05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

A NEW METHODOLOGY FOR EVALUATION AND BENCHMARKING OF SKIN DETECTOR BASED ON AI MODEL USING MULTI CRITERIA ANALYSIS

QAHTAN MAJEED YAS

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENT FOR DOCTOR OF PHILOSOPHY (ARTIFICIAL INTELLIGENCE)

FACULTY OF ART, COMPUTING & CREATIVE INDUSTRY
UNIVERSITI PENDIDIKAN SULTAN IDRIS

2018

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

ABSTRACT

This study aims to develop a new multi-criteria decision analysis methodology for skin detector evaluation and benchmarking based on artificial intelligence models. Two experiments were conducted. The first experiment comprised two stages: (1) Adaptation of the best previous case of skin detection approach utilizes multi-agent learning based on different color spaces. This stage aimed to create a decision matrix of various color spaces, and three groups of criteria (i.e., reliability, time complexity, and error rate within dataset) to test, evaluate and benchmark the adapted skin detection approaches. (2) Performance of multiple evaluation criteria for skin detection engines, this stage included two key stages. First, the correlation between criteria to investigate their relationship and determine their degree of correlation. Second, the performance analysis of criteria to identify the factors that affect the behavior of each criterion. The second experiment utilized a new multi-criteria decision-making by adopting the integration of TOPSIS and AHP to benchmark the results of skin detection approaches. In the validation process, multi-criteria measurement was used to calculate the trade-off for different criteria. Color spaces assessment were conducted to determine the best color spaces with adaptive skin detection engines. Moreover, mean and standard deviation values for thresholds were calculated to select the best color space. Two groups of findings were provided. First, the overall comparison of external and internal aggregation values in selecting the best color space, that is the norm RGB at the sixth threshold. Second, (1) the process proves that the distribution of color spaces with its threshold values affects the behavior of the criteria determined as a trade-off between the criteria according to their weight distribution. (2) The YIQ color space obtains the lowest value and is the worst case, whereas the norm RGB color space receives the highest value and is the most recommended. (3) The best result achieved at the threshold = 0.9. Thus, the implications of this study benefit individuals, research centers, and organizations interested in skin detection applications. Moreover, it provides benefits to software developers working in industrial companies and institutions in developing different techniques and algorithms with different applications.

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

METODOLOGI BARU UNTUK PENILAIAN DAN PENYELESAIAN DETEKSI KULIT BERDASARKAN MODEL AI MENGGUNAKAN ANALISIS KRITERIA MULTI

ABSTRAK

Kajian ini bertujuan untuk membangunkan metodologi baharu bagi menilai dan menanda aras pengesanan kulit berdasarkan model kecedasan buatan menggunakan analisis pelbagai kriteria. Untuk tujuan ini, dua eksperimen telah dijalankan. Eksperimen pertama terdiri daripada dua peringkat: (1) Adaptasi kes terbaik terdahulu dalam mengesan kulit menggunakan pendekatan multi-agen berdasarkan ruang warna yang berbeza. Peringkat ini bertujuan untuk membuat matriks keputusan pelbagai ruang warna dan tiga kumpulan kriteria (iaitu, kebolehpercayaan, kerumitan masa, dan kadar kesilapan dalam set data) untuk menilai dan menanda aras pendekatan pengesanan kulit yang telah disesuaikan. (2) Prestasi kriteria pelbagai penilaian bagi enjin pengesanan kulit, di mana peringkat ini melibatkan dua peringkat kekunci. Pertama, korelasi antara kriteria untuk menyiasat hubungan dan menentukan darjah korelasi. Kedua, analisis prestasi kriteria untuk mengenal pasti faktor kriteria yang mempengaruhi kelakuan setiap kriteria. Eksperimen kedua menggunakan pendekatan membuat-keputusan multi-kriteria baharu melalui integrasi antara TOPSIS dan AHP untuk menanda aras keputusan pendekatan pengesanan kulit. Di dalam proses pengesahan, pengukuran pelbagai kriteria digunakan untuk mengira keseimbangan bagi pelbagai kriteria. Penilaian ruang warna dijalankan untuk menentukan ruang warna yang terbaik dengan enjin pengesanan kulit yang telah diadaptasi. Seterusnya, nilai min dan sisihan piawai dikira untuk memilih ruang warna yang terbaik. Hasil dapatan daripada dua kumpulan adalah seperti berikut. Pertama, perbandingan keseluruhan nilai agregasi luaran dan dalaman dalam memilih ruang warna terbaik, iaitu RGB norma pada ambang keenam. Kedua, (1) proses membuktikan bahawa penagihan ruang warna dengan nilai ambangnya mempengaruhi kelakuan kriteria yang ditentukan sebagai keseimbangan antara kriteria berpandukan pengagihan berat masing-masing. (2) Ruang warna YIQ memperoleh nilai terendah dan merupakan kes terburuk, manakala ruang warna norm-RGB memperoleh nilai tertinggi dan paling disyorkan. (3) Dapatan terbaik dicapai pada ambang = 0.9. Oleh itu, implikasi kajian ini memberi manfaat kepada individu, pusat penyelidikan dan organisasi yang berminat dalam aplikasi pengesanan kulit. Kajian ini turut memberi manfaat kepada pembangun perisian yang bekerja di industri dan institusi dalam membangunkan teknik dan algoritma yang berbeza bagi aplikasi yang berbeza.

