

Contents

List of Figures	xiv
List of Tables	xxviii
Abbreviations	xx
Symbols	xxiv
Preface	xxvii
1 Introduction	1
1.1 Background	1
1.2 Motivation and Scope	3
1.3 Problem Statement	5
1.4 The Research Objectives of Thesis	6
1.5 Contributions to the Thesis	7
1.6 Outline of the Thesis	9
2 Literature Survey	11
2.1 Literature Review	12
2.1.1 Visual Saliency	12
2.1.2 Visual saliency models	13
2.1.2.1 Fixation Prediction models	13
2.1.2.2 Salient object detection models	14
2.1.3 Saliency computation domains:	16
2.1.3.1 Bottom-up-search dependent statistical models: . . .	17
2.1.3.2 Top-down-task dependent deep learning based models:	17
2.1.3.3 Hybrid and Probabilistic models :	18
2.2 Salient Object Detection	20
2.2.1 Statistical and Probabilistic Models(2D Saliency)	20
2.2.1.1 Biologically Inspired Models	21
2.2.1.2 Global Contrast-Based Models	22

2.2.1.3	Regional Contrast-Based Models	23
2.2.1.4	Background Approximation Based Models	29
2.2.1.5	Psychology Inspired Models	31
2.2.1.6	Summary	35
2.2.2	RGBD(3D) Statistical Models	37
2.2.2.1	Summary	43
2.2.3	RGBD(3D) Deep learning Models	43
2.2.3.1	Multi-Layer Perceptron (MLP) Models	44
2.2.3.2	Fully Convolutional Network (FCN) Models	45
2.2.3.3	RGBD 3D SOD Models	46
2.2.3.4	Single-stream Models	47
2.2.3.5	Multi-stream Models	47
2.2.3.6	Fusion Model	48
2.2.3.7	Skip-Connection based models	53
2.2.3.8	Attention mechanism	54
2.2.3.9	Summary	56
2.3	Research Gaps	57
2.4	Dataset	61
2.4.1	RGB(2D) Dataset	62
2.4.2	RGBD(3D) Dataset	65
2.5	Performance Evaluation Metrics	69
2.6	Conclusion	73
3	Probabilistic Contrast and Edge Enhanced Global Topographical Surface based Complex RGB Salient Object Detection	75
3.1	Introduction	77
3.2	The Proposed Models	83
3.2.1	Probabilistic contrast based complex salient object detection .	84
3.2.1.1	Initialization through Poisson probabilistic contrast (PC)	84
3.2.1.2	Poisson probabilistic distribution	86
3.2.1.3	Regional contrast integration with global-PC	87
3.2.1.4	Background suppression model	88
3.2.1.5	Saliency Refinement Model	89
3.2.1.6	Theoretic foundation	90
3.2.2	Experiments and Results of Probabilistic Contrast based Model	94
3.2.2.1	Evaluating Parameters Setting	95
3.2.2.2	Ablation study	98
3.2.2.3	Comparative study	100
3.2.2.4	Comparison with Deep Learning Based Models	106
3.2.3	Edge Enhanced Global Topographical Saliency	108

3.2.3.1	Initial global topographical surface(GTS) through iterative Laplacian of Gaussian (ILG)	108
3.2.3.2	Regional contrast integration within GTS	109
3.2.3.3	Saliency enhancement	112
3.2.4	Result Analysis of Edge Enhanced Global Topographical Saliency	114
3.2.4.1	Evaluating parameters setting	115
3.2.4.2	Successive steps validation	116
3.2.4.3	Comparative analysis	116
3.2.4.4	Comparison of proposed method GTS with Deep Learning Based Models:	122
3.2.4.5	Comparison of average computational speed :	122
3.3	Conclusion	123
4	RGBD Complex salient object detection with improved probabilistic contrast and global concave topographical saliency	125
4.1	Introduction	126
4.2	The proposed method	131
4.2.1	Initialization through global concave surface (GCS)	131
4.2.1.1	Improved Poisson probabilistic contrast (IPC)	132
4.2.1.2	Contour based global surface	133
4.2.2	Regional contrast integration into GCS	134
4.2.2.1	Regional saliency integration with GCS	135
4.2.2.2	Background estimation model	136
4.2.2.3	Saliency enhancement	137
4.2.3	Theoretic foundation	138
4.2.4	The proposed algorithm	140
4.3	Experimental and Result Analysis	141
4.3.1	Data-Set	141
4.3.2	Performance Measures	142
4.3.3	Parameters and Constraints Selection	142
4.3.4	Successive steps validation	143
4.3.5	Comparative Analysis	144
4.3.6	Comparison with other Own proposed methods	148
4.4	Conclusion	149
5	CSA-Net:Deep Cross Complementary Self Attention and Modality-Specific Preservation for Saliency Detection	151
5.1	Introduction	152
5.2	The Proposed Model	157
5.2.1	Modality-specific Saliency Fusion Model(MSF)	159
5.2.2	Cross -Complementary Self-Attention -CSA	160
5.2.3	Complementary Features Fusion Model	163

