	1	0.32

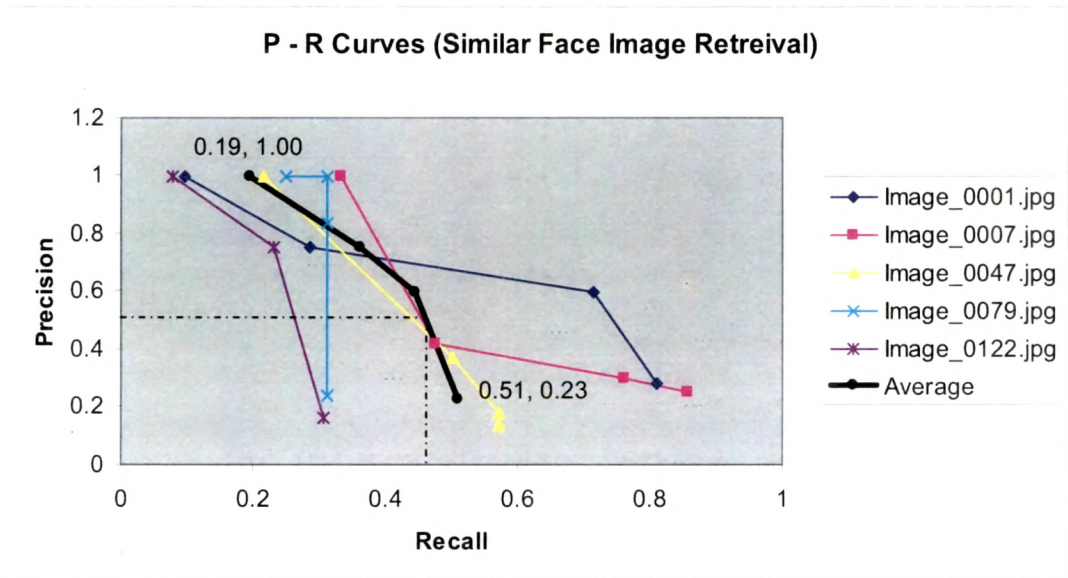


Figure 100. P – R curves. Similar face image retrieval.

Following points are observed:

- The nature of obtained P – R curves matches with the practical P – R curves.
- Stricter similarity cut-off increases the Precision at the cost of Recall.
- Despite vast variations in foreground / background colors, poses and illumination conditions, good Recall with good Precision is achievable for many sample queries.
- For all queries, Precision of 1.0 is achieved at higher cut-off.
- Range of average performance measures for all queries
 - 100 % of average Precision for 19 % of average Recall
 - 23 % of average Precision for 51 % of average Recall
 - Giving 77 % of fall in average Precision to raise average Recall by 32 %
- Lower similarity cut-offs – less than 60, are not recommended for good performance.

- The exponentially extended trend line intersects average Precision = (0.5) line at average Recall at value 0.45 (approx.) implies good performance measures for images of the class.

6.11.2.2 Query Response Example

The query response for retrieving similar-face images with 70 as similarity cut-off is shown in Figure 101. The illustrated response signifies Precision of 75% for 28% of Recall for images having complex backgrounds and face-regions constituting a small portion of images in a database of 115 images.

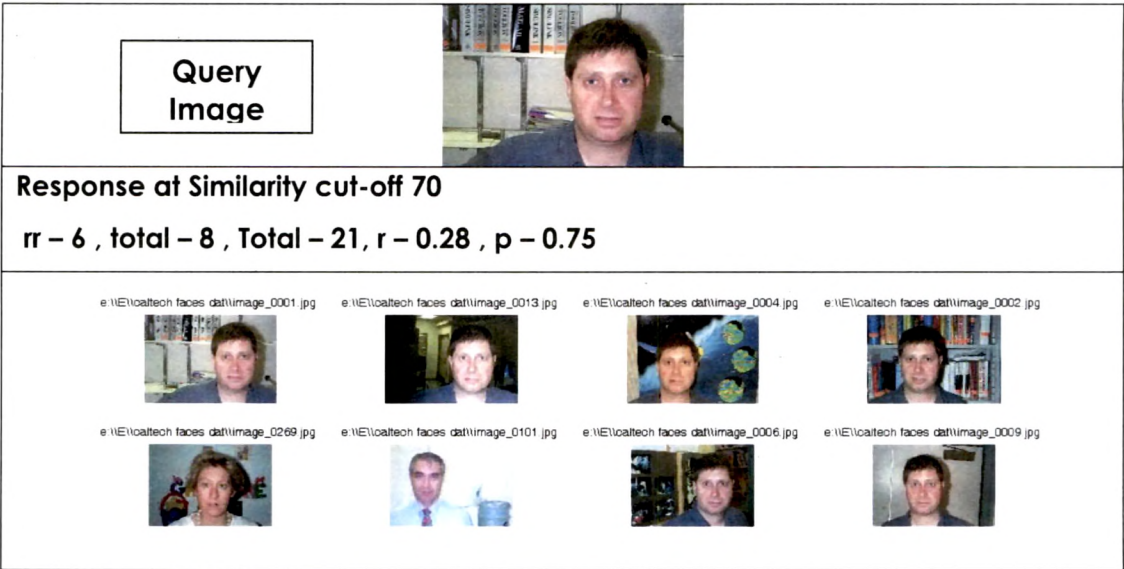


Figure 101. Query Response – similar face-image retrieval.

6.11.2.3 Discussion

- The effectiveness of methods for prominent boundaries detection & foreground separation for precisely extracting face by excluding complex background has been utilized for obtaining face-features and similar-face images. The method incorporates only face shape-feature for similarity comparison.
- Inclusion of other face-features for similarity comparison will improve the performance of the system.

6.12 Performance Comparisons with other CBIR Techniques

The relative performance comparisons of proposed methods with the state of the art CBIR techniques is not feasible because (i) unavailability and un-disclosure of comprehensive technical details of state-of-the-art techniques which are commercial, proprietary or patented. (ii) Available on-line demos of some of the CBIR systems

neither reveal the details of the image database needed for Precision & Recall computations nor permit our exclusive image database to be uploaded for testing and P – R computations.

The performance of proposed CBIR algorithms can be compared and shown superior to CBIR system covered in Chapter 2 having similar test conditions. As mentioned on Page No 27, various CBIR techniques have been proposed in [Kekre, 2010, 1] & [Kekre, 2010, 2] for different classes of SIMPLicity image database [Wang, 2001] [SIMPLicity, on line]. The maximum Recall obtained in all seven methods proposed is about 0.35 (quite low) for 100 retrieved images (relevant + irrelevant) in [Kekre, 2010, 1] & [Kekre, 2010, 2]. And, even for retrieving 2 images, the Precision is less than 0.2 (very low) and drops further sharply with increase in no of retrieved images. *The performance measures for image retrieval with proposed techniques presented in the Chapter are significantly better compared to those of [Kekre, 2010, 1] & [Kekre, 2010, 2].*

As mentioned on Page No 26 & 27, Paitakes et al. [Pratikakis, 2006] proposed a novel unsupervised method for image retrieval based on hierarchical watershed algorithm and presented P - R curves. The mean precision-recall have been measured in [Pratikakis, 2006] for 10 queries per image-class of image database consisting of total 1000 images of 10 different classes with 100 images per class, reading (approximate values) for all categories of images, highest mean Precision of 0.7 at mean Recall of 0.07 and highest mean Recall of 0.425 with mean Precision of 0.41. Further, P – R curves corresponding to all image categories indicate that it is not possible to retrieve images with precision as 1 at any cost of recall, i.e. for no case, only relevant images get retrieved. *The performance measures with proposed techniques of image retrieval are far better than given in [Pratikakis, 2006].*

The foreground extraction and similar face image retrieval applied on uncropped images of Caltech [Caltech, on line] [Fei-Fei, 2004] has not been known reported in the literature.

The ideal P – R curve is characterized by highest and constant Precision value 1 for all Recalls, as shown in Figure 11 on Page 43. But, as described in Section 3.3.3, the practical P – R curves are characterized by fall in Precision for increase in the Recall. Improvement in one measure compromises the other for any Practical CBIR system. In other words, P – R curves of a practical CBIR system indicates how Precision degrades

with increase in Recall. Hence, a CBIR system giving high Precision with high Recall is considered to be a good system in terms of the P – R performance measures.

The exhaustive CBIR results, analysis and inferences of these performance measures for quantitative analysis of proposed methods based on (i) Color codes of entire image (ii) Foreground color codes (iii) Foreground shape correlation (iv) Combination of foreground color codes and shape correlation with selectable percentage proportion of weight of foreground color codes and foreground shape correlation for composite similarity measure (v) Similar face – images containing complex background have been carried out in various sections of the Chapter. The performance of the proposed methods can be judged with following absolute P – R measures:

- The fall in Precision is low for increase in Recall
- The average Precision at 0.5 giving good Recall
- Precision of 1 is attainable for some methods with some queries

The option of selecting the method of image retrieval with broader color descriptors, foreground shape and foreground color descriptor with selectable weights, maps the need and perception of a user giving good absolute performance measures of Precision & Recall.

6.13 Concluding Remark

A user has been offered to opt for the method of retrieval to map needs & choice for Retrieval with good Precision & Recall ...

7. Conclusions & Future Enhancements

7.1 Conclusions

A challenging task of development, implementation and integration of various novel algorithms to result into GUI based, selectable multi-modal processing of selectable single query image for retrieval of similar images has been achieved successfully.

These algorithms include:

- Edges and prominent boundaries detection
- Foreground separation
- Image retrieval based on
 - Color codes of entire image
 - Foreground color codes
 - Foreground shape correlation
 - Combination of foreground color codes and shape correlation
 - With selectable percentage proportion of weight of foreground color codes and foreground shape correlation for composite similarity measure
 - Extracted face region from images containing complex background for similar face-images retrieval

Other points of conclusions are:

- Well localized and well delineated thinned edges detected with proposed method outperform Adobe Photoshop, ACD Editor and MS Editor for i) detection of significant perceptual edges ii) elimination of insignificant edges corresponding background and foreground textures and iii) better preservation of continuity. These edges can be utilized for further reduction of artifacts produced due to watershed algorithm.
- Prominent boundaries, prominence measure, watershed algorithm with various levels of Haar wavelet decompositions are effectively incorporated together for proper segmentation and feature extraction by enforcing reliable processing of low level cues for avoiding breaks as well as under segmentation

by utilizing continuity preserving, well localized visually prominent boundaries for foreground – background separation. The problem of over segmentation is addressed by compositely considering proximity influence and watershed algorithm.

- High performance measures - $Precision_{fg}$ and $Recall_{fg}$, computed with respect to Ground Truth foreground for extracted foreground with proposed method endorses the uniqueness & effectiveness of the method & produced results.
- Prominent boundaries based approach for foreground-background separation is largely insensitive to illumination variations and inter region texture variations.
- *"Relaxed feature description for better Recall"* – the first half of the theme for the approaches has been accomplished with color codes and associated similarity measure for image retrieval.
- *"Simultaneous emphasizing of reliable processing of cues leading to precise feature extraction for better precision"* - the second half of the theme of approached has been accomplished with R,G,B and Gray channel processing for prominent boundaries and subsequent feature extraction.
- Image retrieval with similarity comparison of broader color descriptors - color codes for whole image is robust to illumination, pose and view point variations provides user an option of searching of similar images based on color distribution of whole image.
 - The method possesses conventional shortcomings due to ignorance of shapes and regional features.
 - The method is computationally efficient for feature extraction and similarity comparison.
 - The method provides high Recall and high ranking for nearly similar images.
- The retrieval technique based on foreground color code exploits reliable processing to precisely detect foreground and exclude background & related features to provide foreground color distribution based comparison of foreground objects.
 - The combination of stringent feature for prominent boundaries & foreground detection with broader color descriptor provides remarkable results for wide variety of images.
- Image retrieval with foreground shape correlation is very sensitive to the detected shape of the foreground. Attached shape altering regions adversely

affect the performance of the method. The method may not perform equally well on natural images.