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

TABLE OF CONTENTS

INSTITUTE OF GRADUATE STUDIES DECLARATION OF ORIGINAL WORK i

ACKNOWLEDGEMENT ii

ABSTRACT xvi

ABSTRAK iv

TABLE OF CONTENTS v

LIST OF TABLES xiii

LIST OF FIGURES xvi

LIST OF ABBREVIATION xix

CHAPTER 1 INTRODUCTION

1.1 Introduction 1

1.2 Research Background 2

1.3 Significance of Study 5

1.4 Problem Statement 6

1.5 Research Scope 11

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

1.6 Research Objectives 12

1.7 General View and Scope of the Research 13

1.8 Organization of Thesis 15

1.9 Chapter Summary 17

CHAPTER 2 LITERATURE REVIEW

2.1 Introduction 19

2.2 Evaluation and Benchmarking for Skin Detector Approaches 21

2.2.1 Criteria of Evaluation 21

2.2.1.1 Reliability Group 25

2.2.1.1.1 Matrix of Parameters (MP) Section 26

2.2.1.1.2 Relationship of Parameters (RP) Section 28

2.2.1.1.3 Behaviour of Parameters (BP) Section 32

2.2.1.1.4 Summary of Relationships among Reliability Grou 34

2.2.1.2 Time Complexity Group 42

2.2.1.3 Error Rate within Dataset Group 48

2.2.1.3.1 Cross Validation Pattern 49

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

2.2.1.3.2 Training Pattern 51

2.3 Benchmarking Techniques/ Tools 57

2.3.1 Data Mining Tools Group 58

2.3.2 Computer Vision Tools Group 63

2.4 Open Issues and Challenge for Evaluation and Benchmarking Process 71

2.4.1 Concern for Evaluation Criteria 72

2.4.2 Concern for Criteria Trade-off 73

2.4.3 Concern for Criteria Importance 75

2.5 Theoretical Background about Multi Criteria Decision Making techniques 77

2.5.1 Analytical study of MCDM Techniques 79

2.5.2 Analytic Hierarchy Process (AHP) Method 83

2.5.3 Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) method 86

2.6 Survay of Skin Detection Models 87

2.6.1 Parametric Skin Modelling 88

2.6.1.1 Neural Network Method 90

2.6.2 Non-Parametric Skin Modelling 92

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

2.5.2.1 Naive Bayes Classifier 94

2.5.3 Why Selected the Case Study? 96

2.6 Chapter Summary 100

CHAPTER 3 METHODOLOGY AND DESIGN OF EXPERIMENTS

3.1 Introduction 102

3.2 Preliminary Phase 104

3.3 Identification and Performance Phase 104

3.3.1 Identification of the Decision Matrix 104

3.3.1.1 Development Skin Detector using Multi-Agent Learning based on AI models using different Color Spaces 105

3.3.1.1.1 Multi-Agent Learning Technique 105

3.3.1.1.2 Color Space Adapted 106

3.3.1.1.3 Training operation of Neural Network Model 115

3.3.1.1.4 Training operation of the Bayesian Model 116

3.3.1.1.5 Detection Step of the Skin Detector 119

3.3.1.2 Crossing between Developed Skin Detector and Different Criteria 123

3.3.1.2.1 Procedure for Computation Reliability Group Elements 125

3.3.1.2.2 Procedure Computation for Time Complexity Criterion 128

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

3.3.1.2.3 Error Rate Computation within Dataset Elements 130

3.3.1.3 Evaluation and Testing Skin Detector Based on Three Groups Criteria 130

3.3.2 Performance of Decision Matrix 131

3.3.2.1 Correlation between Criteria 132

3.3.2.2 Performance Analysis of Criteria 133

3.4 Development Phase 134

3.4.1 Development of Decision-making Solution for Skin Detection Approach Based on Integrated ML-AHP&TOPSIS 135

3.4.2 Adaptation of ML-AHP Technique for Weight Investigation of Different Evaluators 137

3.4.2.1 Pairwise Comparisons for Each Criterion 138

3.4.2.2 Design of the ML-AHP measurement Structure 141

3.4.2.3 Weight Calculation of Criteria and Validation of Consistency Value 144