5.2.3.1	Cross-Complementary Features Fusion(CFF)	163
5.2.3.2	Intra-Complementary Features Fusion(IFF)	165
5.2.4	Optimal Selective Saliency Fusion Model-OSS	166
5.2.5	Loss Function	168
5.3	Experimental Details:	169
5.3.1	Data-Set	169
5.3.2	Implementation Details	170
5.3.3	Training Details	171
5.3.4	Evaluation Metrics	172
5.3.4.1	S-Measure	173
5.3.4.2	F-Measure	173
5.3.4.3	Mean Absolute Error(MAE)	174
5.3.4.4	E-Measure(E_ψ)	174
5.4	Comparison and Result Analysis	176
5.4.1	Ablation Analysis	178
5.4.1.1	Validation of Network Structure and Fusion model	178
5.4.1.2	Effectiveness of CSA-Net Model	180
5.4.1.3	Failure Cases and Analyses	181
5.5	Conclusion	182
6	SL-Net: Self-Learning and Mutual Attention based Distinguished Window for RGBD Complex Salient Object Detection	185
6.1	Introduction	186
6.2	The proposed method	192
6.2.1	Overview	192
6.2.2	Enhanced Encoded Feature through Composite Backbone	194
6.2.3	Mutual Attention Based Distinguished Window-MADW	194
6.2.4	Self-Learning based Dense Decoding-SDD	198
6.2.4.1	Global Localized Feature	198
6.2.4.2	Cross-complementary fusion(CF)	199
6.2.4.3	Self Learning Aggregation model	200
6.2.5	Loss Function	201
6.3	Experiment and Result Analysis	202
6.3.1	Data-Set	202
6.3.2	Evaluation metrics	203
6.3.3	Implementation Details	203
6.3.4	Comparison and Result Analysis	204
6.3.5	Ablation Analysis	208
6.3.5.1	Effectiveness of Mutual attention based distinguished window-MADW	209
6.3.5.2	Effectiveness of SDD and Fusion model	210
6.3.5.3	Effectiveness of Proposed Composite model	211

6.3.6	Discussion of Failure case and limitations	212
6.3.7	Comparison of own proposed methods	213
6.4	Conclusion	214
7	Conclusion and Future Work	215
7.1	Conclusions	215
7.2	Future Work	219
	References	221
	List of Publications	245

List of Figures

1.1	A set of images to demonstrate the challenge to extract salient objects in complex and cluttered backgrounds.	4
2.1	Visual saliency and Computational domains.	16
2.2	Computational domains and associated outstanding models in Salient Object Detection.	18
2.3	Regional saliency and Superpixel based salient object detection.	24
2.4	The 3D saliency and Depth Information.	38
2.5	A set of RGB(2D) Datasets.	63
2.6	A set of RGBD(3D) Datasets.	67
3.1	The motivation for proposing probabilistic models to minimize discrepancies in interior and exterior saliency.	78
3.2	The importance of Global Topographical Surface in salient objects from complex and clutter backgrounds	82
3.3	Block diagram of proposed probabilistic model.	83
3.4	Enhancement and suppression of luminance (ESL) with relative variance in chrominance planes, where (a) Input image (b) Input image luminance plane (c) Normalized luminance plane.	86
3.5	Successive saliency computation (a) Input image (b)Initial S_{pc} through, (c) Final S_{pc} through, (d) S_{PR} regional color, spatial, depth enhance saliency (e) S_c central saliency, (f) Final saliency S through	90
3.6	PR-Curves on MSRA dataset for demonstrating the evaluating parameters. (a) Angular threshold, (b) Complex image indicator(λ). Blue line represents the without ESL.	95
3.7	Visual comparison of saliency maps related with contrast based method.	96
3.8	Comparison of purely contrast-based saliency on (1) MSRA (2) PASCAL (3) DUTOMRON, dataset respectively	97
3.9	Comparison of purely contrast-based saliency on (1) ECSSD and (2) IMGSA dataset respectively	98
3.10	Steps wise validation of proposed method using F-Measure.	99
3.11	Visual comparison of saliency maps of a complex image having cluttered background.	101