- Foreground based methods are not suitable for images containing small foreground objects and objects spreading across image boundaries.
- The composite similarity measure constituted with selectable proportion of weights of foreground color codes and foreground shape correlation maps users need for proportionately emphasizing color code distribution and shape of foreground for retrieval of images. The stringent shape feature & broader color code features of foreground and their selectable proportion provide excellent combination for foreground object based image search with high Precision.
- High success ratio of 82% for precise face region extraction from 115 images of standard database Caltech [Caltech, on line] [Fei-Fei, 2004], which possess complex background and various illumination conditions infers the effectiveness of various proposed methods.
- The effectiveness & preciseness of various proposed algorithms have been exploited for development of application specific CBIR for similar face image retrieval on Caltech [Caltech, on line] [Fei-Fei, 2004] image database.
- All algorithms have been tested on variety of images, mainly of standard image databases, viz. BSDb [Fowlkes, on line] [Martin, 2001], SIMPLcity [Wang, 2001] [SIMPLcity, on line], PASCAL challenge 2008 [Everingham, on line], ALOI [ALOI, on line] [Geusebroek, 2001], Caltech [Caltech, on line] [Fei-Fei, 2004], MedPics [MedPics, on line] and University of Washington [University of Washington, on line], establishing suitability of methods for wide varieties and categories of images.
- Obtained P – R curves for query responses with proposed methods of retrieval are close to ideal for many cases and similar to practical P – R curves for the rest proving effectiveness of algorithms and validity of results.
- No method of image retrieval can be generalized to be concluded outperforming others for all queries because of
 - diversities in image characteristics and categories
 - subjectivity in image content description
 - semantic gap

7.2 Limitations

The limitations of the developed CBIR system are as under:

- The developed algorithms and application are resource hungry. Hence, processing of high resolution large sized images is possible only with machines possessing huge memory and computational resources.
- The prominent boundaries detection does not yield expected results for very low resolution images, particularly captured from web-cam.
- Lack of incorporation of indexing technique has caused higher query response time.

7.3 Future Enhancements

Suggested future enhancements are as under.

- Analysis of performance of prominent boundaries detection method with other wavelets.
- Utilization of well localized thin-edges to further reduce artifacts produced due to intrinsic characteristic of watershed algorithm.
- Extraction of texture feature and its incorporation in the CBIR algorithms.
- Incorporation of indexing technique(s) for faster query response.
- Incorporation of database management modules for image and image-feature databases.
- Incorporation of relevance feed-back from user to increase the retrieval performance of the system.
- Incorporation of multiple-queries to refine results for improved retrieval performance.
- Extending the query processing support to all file types.
- Testing of the CBIR algorithms for Precision and Recall performance measures - on very large image databases, of the order of tens of thousands of images. (This testing requires high-end machine(s) to meet computational needs of feature extraction and similarity comparison methods.)
- Testing of CBIR algorithms for databases consisting of transformed images with scaling, cropping, contrast alterations, brightness changes, changes in quality factor of JPEG compression and noise introduction.
- Testing of the CBIR algorithms for precision and recall performance on medical images – particularly X-ray images. The foreground shape based CBIR method should yield very encouraging results because of following reasons:

- The foreground detection should be well localized and precise due to X-ray image characteristics.
- The X-ray images are captured by skilled persons in a controlled illumination conditions eliminating associated challenges.
- The object shape variations due to affine transformations are limited.
- Exploiting regions and region attributes for development of CBIR application for histological and pathological images.
- Inclusion of face-geometric features for face detection & extraction for improvement in face-extraction success ratio.
- Enhancing the application for image annotation.
- Development of face recognition and face expression understanding algorithm/application based on proposed face extraction method.
- Development of application for video abstraction.
- Enhancing the developed application to web-based CBIR system.

7.4 Concluding Remark

A step forward on the road-map of continuous & endless technical evolution towards a perfect & versatile CBIR ...

References

- [Agaian, 2010]** S. S. Agaian, K. A. Panetta, S. C. Nercessian and E. E. Danahy, "Boolean Derivatives With Application to Edge Detection for Imaging Systems," in *IEEE Transactions on Systems, Man, And Cybernetics—Part B: Cybernetics*, Vol. 40, No. 2, April 2010.
- [aliper, on line]** The aliper website:
<http://alipr.com/>
- [ALOI, on line]** Amsterdam Library of Object Images (ALOI). Available:
<http://staff.science.uva.nl/~aloi/>
- [Aptoula, 2009]** E. Aptoula and S. Lefèvre, "Morphological Description of Color Images for Content-Based Image Retrieval," in *IEEE Transactions on Image Processing*, Vol. 18, No. 11, November 2009, pp. 2505-2517.
- [Arbel'aez, 2006]** P. Arbel'aez, "Boundary Extraction in Natural Images Using Ultrametric Contour Maps," in *Proceedings of the 2006 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'06)*, 2006, pp. 182-182.
- [Arbel'aez, 2009]** P. Arbel'aez, M. Maire, C. Fowlkes and J. Malik, "From Contours to Regions: An Empirical Evaluation," in *CVPR*, 2009, pp. 2294-2301.
- [Barber, 1994]** R. Barber, M. Flickner, J. Hafner, D. Lee, W. Niblack, D. Petkovic, J. Ashley, T. McConnell, J. Ho, J. Jang, D. Berkowitz, P. Yanker, M. Vo, D. Ilaas, D. Lassig, S. Tate, A. Chang, P. van Houten, J. Chang, T. Petersen, D. Lutrell, M. Snedden, P. Faust, C. Matteucci, M. Rayner, R. Peters, W. Beck, J. Witsett, "Ultimedia Manager: Query By Image Content And Its Applications," *Digest of Papers, Compcon Spring '94*, 1994, pp. 424-429.
- [Basak, 2006]** J. Basak, K. Bhattacharya and S. Chaudhury, "Multiple Exemplar-Based Facial Image Retrieval Using Independent Component Analysis," in *IEEE Transactions on Image Processing*, Vol. 15, No. 12, December 2006, pp. 3773-3783.
- [Beucher, 1979]** S. Beucher and C. Lantuejoul, "Use of Watersheds in Contour Detection," in *Proceedings of International Workshop Image Processing Real-Time Edge and Motion Detection/Estimation*, 1979.

[Bian, 2010] W. Bian and D. Tao, "Biased Discriminant Euclidean Embedding for Content-Based Image Retrieval," in *IEEE Transactions on Image Processing*, Vol. 19, No. 2, February 2010, pp. 545-554.

[Borenstein, 2008] E. Borenstein and S. Ullman, "Combined Top-Down/Bottom-Up Segmentation," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 30, No 12, Dec 2008, pp.2109-2125.

[Caltech, on line] Caltech 101 face dataset. Available:

http://www.vision.caltech.edu/Image_Datasets/Caltech101/Caltech101.html

[Carson, 2002] C. Carson, S. Belongie, H. Greenspan and J. Malik, "Blobworld: image segmentation using E-M and its application to image querying," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, 24, pp. 1026–1038.

[Cheng, 2007] S. Cheng, W. Huang, Y. Liao and D. Wu, "A Parallel CBIR Implementation Using Perceptual Grouping Of Block-based Visual Patterns," in *IEEE International Conference on Image Processing – ICIP*, 2007, pp. V - 161 - V - 164.

[Chun, 2003] Y. D. Chun, S. Y. Seo and N. C. Kim, "Image Retrieval Using BDIP and BVLC Moments," in *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 13, No. 9, September 2003, pp. 951-957.

[Comaniciu, 2002] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24, No 5, May 2002, pp. 603-619.

[Datta, 2008] R. Datta, D. Joshi, J. Li and J. Z. Wang, "Image Retrieval: Ideas, Influences, and Trends of the New Age," in *ACM Computing Surveys*, vol. 40, no. 2, article 5, 2008, pp. 1-60.

[Deng, on line] Y. Deng and B.S. Manjunath, JSEG software site, available at:

<http://vision.ece.ucsb.edu/segmentation/jseg/software/>

[Deng, 2001] Y. Deng and B. S. Manjunath, "Unsupervised segmentation of color-texture regions in images and video," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2001, 23, (8), pp. 800–810.

[Dutta, 2009] S. Dutta and B. B. Chaudhuri, "A Color Edge Detection Algorithm in RGB Color Space," in *International Conference on Advances in Recent Technologies in Communication and Computing*, 2009, pp. 337-340.

[Everingham, on line] M. Everingham, L. VanGool, C. K. I. Williams, J. Winn and A. Zisserman, PASCAL challenge 2008 image database.

<http://pascallin.ecs.soton.ac.uk/challenges/VOC/voc2008/workshop/index.html>

- [Fei-Fei, 2004]** L. Fei-Fei, R. Fergus and P. Perona, "Learning generative visual models from few training examples: an incremental Bayesian approach tested on 101 object categories," in *IEEE CVPR 2004, Workshop on Generative-Model Based Vision*, 2004.
- [Felzenszwalb, 2004]** P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient graph based image segmentation," in *International Journal of Computer Vision*, 2004, 59,(2), pp. 167–181.
- [Ferrari, 2008]** V. Ferrari, L. Fevrier, F. Jurie, and C. Schmid, "Groups of Adjacent Contour Segments for Object Detection," in *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Volume 30 , Issue 1, Jan. 2008, pp. 36-51.
- [Flickner, 1995]** M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele and P. Yanker, "Query by image and video content: the QBIC system," in *Computer*, Volume: 28, Issue: 9, 1995, pp. 23-32
- [Fire, on line]** The Fire web site:
http://www-i6.informatik.rwth-aachen.de/~deselaers/cgi_bin/fire.cgi
- [Fowlkes, on line]** C. Fowlkes, D. Martin and J. Malik. The Berkeley Segmentation Dataset and Benchmark (BSDb). Available:
<http://www.cs.berkeley.edu/projects/cs/vision/grouping/segbench/>
- [Fukuda, 2008]** K. Fukuda, T. Takiguchi and Y. Ariki, "Graph Cuts by Using Local Texture Features of Wavelet Coefficient for Image Segmentation," in *IEEE International Conference on Multimedia and Expo*, 2008, pp. 881 – 884.
- [Gavrielides, 2006]** M. A. Gavrielides, E. Sikudová, and I. Pitas, "Color-Based Descriptors for Image Fingerprinting," in *IEEE Transactions On Multimedia*, Vol. 8, No. 4, August 2006, pp. 740-748.
- [GazoPa, on line]** The GazoPa website:
www.gazopa.com/
- [Geusebroek, 2001]** J. M. Geusebroek, G. J. Burghouts, and A. W. M. Smeulders, "The Amsterdam library of object images," in *International Journal of Computer Vision*, 61(1), 103-112, January, 2005 .
- [Gudivada, 1995]** V. N. Gudivada and V. V. Raghavan, "Content-based image retrieval systems," in *IEEE Computer* 28(9), 1995, pp. 18-22.
- [Ha, on line]** R. Ha, and J. Romberg, Haar Wavelet Basis. Available:
<http://cnx.org/content/m10764/latest/>
- [Hadjidemetriou, 2004]** E. Hadjidemetriou, M. D. Grossberg, and S.K. Nayar, "Multi-resolution histograms and their use for recognition," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 26 Issue: 7, July 2004, pp. 831 – 847.

- [Hanjalic, 2008]** A. Hanjalic, R. Lienhart, W. Ma and J. R. Smith, "The Holy Grail of Multimedia Information Retrieval: So Close or Yet So Far Away?," in *Proceedings of the IEEE* Vol. 96, No. 4, April 2008, pp. 541-547.
- [Hsu, 2002]** R. Hsu, M. Abdel-Mottaleb and A. K. Jain, "Face Detection in Color Images," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24, No 5 , May 2002.
- [Hsu, 2005]** C. Hsu and C. Li, "Relevance Feedback Using Generalized Bayesian Framework With Region-Based Optimization Learning," in *IEEE Transactions on Image Processing*, Vol. 14, No. 10, October 2005, pp. 1617- 1631.
- [Huang, 1997]** J. Huang, S. R. Kumar, M. Mitra, W. Zhu and R. Zabih, "Image Indexing Using Color Correlograms," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1997, pp. 762 – 768.
- [Kass, 1988]** M. Kass, A. Witkin and D. Terzopoulos, "Snakes: Active Contour Models," in *International Journal of Computer Vision*, 1988, pp. 321-331.
- [Kekre, 2010, 1]** Dr. H. B. Kekre, S. Thepade, P. Mukherjee, M. Kakaiya, S. Wadhwa, and S. Singh, "Image Retrieval with Shape Features Extracted using Gradient Operators and Slope Magnitude Technique with BTC," in *International Journal of Computer Applications (0975 – 8887) Volume 6–No.8, September 2010*.
- [Kekre, 2010, 2]** Dr. H. B. Kekre, S. Thepade, P. Mukherjee, M. Kakaiya, S. Wadhwa and S. Singh, "Image Retrieval with Shape Features Extracted using Morphological Operators with BTC," in *International Journal of Computer Applications (0975 – 8887) Volume 12–No.3, November 2010*, pp. 1-5.
- [Krishnapuram, 2004]** R. Krishnapuram, S. Medasani, S. Jung, Y. Choi, and R. Balasubramaniam, "Content-Based Image Retrieval Based on a Fuzzy Approach," in *IEEE Transactions on Knowledge and Data Engineering*, Vol. 16, No. 10, October 2004, pp. 1185-1199.
- [Laaksonen, 1999]** J. Laaksonen, M. Koskela and E. Oja, "PicSOM: Self-Organizing Maps for Content-Based Image Retrieval," in *International Joint Conference on Neural Networks, IJCNN '99*, Volume 4, July 1999, pp. 2470 – 2473.
- [Laaksonen, 2002]** J. Laaksonen, M. Koskela and E. Oja, "PicSOM Self Organizing Image Retrieval With MPEG-7 Content Descriptors," in *IEEE Transactions on Neural Networks*, Vol. 13, No. 4, July 2002, pp. 841-853.
- [Lew, 2006]** M. S. Lew, N. Sebe, C. Djeraba and R. Jain, "Content-based Multimedia Information Retrieval: State of the Art and Challenges," in *ACM Transactions on Multimedia Computing, Communications, and Applications*, Feb. 2006.