3.4.3 Utilization of the TOPSIS Method for Skin Detection Evaluation and Benchmarking 145

3.4.3.1 Decision Making Context 148

3.5 Validation Phase 151

3.5.1 Validity of the Multi Criteria Measurement Process 151

3.5.2 Comparison between Color Spaces 152

3.5.3 Statistical Measurement for Color Spaces 152

3.6 Chapter Summary 153

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

CHAPTER 4 MULTI CRITERIA ANALYSIS AND COMPARISON

4.1 Background 155

4.2 Results of the Proposed Decision Matrix 156

4.3 Correlation Coefficient 161

4.3.1 Correlation Measurement of the Criteria 161

4.3.1.1 Correlation Analysis in Layer 1 163

4.3.1.2 Correlation Analysis in Layer 2 164

4.3.1.3 Correlation Analysis in Layer 3 167

4.3.1.4 Summery 171

4.4 Performance Analysis of Criteria 171

4.4.1 Reliability Group 172

4.4.1.1 Matrix of Parameters 172

4.4.1.2 Relationshep of Parameters 179

4.4.1.3 Behavior of Parameters 186

4.4.2 Time Complexity Criterion 189

4.4.3 Error Rate within Dataset 191

4.4.4 Summary 193

4.5 Chapter Summery 194

CHAPTER 5 RESULTS AND DISCUSSION

5.1 Introduction 196

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

5.2 Multi-Layer Weight Measurement using AHP 197

5.3 TOPSIS Performance based on Different Evaluator's Weights 199

5.4 Group TOPSIS with Internal and External Aggregation 214

5.5 Chapter Summary 217

CHAPTER 6 VALIDATION AND COMPARISON

6.1 Introduction 219

6.2 Validity of the Multi Criteria Measurement Process 220

6.2.1 Discussion and Evaluation Trade-off for Multi Criteria Measurment 221

6.2.1.1 Scenario Tradeoff for the Reliability Group 222

6.2.1.1.1 Employment of Paired Sample Test of Scenario Reliability Group 224

6.2.1.2 Scenario Tradeoff of the Time Complexity Group 226

6.2.1.2.1 Employment of Paired SampleTest of Scenario Time Complexity Group 228

6.2.1.3 Scenario Tradeoff of Error Rate Group 230

6.2.1.3.1 Employment of Paired Sample Test of Scenario Error . Rate Group .232

6.2.1.4 Summary of Scenarios 234

6.3 Color Spaces Measurement 235

6.4 Threshold Measurements 237

6.5 Chapter Summary 239

CHAPTER 7 CONCLUSIONS AND FUTURE WORK

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

7.1 Introduction 240

7.2 Conclusions 241

7.3 Research Limitation and Issues 242

7.4 Future work 244

REFERENCES **247**

LIST OF PUBLICATION **271**

APPENDIXES

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

LIST OF TABLES

Table No.		Page
2.1	Literature review of evaluation criteria.	23
2.2	Confusion Matrix	27
2.3	Reliability Group for Skin Detection Approach	38
2.4	Time Complexity group for Skin Detection Approaches	44
2.5	Error Rate within Dataset Group for Skin Detection Approach	52
2.6	Summary of Weaknesses of the Tools	67
2.7	Trade-off Problem in the Academic Literature	75
2.8	Example of Multi-Criteria problem	79
2.9	Literature Survey for Various studies in Skin Detection Domain	96
3.1	Establishment of the Decision Matrix	124
3.2	Sample Pairwise Comparison Matrix	139
3.3	Intensity Scale of Criteria	141
3.4	Random Index (Saaty,T.L and Ozdemir,M.S.2003)	145
4.1	Implementation of the Decision Matrix	158
4.2	Comparison of Reliability, Time Complexity, and Error Rate Criteria	163
4.3	Comparison among Matrix of Parameters, Relationship of Parameters, and Behavior of Parameter Sub-Criteria	165
4.4	Comparison Training and Validation Sub-Criteria	166
4.5	Comparison among TP, FP, TN, and FN Sub-Sub-Criteria	167
4.6	Comparison of Accuracy, Precision, Recall and Specificity Sub-Sub -Criteria	169