3.12	Quantitative comparison of saliency Maps on (1)ECSSD (2) IMGSA (3) MSRA dataset respectively	102
3.13	Quantitative comparison of saliency maps on (1) PASCAL and (2) DUTOMRON dataset respectively	103
3.14	Quantitative comparison of saliency maps using F-Measure.	104
3.15	Qualitative comparison of proposed method with deep learning based methods.	106
3.16	Demonstration of successive steps of proposed method GTS (a) Input image, (b) Initialization through iterative ILG, (c) S_{GTS} , (d) Regional saliencies integration S_{FS} (e) $S_{Central}$ central saliency, (f) Gaussian weighted background suppression and central saliency enhancement S	112
3.17	Qualitative visual comparison among Various salient object detection methods.	115
3.18	Comparison of Global-Contrast-based saliency using (a) PR-Curve (b) ROC-Curve on (1) MSRA (2) PASCAL dataset respectively.	117
3.19	Comparison of Global-Contrast-based saliency using (a) PR-Curve (b) ROC-Curve on (1) DUTOMRON and (2) ECSSD dataset respectively.	118
3.20	Comparison of Global-Contrast-based saliency using F-Measurer on (1) MSRA (2) PASCAL (3) DUTOMRON and (4) ECSSD dataset respectively.	120
3.21	Comparison of deep learning based methods with GTS on PASCAL dataset using PR-Curve and F-Measure.	121
4.1	Saliency detection in complex and clutter background with minimize discrepancies in interior, exterior and border regions saliency.	129
4.2	Block diagram of proposed Improved Probabilistic Model.	132
4.3	Visual demonstration of Contribution of each step in GCS	143
4.4	Visual comparison of saliency of proposed method with other state- of-art methods.	144
4.5	Quantitative comparison of proposed method GCS with PR-Curve and ROC-Curve.	146
4.6	Quantitative comparison of proposed method with F-Measure	147
4.7	Comparison of GCS with other proposed own probabilistic contrast based method using PR-Curve	148
5.1	The 2×3 encoder and decoder streams utilize, deeply guided proposed attention map $CSA - Net$ to find the exact salient object.	153
5.2	The proposed two-stage Cross -complementary Self-Attention- CSA model is based on the Non-Local Network.	161
5.3	The proposed architecture of 2×3 encoder and decoder, $CSA - Net$ models.	165
5.4	The block diagram of the Intra-complementary features Fusion model(IFF).	167

5.5	The Visual comparison of the proposed model with other State-of-the-art-methods.	172
5.6	The successive validation of three-stream networks in Complex Images with inferior and low depth images.	178
5.7	The visual demonstration and validation of attention map learned through proposed deeply guided, two-stage additive, Cross-complementary Self-Attention(CSA-Net).	180
5.8	The visual demonstration of failure cases.	181
6.1	The importance of mutual attention mechanism to distinguish the salient object in complex and clutter background.	187
6.2	The illustration of the proposed framework <i>SL-Net</i>	193
6.3	The illustration of the proposed Mutual Attention Based Distinguished Window-MADW.	196
6.4	The design and process of Self-Learning based Dense Decoder- SDD .	197
6.5	Visual Demonstration of proposed method SL-Net with other closely and recent State-of-art-methods.	205
6.6	Visual illustration of the global localized features.	210
6.7	Visual illustration of failure case in incomplete depth maps and complex and cluttered background.	212

List of Tables

2.1	Comprehensive details of exemplary models for salient object detection using statistics and probabilistic approaches.	27
2.2	Comprehensive details of Top Performing conventional 3D model of salient object detection.	41
2.3	Comprehensive Survey of some Deep learning based RGBD Salient object detection model	50
2.4	RGB Dataset used in various saliency computations	62
2.5	RGBD Dataset for Salient Object Detection	66
3.1	Steps wise MAE in proposed Algorithm	100
3.2	Steps wise Mean Absolute Error-MAE in proposed Method-GTS . . .	116
3.3	Mean Absolute Error-MAE of different state-of-the-art methods . . .	121
3.4	Mean Absolute Error-MAE of different state-of-the-art methods . . .	121
3.5	Comparison of average computational run time.	123
3.6	Comparison of average computational run time('*' = GPU time). . . .	123
4.1	Steps wise Mean Absolute Error-MAE in proposed Method-GTS . . .	143
4.2	Mean Absolute Error-MAE of different state-of-the-art methods . . .	147
5.1	The quantitative comparison of proposed framework on seven benchmark RGBD datasets with four recent evaluation parameters.	175
5.2	The validation of the effectiveness of Network Architecture and Streams-wise saliency using Mean Absolute Error-MAE.	179
5.3	The ablation study of each component in the CSA-Net module.	181
6.1	The quantitative comparison of proposed framework on seven benchmark RGBD datasets with four recent evaluation parameters.	207
6.2	The ablation study of each component in the <i>SL - Net</i>	208
6.3	Validation of stage-wise improvements using proposed Mutual Attention based Distinguished Window-MADW	211
6.4	The ablation analysis of Cross-complementary fusion using <i>SDD</i> module and basic fusion strategy.	212
6.5	The comparison of own proposed deep learning based models.	213