- [Li, 2000] J. Li, J. Z. Wang and G. Wiederhold, "IRM: Integrated Region Matching for Image Retrieval," in *Proceeding MULTIMEDIA '00 Proceedings of the eighth ACM international conference on Multimedia*.
- [Li, 2001] Y. Li, T. Zhang and D. Tretter, "An Overview of Video Abstraction Techniques," in *Tech. Rep. Hp-2001-191, HP Laboratory*, July 2001.
- [Li, 2008] F. Li, Q. Dai, W. Xu and G. Er, "Multilabel Neighborhood Propagation for Region-Based Image Retrieval," in *IEEE Transactions on Multimedia*, Vol. 10, No. 8, December 2008, pp. 1592-1604.
- [Ma, 1997] W.Y. Ma and B.S. Manjunath, "NeTra: A Toolbox For Navigating Large Image Databases," in *IEEE International Conference on Image Processing*, 1997, pp. 568-571.
- [Ma, 2000] W. Ma and B. S. Manjunath, "EdgeFlow: A Technique for Boundary Detection and Image Segmentation," in *IEEE transaction on Image Processing*, Vol. 9, Issue 8, August 2000, pp. 1375-1388.
- [Maire, 2008] M. Maire, P. Arbel'aez, C. Fowlkes and J. Malik, "Using Contours to Detect and Localize Junctions in Natural Images," in *Computer Vision and Pattern Recognition*, 2008, pp. 1-8.
- [Malik, 2001] J. Malik, S. Belongie, T. Leung and J. Shi, "Contour and Texture Analysis for Image Segmentation," in *International Journal of Computer Vision*, Vol. 43, Issue 1, 2001, pp. 7-27.
- [Mallat, 1989] S. G. Mallat, "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 11, No. 7, July 1989, pp. 674 – 693.
- [Martin, 2001] D. Martin, C. Fowlkes, D. Tal and J. Malik, "A Database of Human Segmented Natural Images and its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics," in *ICCV*, 2001.
- [Martin, 2004] D. Martin, C. Fowlkes and J. Malik, "Learning to detect natural image boundaries using local brightness, color and texture cues," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(5), 2004, pp. 530–549.
- [MARS on line] The MARS web site:
<http://www.ifp.illinois.edu/~qitian/MARS.html>
- [Medina-Carnicer, 2010] R. Medina-Carnicer, A. Carmona-Poyato, R. Muñoz-Salinas and F. J. Madrid-Cuevas, "Determining Hysteresis Thresholds for Edge Detection by Combining the Advantages and Disadvantages of Thresholding Methods," in *IEEE Transactions on Image Processing*, Vol. 19, No. 1, January 2010, pp. 165-173.

- [MedPics, on line]** MedPics – An Image Library for Medical Education, UCSD – School of Medicine. Available: <http://medpics.ucsd.edu/>
- [Müller, 2001]** H. Müller, W. Müller, D. M. Squire, S. Marchand-Maillet, T. Pun, "Performance Evaluation in Content-Based Image Retrieval: Overview and Proposals" in *Pattern Recognition Letters (Special Issue on Image and Video Indexing)*, 2001; 22(5):593–601. H. Bunke and X. Jiang Eds.
- [Nason, 1995]** G.P. Nason and B.W. Silverman, "The stationary wavelet transform and some statistical applications," in *Lecture Notes in Statistics*, 103, 1995, pp. 281-299.
- [O'Connor, 2006]** N. E. O'Connor and I. Kompatsiaris, "Editorial: Recent Advances in Image and Video Retrieval," in *IEE Proceedings*, December 2006, pp. 851-851.
- [Pass, 1996]** G. Pass and R. Zabih, "Histogram Refinement for Content Based Image Retrieval," in *3rd IEEE Workshop on Applications of Computer Vision, WACV*, 1996, pp. 96-102.
- [Pawan Kumar, 2010]** M. Pawan Kumar, P.H.S. Torr and A. Zisserman, "OBJCUT: Efficient Segmentation Using Top-Down and Bottom-Up Cues," in *IEEE Transactions on Pattern Analysis And Machine Intelligence*, Vol. 32, No. 3, March 2010.
- [Petrakis, on line]** The Intelligent Systems Laboratory, Department of Electronic & Computer Engineering, Technical University of Crete, Greece
<http://www.intelligence.tuc.gr/~petrakis/courses/multimedia/retrieval.pdf>
- [Pavlidis, on line]** T. Pavlidis, "Limitations of Content-based Image Retrieval", invited plenary talk at the *19th International Conference on Pattern Recognition*, Tampa, Florida, Dec. 8-11, 2008. Slides available:
<http://www.theopavlidis.com/technology/CBIR/PaperB/icpr08.htm>.
- [Photobook, on line]** The Photobook web site:
<http://vismod.media.mit.edu/vismod/demos/photobook/ph6>
- [Phung, 2005]** S. L. Phung, A. Bouzerdoum and D. ChaiSkin, "Segmentation Using Color Pixel Classification: Analysis and Comparison," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 27, No. 1, January 2005.
- [Prasad, 2004]** B.G.Prasad, K.K.Biswas and S.K.Gupta, "Image Retrieval Using Integrated Color-Shape-Location Index in Special Issue: Colour for Image Indexing and Retrieval," in *Computer Vision and Image Understanding Volume 94*, Issues 1-3, April-June 2004, pp. 193-233.
- [Pratikakis, 2006]** I. Pratikakis, I. Vanhamel, H. Sahli, B. Gatos and S.J. Perantonis, "Unsupervised watershed-driven region-based image retrieval," in *IEE Proceedings on Vision, Image and Signal Processing*, Vol. 153, No. 3, June 2006 pp. 313-322.

- [QBIC, on line]** The QBIC web site: <http://www.qbic.almaden.ibm.com/>
- [Ren, 2008]** X. Ren, C. C. Fowlkes and J. Malik, "Learning Probabilistic Models for Contour Completion in Natural Images," in *International Journal of Computer Vision*, 2008, 77, pp. 47–63.
- [Rao, 1999]** A. Rao, R. K. Srihari and Z. Zhang, "Spatial Color Histograms for Content-Based Image Retrieval," *11th IEEE International Conference on Tools with Artificial Intelligence*, 1999, pp. 183 – 186.
- [Rasiwasia, 2007]** N. Rasiwasia, P. J. Moreno and N. Vasconcelos, "Bridging the Gap: Query by Semantic Example," in *IEEE Transactions on Multimedia*, Vol. 9, No. 5, August 2007, pp. 923-938.
- [Rasiwasia, 2008]** N. Rasiwasia and N. Vasconcelos, "A Study of Query by Semantic Example," in *3rd International Workshop on Semantic Learning and Applications in Multimedia*, Anchorage, June 2008, pp. 1-8.
- [Rui, 1998]** Y. Rui, T. S. Huang, M. Ortega and S. Mehrotra, "Relevance Feedback: A Power Tool for Interactive Content-Based Image Retrieval," in *IEEE Transactions on Circuits and Systems For Video Technology*, Vol. 8, No. 5, September 1998, pp. 644 – 655.
- [Saadatmand-Tarzjan, 2007]** M. Saadatmand-Tarzjan and H. A. Moghaddam, "A Novel Evolutionary Approach for Optimizing Content-Based Image Indexing Algorithms," in *IEEE Transactions on Systems, Man and Cybernetics—Part B: Cybernetics*, Vol. 37, No. 1, February 2007, pp. 139-153.
- [Schettini, 2001]** R. Schettini, G. Ciocca and S. Zuffi, "A Survey of Methods for Colour Image Indexing and Retrieval in Image Databases, in *Color Imaging Science: Exploiting Digital Media*," R. Luo and L. MacDonald, Eds., J. Wiley, New York ~2001.
- [Shi, 2000]** J. Shi and J. Malik, "Normalized Cuts and Image Segmentation," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, Aug. 2000, pp. 888-905.
- [Shih, 2005]** F. Y. Shih and S. Cheng, "Automatic seeded region growing for color image segmentation," in *Image and Vision Computing*, 23 (2005), 877–886.
- [SIMPLcity, on line]** The SIMPLcity web site:
http://wang14.ist.psu.edu/cgi-bin/zwang/regionsearch_show.cgi
- [Smith, 1996]** J. R. Smith and S. F. Chang, "VisualSEEK: A Fully Automated Content-Based Image Query System", in *Proceedings of ACM International Conference on Multimedia*, Boston, MA, 1996, pp. 87–98.

- [Smeulders, 2000]** A. W. Smeulders, M. M. Worring, A. Gupta and R. Jain, "Content-Based Image Retrieval at the End of the Early Years," in *IEEE Transaction on Pattern Analysis Machine Intelligence*, Vol. 22 No.12, 2000, pp. 1349-1380.
- [Stricker, 1995]** M. A. Stricker and M. Orengo, "Similarity of Color Images," in *Proceedings of the SPIE conference on the Storage and Retrieval for Image and Video Databases III*, 1995, pp. 381-392.
- [Swain, 1991]** M. Swain and D. Ballard, "Color Indexing," in *International Journal of Computer Vision*, 7(1), 1991, pp. 11-32.
- [Tao, 2008]** D. Tao, X. Tang and X. Li, "Which Components are Important for Interactive Image Searching?," in *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 18, No. 1, January 2008, pp. 3-11.
- [tiltomo, on line]** The tiltomo website:
<http://www.tiltomo.com/>
- [Thakore, 2010, 1]** D. G. Thakore and A. I. Trivedi, "Content Based Image Retrieval Techniques – Issues, Analysis and the State Of The Art," in *Proceedings of International Symposium on Computer Engineering & Technology – ISCET 2010*, March 2010, pp. 98-102.
- [Thakore, 2010, 2]** D. G. Thakore and A. I. Trivedi, "Edge Detection from Candidate Boundaries and Qualitative Comparison of Results for Color Images," in *International Conference on Emerging Trends on Engineering and Technology (ICETET-2010)*, November 2010, pp. 229-234.
- [Thakore, 2010, 3]** D. G. Thakore and A. I. Trivedi, "Prominent Boundaries Detection Technique for Color Images," in *Proceedings of Second International Conference on Computing Communication and Networking Technologies (ICCCNT-2010)*, July 2010, pp. 1-5.
- [Thakore, 2010, 4]** D. G. Thakore and A. I. Trivedi, "Foreground Objects Revealing for Prominent Boundaries Detected Color Images," in *Proceedings of Second International Conference on Multimedia and Content Based Image Retrieval (ICMCBIR-2010)*, July 2010, pp. 121-128.
- [Toshev, 2010]** A. Toshev, B. Taskar and K. Daniilidis, "Object Detection via Boundary Structure Segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2010, pp. ---.
- [Ugarriza, 2009]** L. G. Ugarriza, E. Saber, S. R. Vantaram, V. Amuso, M. Shaw and R. Bhaskar, "Automatic Image Segmentation by Dynamic Region Growth and

Multiresolution Merging," in *IEEE Transactions on Image Processing*, Vol. 18, No. 10, October 2009.

[University of Washington, on line] University of Washington, Image database. Available: <http://www.cs.washington.edu/research/imagedatabase/demo/seg/>

[Vasconcelos, 2007] N. Vasconcelos, "From Pixels to Semantic Spaces: Advances in Content-Based Image Retrieval," in *Computer Volume: 40, Issue: 7, 2007*, pp. 20-26.

[Veltkamp, 2000] R. C. Veltkamp and M. Hagendoorn, "State-of-the-Art in Shape Matching," in *Multimedia Search: State of the Art*, Springer-Verlag, 2000.

[Veltkamp, on line] R. C. Veltkamp and M. Tanase, "Content-Based Image Retrieval Systems: A Survey," Available: <http://aa-lab.cs.uu.nl/cbirsurvey/cbir-survey/>

[Virage, on line] The Virage web site: <http://www.virage.com>

[VisualSEEk, on line] The VisualSEEk web site: <http://www.ee.columbia.edu/ln/dvmm/researchProjects/MultimediaIndexing/VisualSEEk/VisualSEEk.htm>

[Vincent, 1991] L. Vincent and P. Soille, "Watersheds in Digital Spaces: An Efficient Algorithm Based on Immersion Simulations," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 13, No. 6, June 1991, pp. 583-598.

[Wang, 2001] J. Z. Wang, J. Li and G. Wiederhold, "SIMPLicity: Semantics-Sensitive Integrated Matching for Picture Libraries," in *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Volume 23, No 9, Sept 2001, pp. 947-963.