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

4.7 Comparison between F-measure and G-measure Sub-Sub-Criteria 170

5.1 ML-AHP measurement for Weights Preferences 199

5.2 First Evaluator Result to Evaluate and Benchmark for Different Color Space Algorithm 201

5.3 Second Evaluator Result to Evaluate and Benchmark for Different Color Space Algorithms 203

5.4 Third Evaluator Result to Evaluate and Benchmark for Different Color Space Algorithms 205

5.5 Fourth Evaluator Result to Evaluate and Benchmark for Different Color Space Algorithms 207

5.6 Fifth Evaluator Result to Evaluate and Benchmark for Different Color Space Algorithms 209

5.7 Sixth evaluator Result to Evaluate and Benchmark for Different Color Space Algorithms 211

5.8 Group Decision-maker of TOPSIS method with Internal and External Aggregations 215

6.1 Implementing Numerical Sequence Process for Reliability Criterion 222

6.2 P-value for Different Weights of the Reliability Criterion 224

6.3 Results of the Correlation between Different Weights for Reliability Criterion 225

6.4 Implementing Numerical Sequence Process for Time Complexity Criterion 226

6.5 P-value for Different Weights of the Time Complexity Criterion 228

6.6 Results of the Correlation between Different Weights for Time Complexity Criterion 229

6.7 Implementation of Multi Criteria Measurements in Error Rate Criterion 230

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

6.8 P-value for Different Weights of the Error Rate Criterion 232

6.9 Results of the Correlation between Different Weights for Error Rate Criterion 233

6.10 Mean and Stander Division for Threshold values 238

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

LIST OF FIGURES

No. Figures **page**

1.1 Significance of Study 6

1.2 Research Problem and Gap 8

1.3 Magic triangle of skin detection requirement 9

1.4 General View of the Research 14

2.1 Taxonomy for Chapter Direction of the Research 20

2.2 Sections Reliability Group 37

2.3 Time Complexity process for Skin Segmentation 44

2.4 Presentation of the Example in Table (2.8) 79

2.5 Initial Decision in the Hierarchical Structure 84

3.1 Research Methodology of Design Phases 103

3.2 Multi-Agent Learning of Skin Detection 106

3.3 Development of Case Study using Different Color Spaces 114

3.4 Training Process of the Bayesian Model (Zaidan,A.A.et al.2014b) 117

3.5 Skin Segmentation and Detection Processes 120

3.6 Matching Process for Different Objects 127

3.7 Procedure of Time Complexity 129

3.8 New Methodology for Skin Detector 134

3.9 Integration of ML-AHP and TOPSIS methods for Skin Detection Approaches 136

3.10 AHP method based on Multi-Layer Structure 138

3.11 Pairwise Answer from Evaluators 140

3.12 ML-AHP Steps Used to Account for Multi-layer Matrix 143

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

3.13 Group Decision Maker Process 149

3.14 Individual Decision Maker Process 150

4.1 Overview of the Results and Evaluation of Different Criteria 156

4.2 Taxonomy of Criteria Distribution into Three Layers 162

4.3 Behavior of True Negative Criterion with Different Colors 173

4.4 Behavior of the True Positive Criterion with Different Color Spaces 174

4.5 Behavior of the False Positive Criterion with Different Color Spaces 176

4.6 Behavior of the False Negative Criterion with Different Color Spaces 178

4.7 Behavior of the Accuracy Criterion with Different Color Spaces 180

4.8 Behavior of the Recall Criterion with Different Color Spaces 182

4.9 Behavior of the Precision Criterion with Different Color Spaces 184

4.10 Behavior of the Specificity Criterion with Different Color Spaces 185

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

4.11 Behavior of the F-measure Criterion with Different Color Spaces 187

4.12 Behavior of the G-measure Criterion with Different Color Spaces 189

4.13 Behavior of the Time Complexity Criterion with Different Color Spaces 190

4.14 Behavior of the Error Rate of Validation Criterion with Different Color Spaces 191

4.15 Behavior of the Error Rate of Training Criterion with Different Color Spaces 192

5.1 Overview of Results and Evaluation of the skin detector 197

5.2 Virtualize Ranking for Six Evaluators 213

5.3 Internal and External Aggregation Ranking 217

6.1 Overview of the Design and Implementation of the validation process 220

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

6.2 Trade-off Scenario for Reliability Group in Comparison with Other groups 223

6.3 Trade-off Scenario for Time Complexity Group in Comparison with Other groups 227

6.4 Tradeoff Scenario for Error Rate Group in Comparison with Other groups 231

6.5 Color Space Measurement 235

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

LIST OF ABBREVIATION

ANN	Artificial Neural Network
AHP	Analytic Hierarchy Process
ANP	Analytic Network Process
CPU	Central Processing Unit
CIE	Commission International de L'Eclairage
CR	Consistency Ratio
DM	Decision Matrix
EM	Evaluation Matrix
FP	False Positive
FN	False Negative
GH	Grouping Histogram
GDM	Group Decision Making
HAW	Hierarchical Adaptive Weighting
IT	Information Technology
KNIME	Konstanz Information Miner
KEEL	Knowledge Extraction based on Evolutionary Learning
LUT	Lookup Table
MCDM	Multi- Criteria Decision Making
MADM	Multi- Attribute Decision Making
MCDA	Multi-Criteria Decision Analysis

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

MEW	Multiplicative Exponential Weighting
RI	Random Index
SVM	Support Vector Machine
SAN	Segment Adjacent-Nested
SAW	Simple Additive Weighting
TP	True Positive
TN	True Negative
TOPSIS	Technique for Order Preference by Similarity to Ideal Solution
WEKA	Waikato Environment for Knowledge Analysis
WSM	Weighted Sum Model

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

CHAPTER 1

INTRODUCTION

1.1 Introduction

This chapter introduces the research direction, research background, and a statement of the problem. This chapter also presents the ambitions, motivations, and objectives of this research are also presented.

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

Section 1.2 presents a brief background of the research components. Section 1.3 introduces the statement of the problem, which is the basis of the research direction. Section 1.4 discusses the scope of the research. Section 1.5 describes the research objectives. Section 1.6 presents a general view of the research. Finally, Section 1.7 briefly outlines the main structure of the research.

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

1.2 Research Background

Decades ago, the skin detection approach has been considered an important platform for various fields, such as medical and several scientific disciplines (L. Huang et al. 2015). In other words, skin detection has gained an important function in a wide range of image or video processes for various applications. A few factors that directly impact skin appearance include illumination, background, camera characteristics, and ethnicity (Kakumanu, Makrogiannis, and Bourbakis 2007). Elgammal, Muang, and Hu (2009) defined the skin detection approach as a process of finding skin-colored pixels and regions in an image or a video into a specific region. This process is typically used as a preprocessing step to finding regions in images that potentially detect the human face and limbs. The skin detection approach includes various applications, such as face detection, (Zhipeng, C., Junda, H., & Wenbin 2010), face tracking (Tsai 2012), gesture analysis (Hussain, I., Talukdar, A. K., & Sarma 2014), Internet pornographic image filtering (Lee, Kuo, and Chung 2010), surveillance systems (Zui Zhang 2009), content-based image retrieval systems (Patil, C. G., Kolte, M. T., Chatur, P. N., & Chaudhari 2014), and various human–computer interaction domains (Hollender et al. 2010). The most practical and effective techniques are used in developing skin detector artificial intelligence (AI) algorithms according to the literature on skin detection for skin pixel and non-skin pixel features based on color features. On the contrary, many researchers have applied hybrid algorithms in AI models (Singh Sisodia and Verma 2011; (Shruthi, M. L. J., & Harsha 2013; Zaidan et al. 2014b). However, with the current rapid development of the skin detection approach in various applications, finding an evaluation and benchmarking