[Xu, 2010] J. Xu, A. Janowczyk, S. Chandran and A. Madabhushi, "A Weighted Mean Shift, Normalized Cuts Initialized Color Gradient Based Geodesic Active Contour Model: Applications to Histopathology Image Segmentation," in *Proceedings of SPIE Medical Imaging 2010: Image Processing*, Vol. 7623, 76230Y.

[Yang, 2002] M. Yang, D. J. Kriegman and N. Ahuja, "Detecting Faces in Images: A Survey," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24, No. 1, January 2002.

Annexure 1 - Publications

- D. G. Thakore and A. I. Trivedi, "Content Based Image Retrieval Techniques – Issues, Analysis and the State Of The Art," in *Proceedings of International Symposium on Computer Engineering & Technology – ISCET 2010*, March 2010, pp. 98-102.

Accessible at:

<http://www.rimtengg.com/iscet/proceedings/pdfs/image%20proc/47.pdf>

- D. G. Thakore and A. I. Trivedi, "Prominent Boundaries Detection Technique for Color Images," in *Proceedings of Second International Conference on Computing Communication and Networking Technologies (ICCCNT-2010)*, July 2010, pp. 1-5.

Accessible at:

http://www.ieeeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5592598

- D. G. Thakore and A. I. Trivedi, "Foreground Objects Revealing for Prominent Boundaries Detected Color Images," in *Proceedings of Second International Conference on Multimedia and Content Based Image Retrieval (ICMCBIR-2010)*, July 2010, pp. 121-128.
- Darshak G. Thakore and A. I. Trivedi, "Edge Detection from Candidate Boundaries and Qualitative Comparison of Results for Color Images," in *Proceedings of Third International Conference on Emerging Trends on Engineering and Technology, (ICETET-2010)*, November 2010, pp. 229-234.

Accessible at:

<http://doi.ieeecomputersociety.org/10.1109/ICETET.2010.119>

& at: <http://www.ieeeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5698325>

- Darshak G. Thakore and A. I. Trivedi, "Prominent Boundaries and Foreground Detection Based Technique for Human Face Extraction from Color Images Containing Complex Background," in *Proceedings of Third National Conference on Computer Vision, Pattern Recognition (NCVPRIPG-2011)*, Image Processing and Graphics, Hubli, Karnataka, India, December 2011, pp. 15-20.

Accessible at:

<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6132990>

Annexure 2 - GUI & Description

A-2.1 Graphical User Interface

The Figure 102 shows Graphical User Interface (GUI) of developed CBIR application. The full-fledged GUI enables user to

- Select query image and target (search) folder
- Extract features of image(s)
- Specify input parameters
- Perform segmentation with two different methods
- Retrieve images with various algorithms

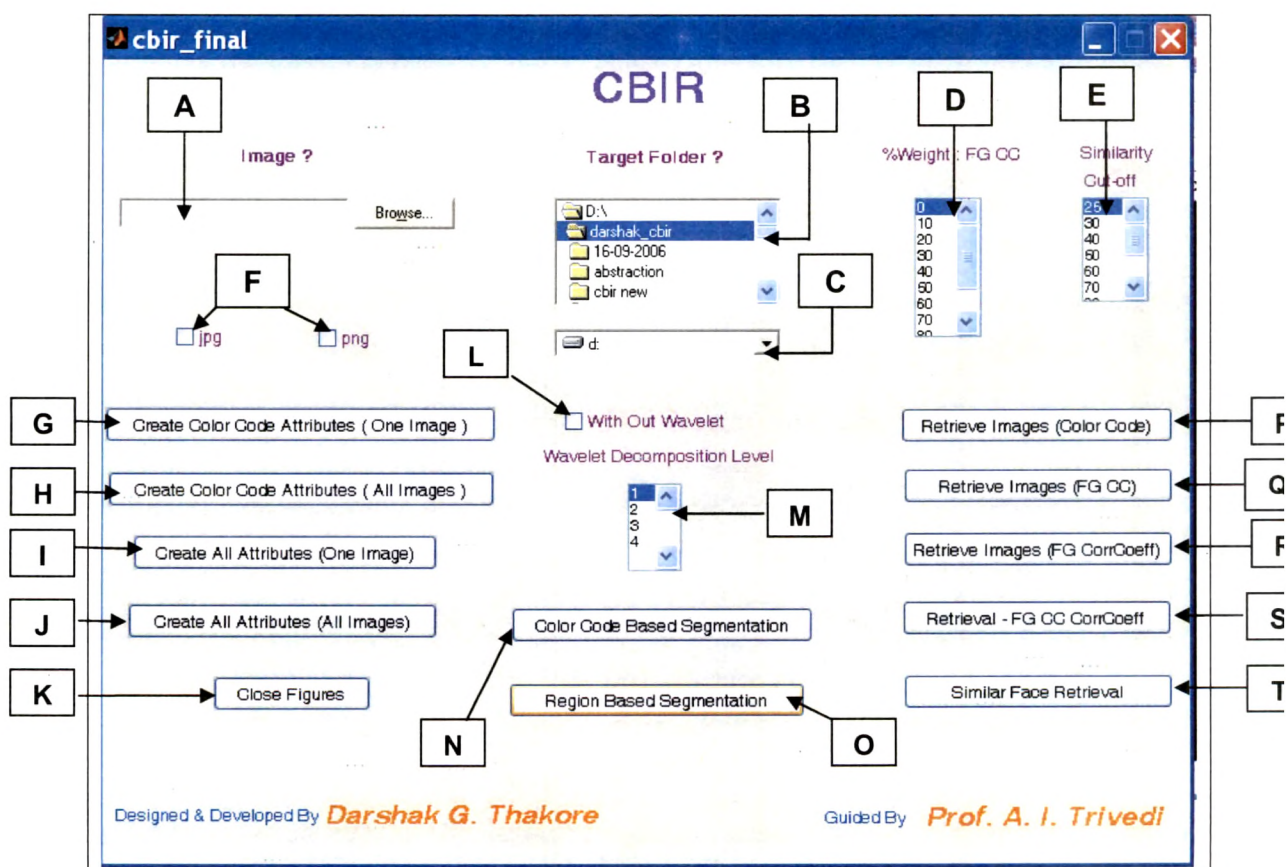


Figure 102. GUI of the CBIR

A-2.2 Component Description

The functionality of various components of GUI are describes below.

A. Image selection

The component enables user to select image by browsing files as shown in following Figure 103. The selected image will be processed for feature extraction or segmentation based on the task selected. The image will be treated as a query for image retrieval purpose.

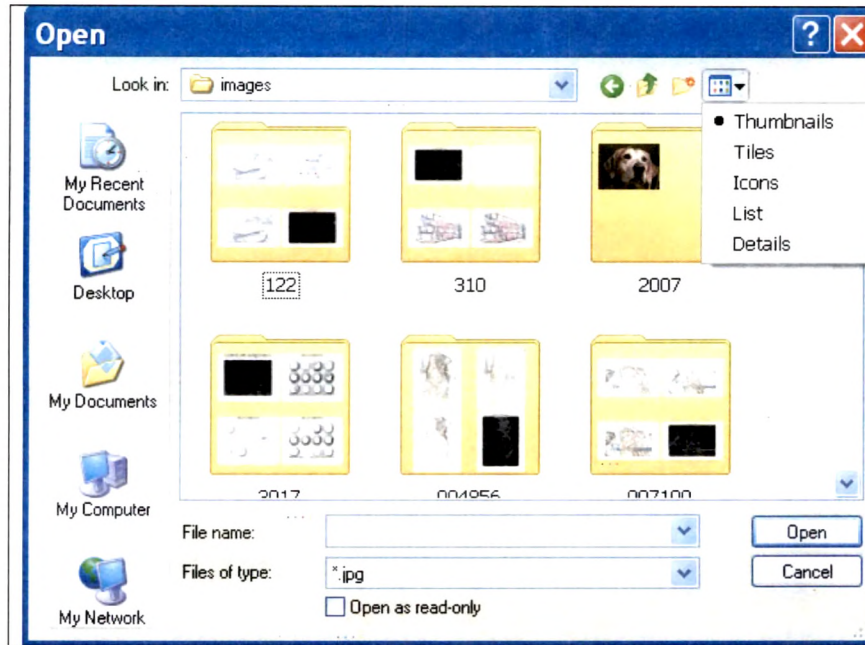


Figure 103. Image Selection by Browsing

B. Target folder selection

The folder selected specifies (i) feature extraction of all files of folder or (ii) target folder for image search.

C. Drive selection

The disk-drive selection for target folder is achieved with the GUI component.

D. Weight Selection for Foreground Color Codes (%)

The input parameter is used for specifying % proportion of foreground color code similarity in composite similarity measure.

E. Similarity Cut-off

The input parameter specifies cut-off of similarity measure for image retrieval.

F. File type selection

The jpg or png file selection is achieved with the checkboxes.

G. Image feature extraction - Color Codes, one image

The color code attributes are formed for an image specified at GUI component A.

H. Image feature extraction - Color Codes, all images

The color code features extraction is performed for all files of folder selected with GUI component marked as B. The confirmation dialogue box, as shown in Figure 106 is prompted before proceeding for a time expensive processing.

I. Extraction of all features – one image

All features of an image are extracted with the component.

J. Extraction of all features – all images

All features of all images of folder selected with GUI component marked with B are extracted with the component. The confirmation dialogue box, as shown in Figure 106 is prompted before proceeding for a time expensive processing.

K. Close all figures

The button is used to close all windows (figures) produced for output.

L. Exclude Wavelet decomposition

The checkbox excludes wavelet decomposition step when checked.

M. Wavelet decomposition level

The selection specifies wavelet decomposition level.

N. Color codes based segmentation

The button-click performs color codes based segmentation of selected image.

O. Region based segmentation

The button-click performs prominent boundaries detection based segmentation.

P. Color codes based image retrieval

The image retrieval based on similarity measures of color code attributes will be performed with the button-click.

Q. Foreground Color codes based image retrieval

The image retrieval based on similarity measures of foreground color code attributes will be performed with the button-click.

R. Foreground Correlation Coefficient based image retrieval

The image retrieval based on similarity measures of correlation coefficients of foreground will be performed with the click of the button.

S. Foreground Color Codes and Correlation Coefficient based image retrieval

The image retrieval based on composite similarity measures of foreground color codes and foreground correlation coefficients will be performed with the click of the button. The proportion of percentage weight is specified with component marked as D.

T. Similar face-image retrieval

The similar face-image retrieval is performed with the button-click.

Various Dialogue boxes

Figure 104 shows two dialogue boxes prompted when image for feature extraction or query image for similar image retrieval is not specified at component marked as A.

Figure 105 shows prompted dialogue box when image type selection is not made.

Figure 106 shows prompted dialogue box for confirming proceeding for time consuming process of extracting features of all files of specified folder.

Figure 107 shows prompted dialogue box when wrong selection of target folder is made.

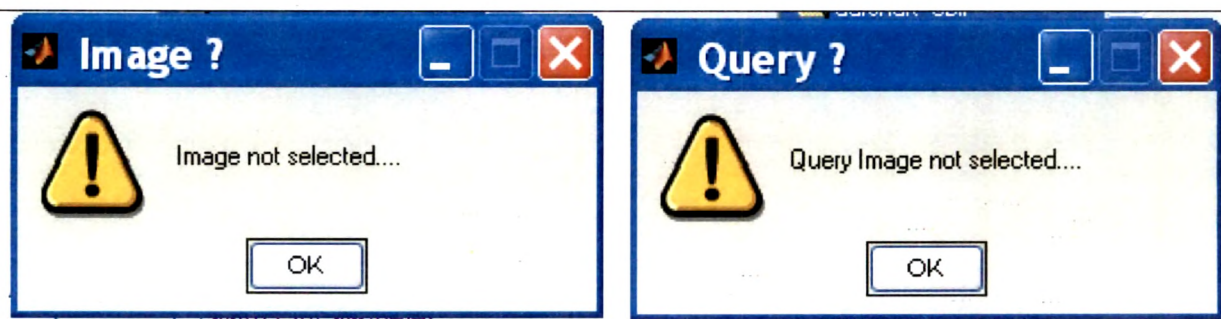


Figure 104. Dialogue box for unselected image



Figure 105. Dialogue box for unselected check box for image type

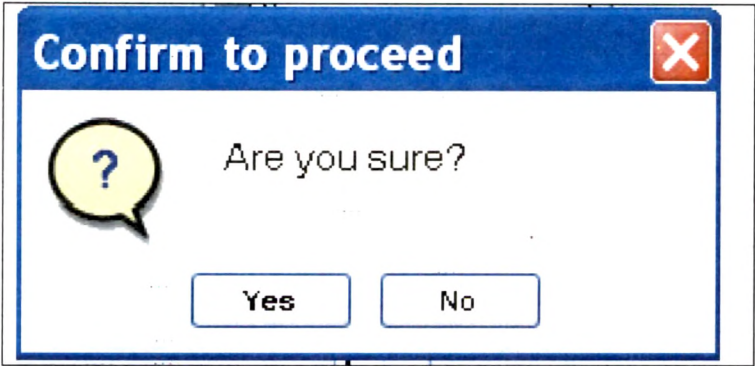


Figure 106. Dialogue box for confirming all image processing for all attributes



Figure 107. Dialogue box for wrongly selected target folder for image retrieval

Annexure 3 – Results: Miscellaneous

A-3.1 Suitability of Proposed Methods for Character Recognition

Following Figure 108 shows prominent boundaries, extracted foreground, background and resultant watershed pixels of a scanned image of a printed Gujarati alphabet. The shape preserving feature extraction is a major concern in character recognition, which is challenged by breaks present & induced because of processing of scanned image. Continuity and shape preserving watershed pixels of one pixel width produced with proposed algorithm provide the most suitable & required features for character recognition.

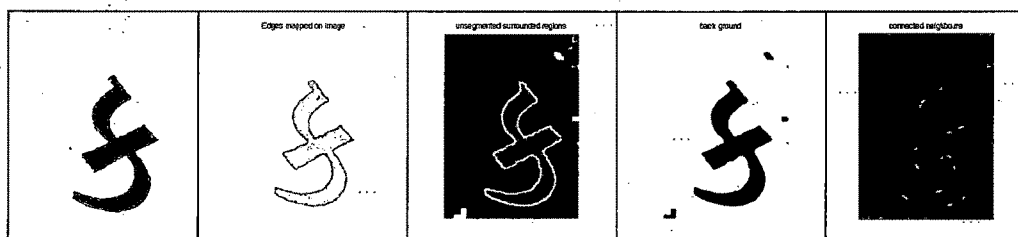


Figure 108. From Left to Right – Scanned alphabet, Mapped edges, Foreground, Background and Watershed pixels.

A-3.2 Foreground Extraction - Tiger Images of BSDB [Fowlkes, on line] [Martin, 2001]

Tiger images captured in natural conditions of forest provide all possible challenges for foreground extraction. The tiger images of BSDB [Fowlkes, on line] [Martin, 2001] are few of the toughest among all. The effectiveness of proposed methods leading to foreground extraction is shown with the results in Figure 108.

A-3.3 Query Response Examples – Tiger Images BSDB [Fowlkes, on line] [Martin, 2001]

The image database BSDB [Fowlkes, on line] [Martin, 2001]- a collection of challenging images for segmentation is not meant for image retrieval as many images belonging to same class are not available. Still, on available 6 tiger images of the database, experiments of image retrieval were carried out to target retrieval of tiger images.


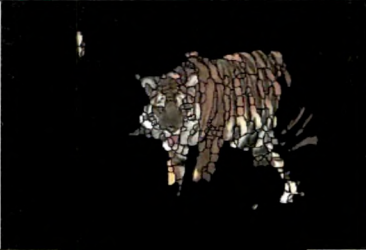

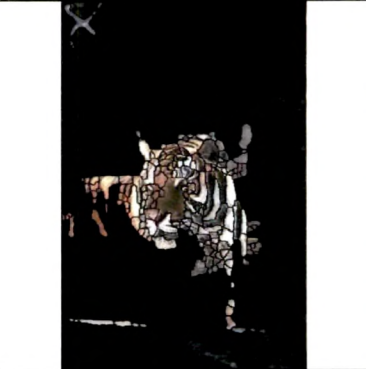

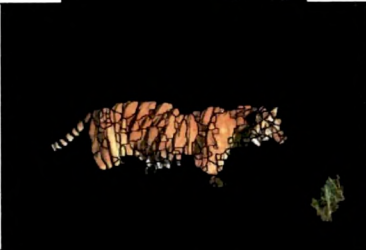





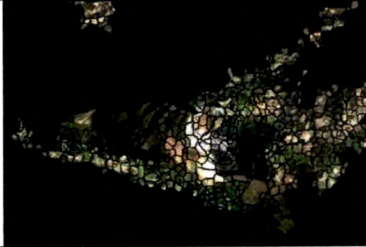
	Image	Extracted foreground
a)		
b)		
c)		
d)		
e)		
f)		

Figure 109. Tiger images of BSDB [Fowlkes, on line] [Martin, 2001] and extracted foreground.

Queries	Responses
	
	
	
	
	
	

Figure 110. Queries & responses of Tiger images BSDB [Fowlkes, on line] [Martin, 2001].

The foreground color codes based method was used for the purpose of retrieval with similarity cut-off set to 50. Figure 109 illustrates the effectiveness of the method for extracting foreground tiger for challenging complex background with different pose/illumination/texture variations. The screen shot of results for each of tiger image given as query are shown in Figure 110. Except for the last query image containing disguised tiger, retrieval of target images for all other queries have been quite remarkable.

A.3.4 Face Extraction - Face Image BSDB [Fowlkes, on line] [Martin, 2001]

A typical image of BSDB database [Fowlkes, on line] [Martin, 2001] has been presented with face extraction results (Figure 111). The image contains skin-colored head-cape & cloths with typical pose of hands and face. Exclusion of face-touching hands in the detected face region is note-worthy. The precise face region extraction proves the effectiveness of proposed methods.



Figure 111. Image of BSDB [Fowlkes, on line] [Martin, 2001] and extracted face

Annexure 4 – Distinctive Issues

A-4.1 Search Engines - Google Image Search and vSearch

The topic presents query responses & related issues of state of the art image search technologies - Google Image Search and vSearch. Current version of the image search engine of Google, supporting image query or image url was launched recently in June 2011. Sources: (i) <http://www.google.com/insidesearch/press/launch.html> and (ii) <http://computervisioncentral.com/content/google-rolling-out-content-based-image-search01668>. Prior versions were supporting only textual queries. As being proprietary / commercial products, no authentic technical details of Google image search and vSearch are available. The inferred block diagram and related issues of Google search engine are illustrated below.

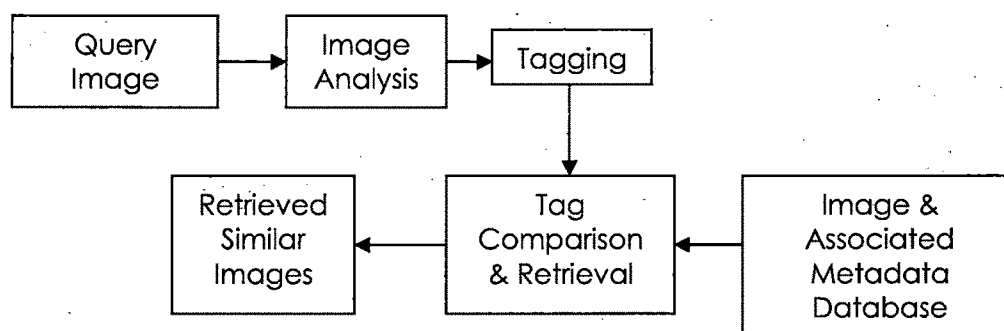


Figure 112. Inferred block diagram for Google like search engines

The Google achieves similarity retrieval by comparing Tags of the query image with pre-generated, stored & indexed Tags of database images. The Google addressed the issue of diversities in the human perception with the help of Google labeler. Google labeler was on line during 2006 to 2011. The Google labeler was a game to be played by multiple players who were given same images for manual labeling with all possible Tags. The same images would be given to a large number of users (players) to have exhaustive labeling. The exhaustive collections of image-tags have been utilized for the purpose of image retrieval. A tag for an image was subsequently weighted proportionate to the number of times it was perceived by the human beings. Hence, the label due to rarely or

wrongly perceived contents would be given less weight, putting such image at the lower rank at the time of retrieval for a match due to least important tag.

Though the current version of the Google gives better results in terms of the Precision for many queries, including the illustrative one – *black rose*, it is not free from the limitations. Two typical query examples to illustrate current state of the art and limitations of the search engine are shown below in Figure 113 (as on 10-04-2012), where automatic tagging of query images was not performed by Google and the user was asked to describe the image. The selected query images are from the standard database - Berkeley Segmentation Dataset and Benchmark (BSDB) [Fowlkes, on line]. Without describing prompted image contents, Google produced results of visually similar images which **were not containing** any images of Baby girl and Tigers respectively for given two queries on the first page of retrieved results. As observed, retrieval was mainly based on the color distributions.

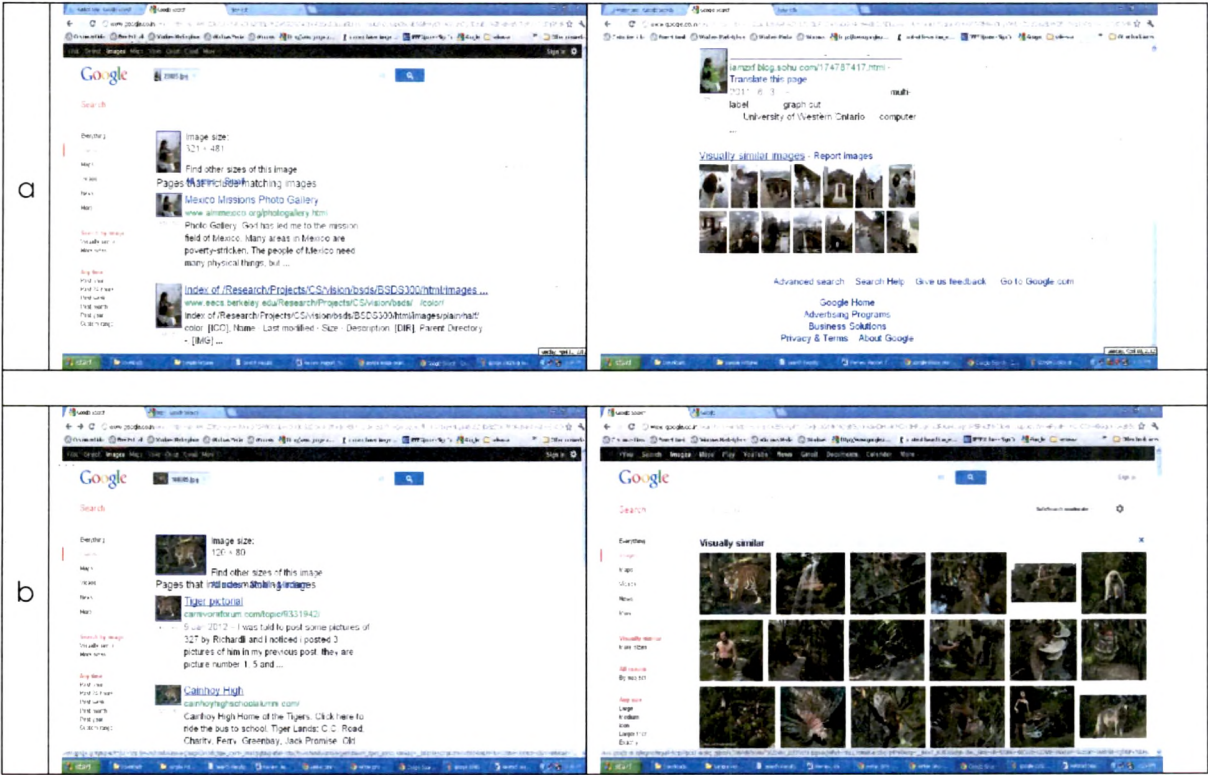


Figure 113. Query response for Google – could not tag & retrieving irrelevant images

Third typical Google query example shown below in Figure 114 is for the winter image of MS Windows operating system. The resulted automatic tags for the query image were winter & pins. The first tag is pertaining to a concept where as the second is a wrongly annotated tag, producing many dissimilar images on a first page.

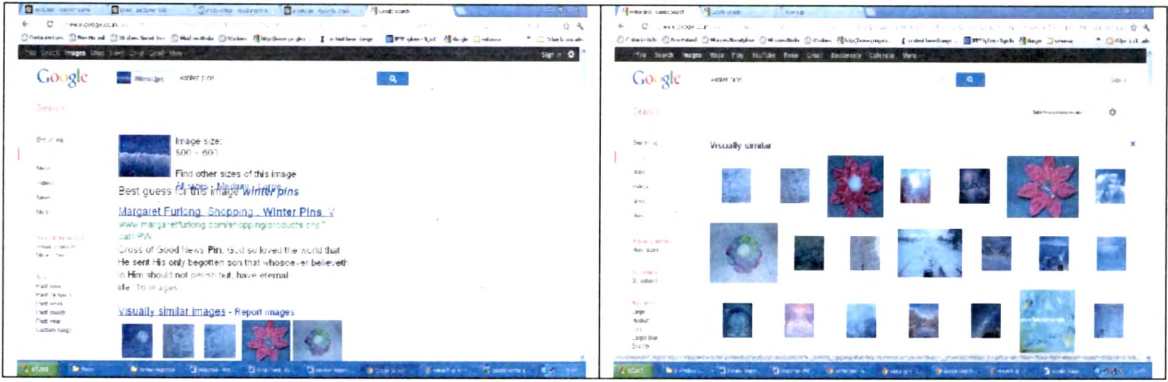


Figure 114. Query response for Google - tagging wrongly & retrieving many irrelevant images

Typical vSearch query results for 4 different queries are shown below in Figure 115. The top-left image in the table cells are respective query images. Resulted images of all 4 queries contain many dissimilar images. The inferred dominant method of similarity comparison is based on color distributions.

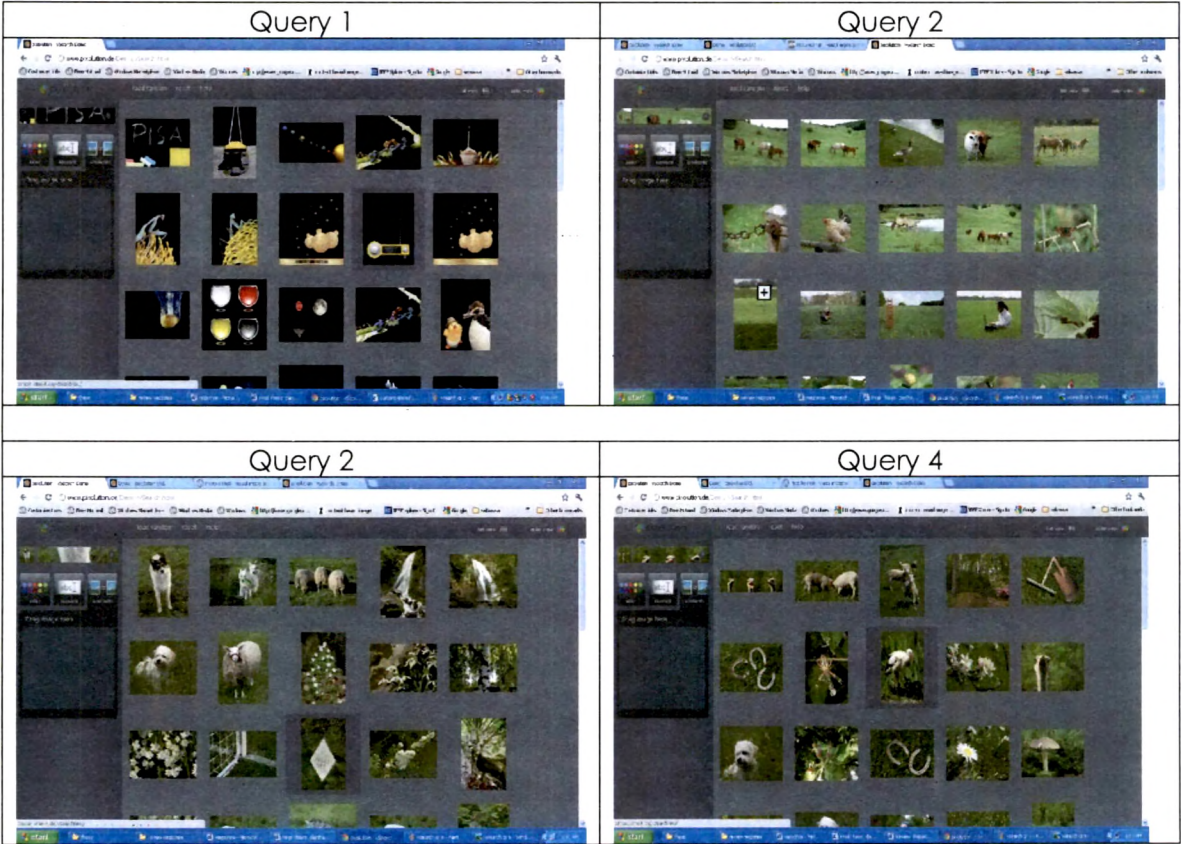


Figure 115. Query response for vSearch - retrieving many irrelevant (?) images

Our proposed novel techniques are based on the theme – “Relaxed feature description for better Recall and simultaneous emphasizing of reliable processing of cues leading to precise feature extraction for better Precision.” The user has been given choice to select method of image retrieval to map his needs & perception. The broader

color code description of whole image performs search on color similarity with higher Recall. The other options are foreground shape based technique, performing similarity comparison based on the detected foreground shape and a combinational technique that enables user to select proportionate weight of foreground shape and foreground color descriptors for similarity comparisons.

A-4.2 Quantitative Analysis & Comparisons of Edge Responses

Qualitative comparisons of edge responses of the proposed method with ACD Photo Editor, Adobe Photoshop and MS Photo Editor have been presented in Section 4.3.3. Figure 116 & Figure 117 show quantitative analysis of edge responses with performance measures Precision_e (P_e), Recall_e (R_e) and F – measure (F_e) along with qualitative comparisons for two sample images of BSDB [Fowlkes, on line] [Martin, 2001]. Precision_e is a measure of how many detected edges are correct and Recall_e is a measure of how many correct edges are detected with reference to ground truth. F-measure_e (F_e) combines Precision_e and Recall_e to yield performance reflective single number given by $2 / (1/P_e + 1/R_e)$. These measures are computed for detected perceptually significant edges with reference to human segmented image of BSDB [Fowlkes, on line] [Martin, 2001]. The computation of Precision_e (P_e) and Recall_e (R_e) are carried out by locating detected edges in a vicinity of +/- one pixel in all directions with reference to ground truth edges.

Figure 116 (a) left and Figure 117 (a) left show original images of BSDB [Fowlkes, on line] [Martin, 2001]. Corresponding human segmented images are shown in Figure 116 (a) middle & Figure 117 (a) middle respectively. Figure 116 (a) right & Figure 117 (a) right present edge responses of proposed method with threshold 25. All three leading tools - ACD Photo Editor, Adobe Photoshop and MS Photo Editor produce & present edge response as a color image as shown in Figure 116 (b) to (d) & Figure 117 (b) to (d) at first column. These responses are converted to Gray images and thresholded with thresholds 25, 64 & 128 for precision & recall computations with white representing edge pixel. The RGB to Gray conversion & thresholding is performed with a Matlab program. It should also be noted that Adobe & MS Photo produce color edge responses characterized by white background with colored edges as shown in Figure 116 (c) & (d) and Figure 177 (c) & (d). And hence, corresponding Gray images are required to be negated before thresholding for carrying out quantitative comparisons.

Precision, Recall and F – measure plotted at different thresholds for quantitative comparisons of edge responses of the proposed method with ACD Photo editor, Adobe

Photoshop and MS Photo editor have been presented in Figure 118, Figure 119 & Figure 120 respectively.





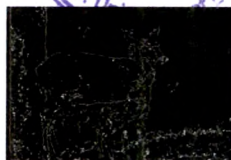
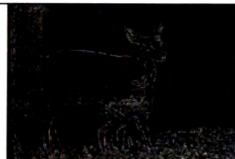











a)	 Original Image BSDB[Fowlkes, on line] [Martin, 2001]	 Human Segmented Image	 Edge Response of proposed method, SWT level 2, $P_e = 0.31$, $R_e = 0.52$, $F_e = 0.39$	
b)	 ACD Edge response (Color)	 Processed Edge response of ACD, Threshold 25, $P_e = 0.15$, $R_e = 0.86$, $F_e = 0.24$	 Processed Edge response of ACD, Threshold 64 $P_e = 0.20$, $R_e = 0.63$, $F_e = 0.3$	 Processed Edge response of ACD, Threshold 128, $P_e = 0.28$, $R_e = 0.40$, $F_e = 0.32$
c)	 Adobe Edge response (Color)	 Processed Edge response of Adobe, Threshold 25, $P_e = 0.07$, $R_e = 0.99$, $F_e = 0.12$	 Processed Edge response of Adobe, Threshold 64, $P_e = 0.11$, $R_e = 0.92$, $F_e = 0.18$	 Processed Edge response of Adobe, Threshold 128, $P_e = 0.19$, $R_e = 0.69$, $F_e = 0.30$
d)	 MS Photo Editor Edge response (Color)	 Processed Edge response of MS Photo Editor, Threshold 25, $P_e = 0.15$, $R_e = 0.80$, $F_e = 0.26$	 Processed Edge response of MS Photo Editor, Threshold 64, $P_e = 0.19$, $R_e = 0.57$, $F_e = 0.28$	 Processed Edge response of MS Photo Editor, Threshold 128, $P_e = 0.23$, $R_e = 0.36$, $F_e = 0.28$

Figure 116. Edge Response Comparison & quantitative analysis – example 1











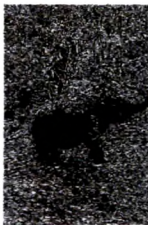

a)			
	Original Image BSDDB[Fowlkes, on line] [Martin, 2001]	Human Segmented Image BSDDB[Fowlkes, on line] [Martin, 2001]	Edge Response of proposed method, SWT level 2, $P_e = 0.17$, $R_e = 0.53$, $F_e = 0.26$
c)			
	ACD Edge response (Color)	Processed Edge response of ACD, Threshold 25, $P_e = 0.10$, $R_e = 0.93$, $F_e = 0.18$	Processed Edge response of ACD, Threshold 64 $P_e = 0.12$, $R_e = 0.87$, $F_e = 0.22$
			Processed Edge response of ACD, Threshold 128, $P_e = 0.14$, $R_e = 0.75$, $F_e = 0.24$
d)			
	Adobe Edge response (Color)	Processed Edge response of Adobe, Threshold 25, $P_e = 0.05$, $R_e = 0.98$, $F_e = 0.09$	Processed Edge response of Adobe, Threshold 64, $P_e = 0.06$, $R_e = 0.96$, $F_e = 0.11$
			Processed Edge response of Adobe, Threshold 128, $P_e = 0.08$, $R_e = 0.92$, $F_e = 0.14$
e)			
	MS Photo Editor Edge response (Color)	Processed Edge response of MS Photo Editor, Threshold 25, $P_e = 0.10$, $R_e = 0.91$, $F_e = 0.18$	Processed Edge response of MS Photo Editor, Threshold 64, $P_e = 0.12$, $R_e = 0.83$, $F_e = 0.2$
			Processed Edge response of MS Photo Editor, Threshold 128, $P_e = 0.14$, $R_e = 0.70$, $F_e = 0.24$

Figure 117. Edge response comparison & quantitative analysis – example 2

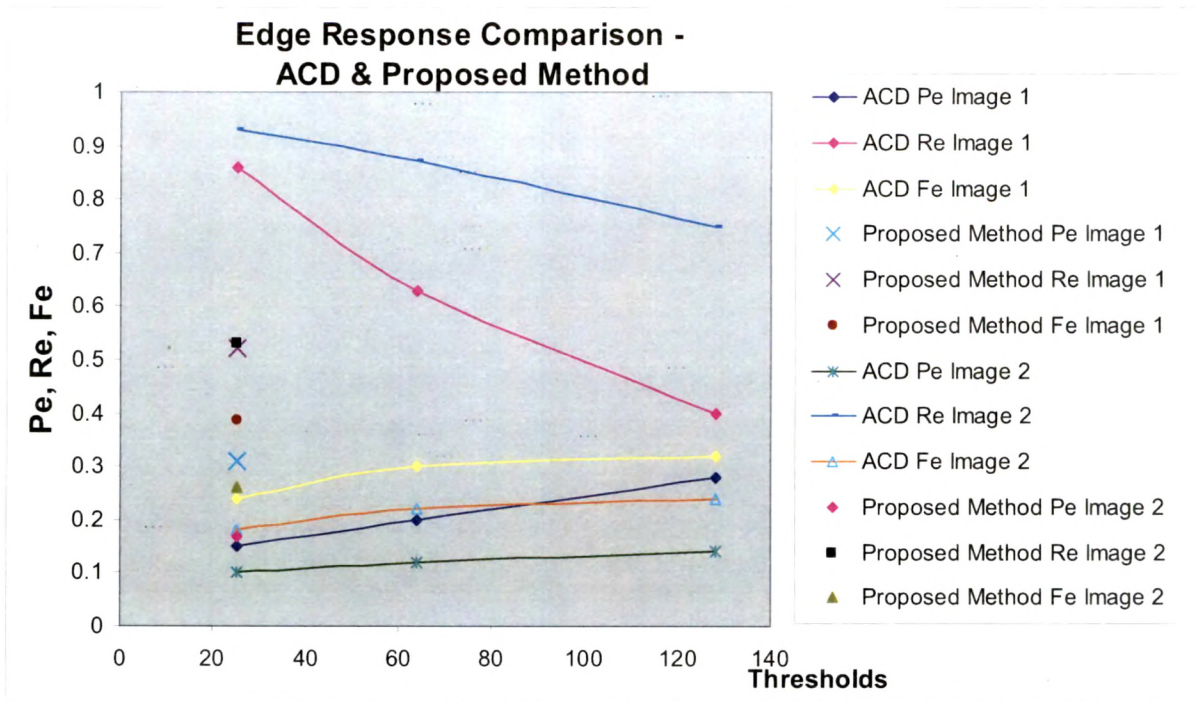


Figure 118. Quantitative comparison of edge responses: ACD Photo editor & proposed method

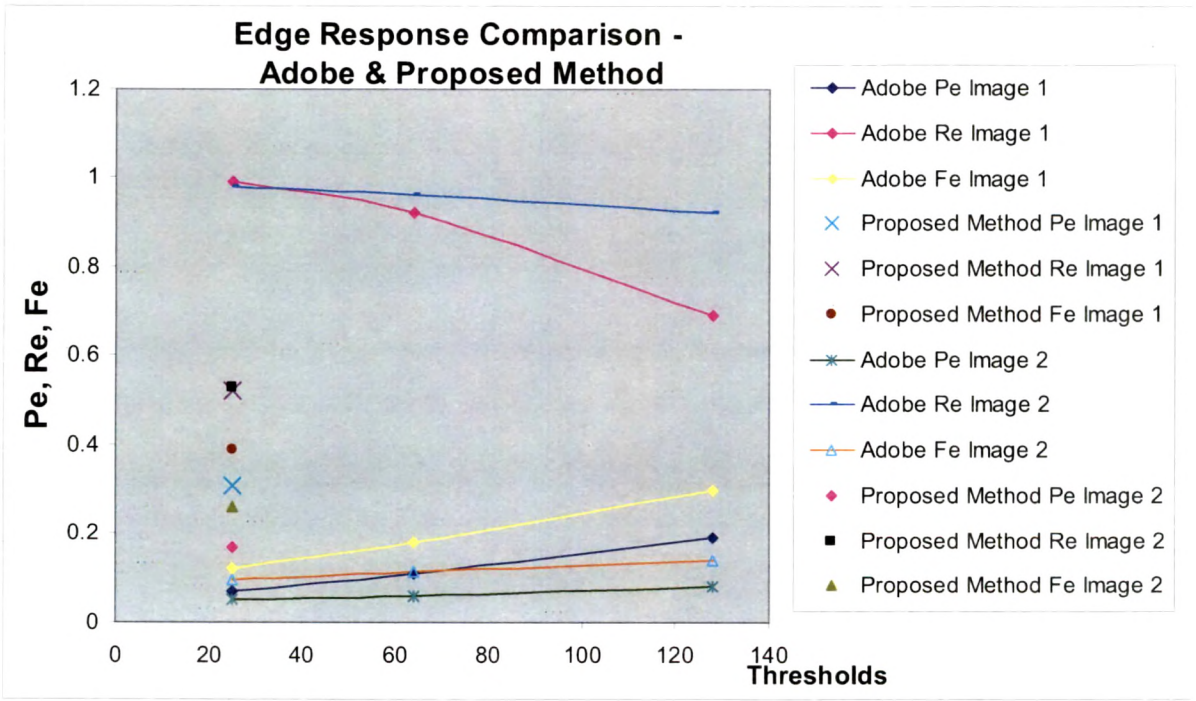


Figure 119. Quantitative comparison of edge responses: Adobe Photoshop & proposed method

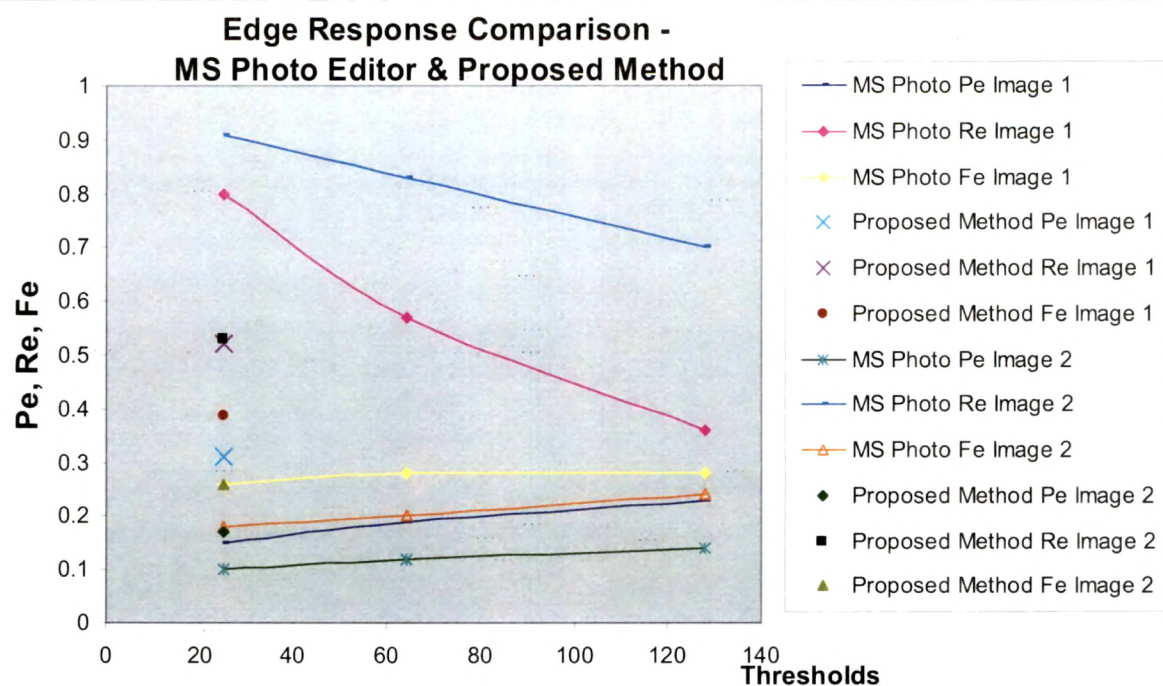


Figure 120. Quantitative comparison of edge responses: MS Photo editor & proposed method

A-4.2.1 Discussion

- Edge responses of the three software packages are characterized by low Precision resulted due to detection of perceptually significant as well as insignificant edges. Precision of Adobe is the lowest. Precision of the proposed method is better than others. The edge responses of the three packages are to be thresholded explicitly at different required levels for better Precisions & F-measures.
- Large numbers of detected edges yield better Recall for the three software packages compared to the proposed method. The higher Recall obtained in the three software packages is at a cost of Precision. F-measures of the proposed method is better than the three software packages.
- The results of the proposed method outperform others for i) detection of significant perceptual edges ii) elimination of insignificant edges corresponding background and foreground textures.

A-4.3 Quantitative Analysis of Results of Proposed Method for Foreground Extraction w. r. t. Ground Truth

The quantitative comparisons of the results of proposed method for foreground extraction have been carried out with performance measures $Precision_{fg}$ and $Recall_{fg}$, computed with respect to Ground Truth foreground. High values of the performance measures are noteworthy (Figure 121 to Figure 126).

The computation of $Precision_{rg}$ and $Recall_{rg}$ differs for region based segmentation compared to the Precision and Recall used for measuring image retrieval performances. They are defined as follows:

$Precision_{rg}$ is a ratio of area of intersection of detected foreground regions with Ground Truth foreground region to area of Detected foreground regions.

$Recall_{rg}$ is a ratio of area of intersection of detected foreground regions with Ground Truth foreground regions to area of Ground Truth foreground regions.

E.g., $Precision_{rg}$ of 0.5 indicates that correctly detected foreground (with reference to Ground-Truth) is 50 % of the total detected foreground. $Recall_{rg}$ of 0.6 indicates that the correctly detected foreground (with reference to Ground-Truth) is 60% of the total correct (Ground-Truth) foreground.

In addition to a large number of results presented in the thesis for foreground extraction for qualitative comparisons with the Human segmented images of BSDb [Fowlkes, on line] [Martin, 2001], qualitative & quantitative analysis for performance measures have been carried on sample images of BSDb [Fowlkes, on line] [Martin, 2001], SIMPLcity [SIMPLcity, on line] and ALOI [ALOI, on line] [Geusebroek, 2001].

Ground Truth Foreground Images: The BSDb [Fowlkes, on line] [Martin, 2001] provides Human segmented Ground Truth images. These images contain segmented foreground and background regions. And hence, the Ground Truth foreground images have been produced manually using Adobe Photoshop from these Human segmented images. The images from other databases have been also processed with Adobe Photoshop to generate Ground Truth foreground images.

Following Figures give the qualitative and quantitative comparisons of the results with the Ground Truth. The original images and corresponding human segmented images of BSDb [Fowlkes, on line] [Martin, 2001] are shown in Figure 121 (a) and Figure 122 (b) respectively. Figure 121 (c) shows Ground-Truth foreground images produced with Adobe Photoshop from respective images of Figure 121 (b). The foreground regions are marked with White. Figure 121 (d) indicates level of Haar SWT used for proposed algorithm. Figure 121 (e) shows the extracted foreground regions from original images of (a) with proposed foreground extraction algorithm. These extracted foreground regions are marked with White and can be qualitatively compared with the corresponding Ground Truth foreground shown in Figure 121 (c). These foreground regions (White) of Figure 121 (e) are mapped to images to yield foreground images shown in Figure 121 (f) containing background marked as Black. Figure 121 (g) and Figure 121 (h) are the

quantitative measures of P_{fg} and R_{fg} respectively, for extracted foreground with proposed algorithm with reference to the Ground Truth foreground.

Figure 122 illustrates results and quantitative & qualitative analysis for foreground extraction carried out at two different levels of Haar SWT decomposition for images of BSDb [Fowlkes, on line] [Martin, 2001]. The chart for P_{fg} & R_{fg} for results of Figure 121 and Figure 122 have been presented in Figure 123 indicative of high average P_{fg} & high average R_{fg} .

Figure 124 and Figure 125 illustrate high performance measures with respect to ground truth for images of other databases - SIMPLicity [Wang, 2001] [SIMPLicity, on line] and ALOI [ALOI, on line] [Geusebroek, 2001]. The corresponding chart has been presented in Figure 126.








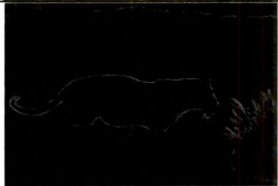












a)				
b)				
c)				
d)	Level 2	Level 3	Level 2	Level 2
e)				
f)				
g)	$P_{fg} = 0.55$	$P_{fg} = 0.85$	$P_{fg} = 0.80$	$P_{fg} = 0.73$
h)	$R_{fg} = 0.91$	$R_{fg} = 0.97$	$R_{fg} = 0.66$	$R_{fg} = 0.85$

Figure 121. Qualitative & Quantitative Performance Comparisons for foreground extraction. (a) Original BSDb Images BSDb [Fowlkes, on line] [Martin, 2001] (b) Human segmented Ground Truth images at BSDb [Fowlkes, on line] [Martin, 2001]. (c) Ground Truth foreground from (b), produced with Adobe Photoshop. (d) Level of Haar SWT used. (e) Extracted foreground regions from Original images of (a) produced with proposed algorithm. (f) Corresponding foreground image, mapped from (e). (g) and (h) P_{fg} and R_{fg} respectively for extracted foreground regions of (e).















a)				
b)				
c)				
d)	Level 2	Level 3	Level 2	Level 3
e)				
f)				
g)	$P_{fg} = 0.41$	$P_{fg} = 0.52$	$P_{fg} = 0.37$	$P_{fg} = 0.53$
h)	$R_{fg} = 0.98$	$R_{fg} = 0.95$	$R_{fg} = 0.81$	$R_{fg} = 0.67$

Figure 122. Qualitative & Quantitative Performance Comparisons for foreground extraction with respect to different levels of Haar SWT. (a) Original BSD Images BSDB [Fowlkes, on line] [Martin, 2001]. (b) Human segmented Ground Truth images at BSDB[Fowlkes, on line] [Martin, 2001]. (c) Ground Truth foreground from (b), produced with Adobe Photoshop. (d) Level of Haar SWT used. (e) Extracted foreground regions from Original images of (a) produced with proposed algorithm. (f) Corresponding foreground image, mapped from (e). (g) and (h) Precisionfg and Recallfg respectively for extracted foreground regions of (e).

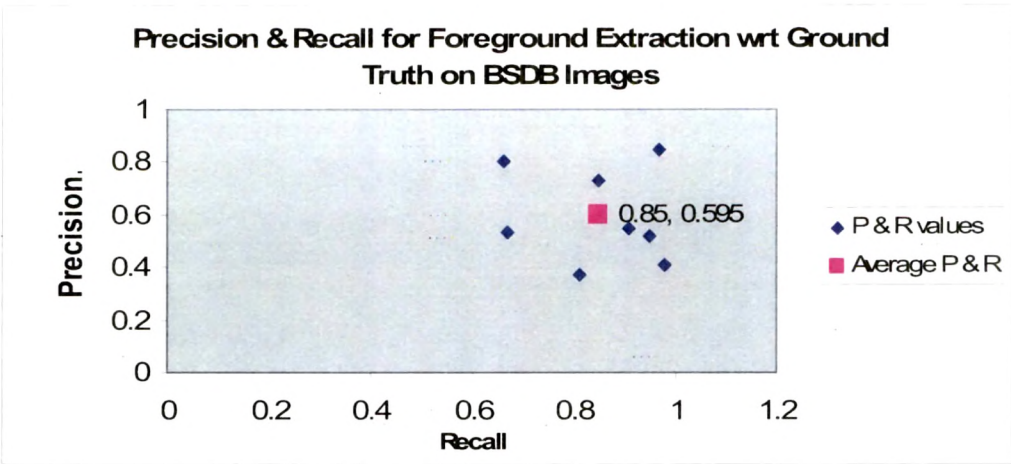


Figure 123. Quantitative analysis w.r.t. ground truth for foreground extraction on BSDB Images

a)			
b)			
c)	Level 2	Level 1	Level 1
d)			
e)			
f)	$P_{fg} = 0.70$	$P_{fg} = 0.92$	$P_{fg} = 0.73$
g)	$R_{fg} = 0.92$	$R_{fg} = 0.95$	$R_{fg} = 0.93$

Figure 124. Qualitative & Quantitative Performance Comparisons for foreground extraction on images with illumination variations. (a) Original images, Left - size reduced image photographed by an amateur, Middle & Right from ALOI [ALOI, on line] [Geusebroek, 2001]. (b) Ground Truth foreground from (a) produced with Adobe Photoshop. (c) Level of Haar SWT used. (d) Extracted foreground regions from Original images of (a) produced with proposed algorithm. (e) Corresponding foreground image, mapped from (d). (f) and (g) Precisionfg and Recallfg respectively for extracted foreground regions of (d).

a)			
b)			
c)	Level 2	Level 3	Level 2
d)			
e)			
f)	$P_{fg} = 0.64$	$P_{fg} = 0.43$	$P_{fg} = 0.74$
g)	$R_{fg} = 0.93$	$R_{fg} = 0.90$	$R_{fg} = 0.98$

Figure 125. Qualitative & Quantitative Performance Comparisons for foreground extraction. (a) Original images [Wang, 2001] [SIMPLIcity, on line]. (b) Ground Truth foreground from (a) produced with Adobe Photoshop. (c) Level of Haar SWT used. (d) Extracted foreground regions from Original images of (a) produced with proposed algorithm. (e) Corresponding foreground image, mapped from (d). (f) and (g) Precision_{fg} and Recall_{fg} respectively for extracted foreground regions of (d).

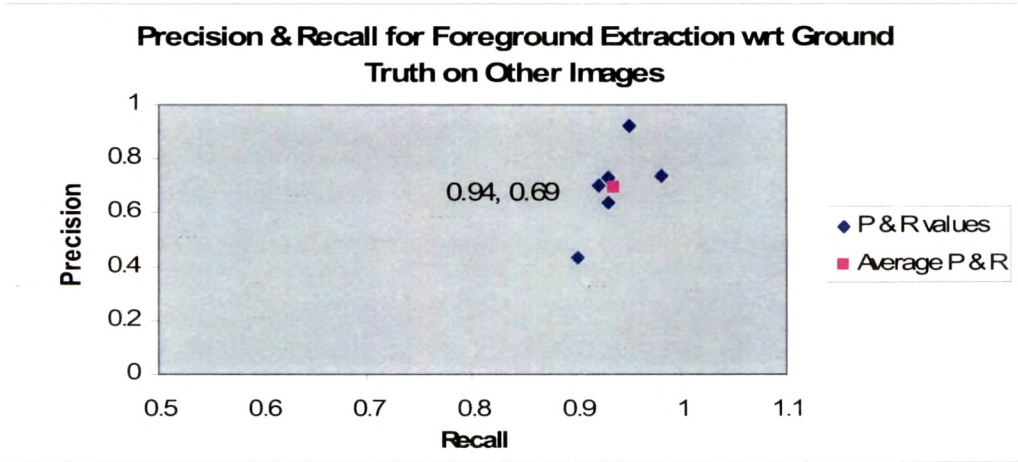


Figure 126. Precision – Recall analysis w. r. t. ground truth for foreground extraction on BSDB images [Fowlkes, on line]

A-4.3.1 Discussion

- The proposed method of foreground extraction is effective on diversified images as can be concluded by qualitative & quantitative comparisons of Ground Truth foregrounds with extracted foregrounds with proposed method. $Precision_{fg}$ & $Recall_{fg}$ - quantitative performance measures with the respect to Ground truth are quite high.
- The $Precision_{fg}$ of extracted foreground is high even for the complex natural images of BSDb [Fowlkes, on line] [Martin, 2001]. The average Precision of 0.595 with average Recall of 0.85 for sample image set is quite significant.
- The work reported so far in the literature for foreground extraction of challenging images of BSDb [Fowlkes, on line] [Martin, 2001] is mainly based on Graph cuts. The proposed method and the qualitative results with high Precision & Recall are unique, novel & not reported so far.

A-4.4 Color Codes

Innovative and unique 27 ($25 + 2$) color codes to effectively represent entire spectrum of RGB color space (2^{24} colors) are formulated and used for the purpose of image segmentation and feature extraction leading to color similarity based image retrieval. The codes are formulated by exploiting intra-tuple RGB relationship as shown in Table 27. A Color Code represents a set of colors satisfying corresponding intra-tuple RGB relationship. These color codes are the broadest color descriptors used to represent color attributes of images. Color code_0 represents pure black (with $R = G = B = 0$) differentiating it from Code_1. Code_26 is a special code to represent colors around boundaries of color codes.

A-4.4.1 Results - Color Code Based Segmentation

Figure 127 illustrates effectiveness of proposed novel color codes to represent image color attributes leading to image segmentation. The depicted sample images are some of the most challenging images of standard databases of BSDb [Fowlkes, on line] [Martin, 2001] and SIMPLicity [Wang, 2001]. A standard image of Baboon possessing typical textures and color combinations is also segmented effectively with 4.1 seconds as segmentation time on the dual core processor with 1.49 GB of RAM. Segmentation time analysis is further reported in Section A-4.5.

Despite being broadest color descriptors, their effective representations for colors of images have been exploited for image retrieval to increase the Recall. The combination of foreground color codes and foreground shape with selectable

percentage weight for image retrieval not only maps the need & perception of a user but also increases the Precision without sacrificing Recall much.

Table 27. Formulation of Color Codes

Sr. No.	Code	R G B Relationship Deciding Set of Colors Mapping to Respective Color Code
1.	Code_1	$R = B = G \text{ \& } R \neq 0$
2.	Code_2	$R = G \text{ \& } R > B$
3.	Code_3	$R = B \text{ \& } R > G$
4.	Code_4	$G = B \text{ \& } G > R$
5.	Code_5	$R = G \text{ \& } B > R$
6.	Code_6	$R = B \text{ \& } G > R$
7.	Code_7	$G = B \text{ \& } R > G$
8.	Code_8	$R > B > G \text{ \& } (R - B) = (B - G)$
9.	Code_9	$R > B > G \text{ \& } (R - B) > (B - G)$
10.	Code_10	$R > B > G \text{ \& } (R - B) < (B - G)$
11.	Code_11	$R > G > B \text{ \& } (R - G) = (G - B)$
12.	Code_12	$R > G > B \text{ \& } (R - G) > (G - B)$
13.	Code_13	$R > G > B \text{ \& } (R - G) < (G - B)$
14.	Code_14	$G > R > B \text{ \& } (G - R) = (R - B)$
15.	Code_15	$G > R > B \text{ \& } (G - R) > (R - B)$
16.	Code_16	$G > R > B \text{ \& } (G - R) < (R - B)$
17.	Code_17	$G > B > R \text{ \& } (G - B) = (B - R)$
18.	Code_18	$G > B > R \text{ \& } (G - B) > (B - R)$
19.	Code_19	$G > B > R \text{ \& } (G - B) < (B - R)$
20.	Code_20	$B > G > R \text{ \& } (B - G) = (G - R)$
21.	Code_21	$B > G > R \text{ \& } (B - G) > (G - R)$
22.	Code_22	$B > G > R \text{ \& } (B - G) < (G - R)$
23.	Code_23	$B > R > G \text{ \& } (B - R) = (R - G)$
24.	Code_24	$B > R > G \text{ \& } (B - R) > (R - G)$
25.	Code_25	$B > R > G \text{ \& } (B - R) < (R - G)$
26.	Code_0	$R = B = G = 0$
27.	Code_26	Special


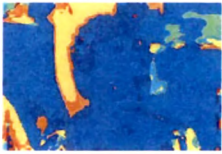

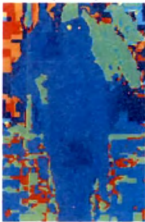

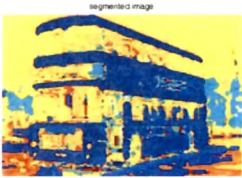

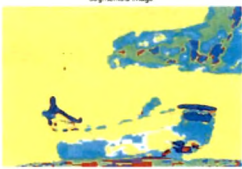

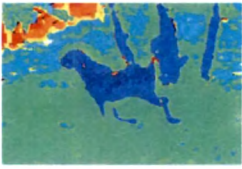

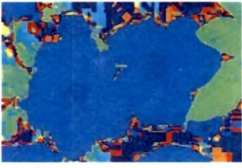

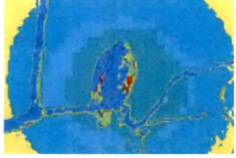
Sr. No.	Original Images SIMPLcity [Wang, 2001], BSD3 [Fowlkes, on line] [Martin, 2001]	Color code based segmented images
1)	1.jpg 	
2)	44.jpg 	
3)	310.jpg 	
4)	122.jpg 	
5)	740.jpg 	
6)	670.jpg 	
7)	42049.jpg 	

Figure 127. Results: Color codes based segmentation.


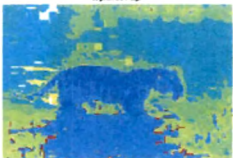

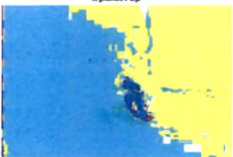


Sr. No.	Original Images SIMPLcity [Wang, 2001], BSDb [Fowlkes, on line] [Martin, 2001]	Color code based segmented images
9)	108073.jpg 	
10)	30091.jpg 	
11)	BABOON.jpg 	

Figure 127 (Contd.). Results: Color codes based segmentation.

A-4.5 Processing Time Analysis

The development, testing & implementation of algorithms has been carried out on a machine having dual core Intel processor (T2050 @ 1.6 GHz) with 1.49 GB of RAM. The high processing time required particularly for prominent boundaries based algorithms demands high end servers for their deployment at real time.

Table 28 presents processing time for segmentation with Color codes and total processing time for edge & prominent boundaries detection and foreground extraction on sample images of SIMPLcity images [Wang, 2001] of size 384 x 256. The processing time for boundary detection based approach is tabulated for SWT Haar level 1 and level 2. Whereas Table 29 summarizes processing time for various images of BSDb images [Fowlkes, on line] [Martin, 2001] of size 421 X 381 processed for color codes based segmentation and boundary based algorithms with SWT Haar decomposition at level 2 and level 3.

Table 28. Processing time analysis for SIMPLIcity images













Sr. No	SIMPLIcity Images [Wang, 2001] Size - 384 x 256	Processing Time Seconds	
		Color Code Based Segmentation Algorithm	Boundary Detection Based Algorithms
		Attempt 1	SWT Haar Level 1
		Attempt 2	SWT Haar Level 2
1	1.jpg 	1.9	234.9
		1.65	205.2
2	44.jpg 	1.5	154.1
		1.85	103.3
3	310.jpg 	1.68	2177.7
		2.32	605.2
4	122.jpg 	1.54	677.6
		1.77	809
5	740.jpg 	1.90	91.1
		1.79	177.76
6	670.jpg 	2.1	229.75
		2.7	141.3

Table 29. Processing time analysis for BSDb images

Sr. No	BSDb Images [Fowlkes, on line] [Martin, 2001] Size - 421 X 381	Processing Time Seconds	
		Color Code Based Segmentation Algorithm	Boundary Detection Based Algorithms
		Attempt 1	SWT Haar Level 2
		Attempt 2	SWT Haar Level 3
1	42049.jpg 	1.64	544.7
		1.89	374.3
2	300091.jpg 	1.93	232.2
		1.7	175.7
3	296059.jpg 	1.78	246.3
		2.03	226.7
4	38092.jpg 	1.65	438.5
		1.70	290.5
5	108005.jpg 	1.70	547.5
		1.68	245.3
6	108073.jpg 	1.71	393.7
		1.78	228.2

A-4.5.1 Discussion

- Color code based segmentation requires less time compared to boundary detection based algorithms.
- Processing time of images containing textures is significantly high for boundary detection based algorithms. Processing time reduces at higher levels of SWT.
- Processing time of color codes based segmentation does not depend on textures or categories of images. It is merely proportionate to the size of the image.
- Considering effectiveness and suitability of proposed boundary detection based algorithms, high end servers are needed & recommended to meet their computational requirements.