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

methodology that is reliable, effective, and comprehensive has become critical (Jones and Rehg 1999; Phung, Bouzerdoum, and Chai, D. 2005; Gamage, Akmeliawati, and Chow 2009; Taqa and Jalab 2010a).

Considering the basic criteria evaluation of reliability, time complexity, and error rate within the dataset in the design of any skin detector application, (Jones and Rehg (1999) adapted three criteria, namely reliability, computational cost, and error rate of skin detection. In one of the earliest works that highlight the problem of skin detection evaluation and benchmarking, three general requirements for the skin detection approach are reported: adapted reliability (i.e., the obtained skin detection rate and false positives) and datasets (i.e., the obtained equal error rate comparison of AI models) with less time-consuming requirements to process web images.

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

Despite the importance of the remaining criteria, Phung, Bouzerdoum, and Chai, D. (2005) highlighted the dataset criterion by comparing two algorithms. The dataset is represented by training and testing for skin and non-skin pixels for skin segmentation images. However, the output images created through a classifier are compared pixel-wise with the ground truth of skin segmentation. Gamage, Akmeliawati, and Chow (2009) reported a skin detection algorithm that has been tested with images through independent databases. They investigated the size of the image, which has a significant impact on time complexity. Thus, they proved that increasing image size leads to low accuracy than increase time complexity of the experiment. Finally, Taqa and Jalab (2010a) stated that reliability is a prerequisite for

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

skin detection evaluation. They highlighted a reliability criterion based on accuracy, precision, and recall of the image color despite the importance of the remaining criteria. However, the quality assessment of skin detection requires attention.

Consequently, two key problems are encountered by skin detection developers. One is the evaluation of skin detection approaches based on the abovementioned evaluation criteria and benchmark new skin detection approach versus existing approaches. Therefore, the evaluation and benchmarking process need to consider these requirements. Despite the tradeoff among various criteria, (Jones and Rehg (1999); Phung, Bouzerdoum, and Chai, D. (2005); Gamage, Akmeliawati, and Chow (2009); Taqa and Jalab (2010a) have adopted each of the proposed criteria. They attempted to evaluate the reliability criterion for a given time complexity based on different datasets. However, the term “reliability” is unclearly defined in the literature. According to the preceding studies mentioned, the percentage of reliability varies depending on different adapted algorithms and thus exhibit an inconsistent level. Meanwhile, Fernandes, Cavalcanti, and Ren (2013) reported time complexity variation between the algorithms, which depend on the CPU time. Consequently, the processing time of an image is affected, but this aspect is excluded in the scope of the present research. Therefore, the calculation should be the highest percentage of reliability compared with the lowest time complexity of the output image. Kawulok (2013) mentioned that the dataset can be divided into two classes, namely training and validation data, to find the minimum detection error. In general, all these studies have

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

proven the evaluation and benchmarking process of each of these criteria based on independent guidelines.

Therefore, conducting further investigations and developing a clear methodology for testing, evaluation, and benchmarking are necessary to standardize basic and advanced requirements for the skin detection approach. Redefining the problem of evaluation and benchmarking need is also necessary. Moreover, the new evaluation methodology must be flexible to handle the conflicting criteria problem and must have the capability to maintain the current criteria.

1.3 Significance of Study

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

The evaluation and benchmarking of skin detection approaches are important areas for many researchers and organizations interested in their applications. Many individuals and organizations are interested in the applications of skin detection approaches, such as researchers working in scientific research centers, developers working in industrial companies and institutions, and graduate students enrolled in schools that develop various applications of skin detection approaches. Thus, the importance of the study is the development of multiple applications of skin detection approaches, including face detection, face tracking, gesture analysis, Internet pornographic image filtering, surveillance systems, content-based image retrieval systems, and various human-computer interaction domains. Moreover, this study

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi