PRE-PROCESSING STRATEGIES FOR SKIN DETECTION USING MLP

CHELSIA AMY DOUKIM

SCHOOL OF ENGINEERING AND
INFORMATION TECHNOLOGY
UNIVERSITI MALAYSIA SABAH
2011

PRE-PROCESSING STRATEGIES FOR SKIN DETECTION USING MLP

CHELSIA AMY DOUKIM

THESIS SUBMITTED IN FULFILLMENT FOR THE DEGREE OF MASTER OF ENGINEERING

**SCHOOL OF ENGINEERING AND INFORMATION TECHNOLOGY
UNIVERSITI MALAYSIA SABAH**

2011

DECLARATION

I hereby declare that the material in this thesis is my own except for quotations, excerpts, equations, summaries and references, which have been dully acknowledged.

29 NOVEMBER 2010

CHELSIA AMY DOUKIM
PK20068608

CERTIFICATION

NAME	:	**CHELSIA AMY DOUKIM**
MATRIC NO.	:	**PK20068608**
TITLE	:	**PRE-PROCESSING STRATEGIES FOR SKIN DETECTION USING MLP**
DEGREE	:	**MASTER OF ENGINEERING (ARTIFICIAL INTELLIGENCE)**
VIVA DATE	:	**29 NOVEMBER 2010**

DECLARED BY

1. SUPERVISOR

Dr. Jamal Ahmad Dargham — Signature

2. CO-SUPERVISOR

Assoc. Prof. Dr. Ali Chekima — Signature

ACKNOWLEDGEMENT

First and foremost, thanks to God for giving me the strength to complete this thesis.

I would like to acknowledge and extend my heartfelt gratitude to my supervisor, Dr. Jamal Dargham of the School of Engineering and Information Technology, Universiti Malaysia Sabah for his vital advice and encouragement throughout the past few years. I would also like to thank him for providing the necessary thrust upon completion of this research and publish this thesis.

I wish to express my special sense of gratitude to my co-supervisor Associate Professor Dr. Ali Chekima of the School of Engineering and Information Technology, Universiti Malaysia Sabah for his constant advices and motivation.

A trillion thanks to my family especially my parents, for their unconditional love, patience, encouragement and understanding, and of course for being the greatest financial sponsorship. Thanks to my siblings and cousins for their supports and constant amusement.

Thanks to Pn. Victoria Asoi of PGA-SKTM for her continuous assistance. Thanks to my fellow scholars and undergraduates in the Computer Engineering Program.

Finally, thanks to my dear friends for their encouragement and words of wisdom.

Chelsia Amy Doukim
29 NOVEMBER 2010

ABSTRACT

PRE-PROCESSING STRATEGIES FOR SKIN DETECTION USING MLP

Skin detection is an important preliminary step in a wide range of image processing applications such as face detection, person identification, gesture analysis and access control. Several techniques have been used for skin detection. In this thesis, the multilayer perceptron (MLP) neural network and histogram thresholding techniques were used. Recent studies have shown that combining skin features and/or skin classifiers can further improve the performance of the skin detection system. Thus, the main objective of this research is to evaluate the effect of several combination strategies on the performance of a skin detection system based on the MLP. To achieve this goal, first the histogram thresholding technique was used to select skin features (chrominance component in a given colour space) that give the highest correct skin detection. These features will be used as inputs to the MLP classifiers. A modified Growing algorithm for finding the number of neurons in the hidden layer of a neural network was also developed it was able to reduce the computational time compared to the conventional Growing algorithm. The combination strategies were done by combining the skin features as well as the skin classifiers. Three skin features (chrominance component from the selected colour space) that gave the highest correct skin detection on a single input MLP classifier were used for these strategies. The strategy of combining skin features or inputs was done using two and three skin features. For combining skin classifiers strategy, several combining rules such as binary operators AND and OR were used to combine two and three classifiers, while combining rules namely Voting, Sum of Weights and New Neural Network were used to combine three classifiers. The Sum of Weights and New Neural Network were the proposed combining rules in this thesis. In order to evaluate the performances of the skin detection systems, the images from Compaq database were used. The strategy of combining two skin features C_b/C_r gave the best performance for combining skin feature strategy with 3.01% more correct detection compared with the best performance given by a single input MLP classifier given by C_b-C_r. The strategy of combining three classifiers using the Sum of Weights gave the best performance for its combining strategy with an improvement of 4.38% more correct detection compared to the best single input MLP classifier given by C_b-C_r. The Sum of Weights strategy also gave 1.37% more correct detection than the best combining skin feature strategy. The other proposed combining strategy called New Neural Network has managed to achieve 82.21% of correct detection. The best performance results obtained in this thesis were considerably good considering the unconstrained nature of the images from the Compaq database.

ABSTRAK

Sistem pengesanan kulit merupakan satu proses utama yang penting dalam aplikasi pemprosesan imej seperti sistem pengesanan wajah, pengenalan diri, penganalisis isyarat dan kawalan akses. Beberapa teknik pengesanan kulit telah digunakan, namun demikian rangkaian neural jenis "multilayer perceptron" (MLP) dan kaedah histogram dengan teknik "threshold" telah digunakan dalam tesis ini. Sejak kebelakangan ini, kajian melaporkan bahawa strategi penggabungan seperti menggabungkan sifat kulit atau sistem pengesanan kulit yang berbeza berupaya meningkatkan keberkesanan sesuatu sistem pengesanan kulit. Oleh yang demikian, objektif utama kajian ini adalah untuk mengkaji keberkesanan beberapa strategi penggabungan berasaskan rangkaian neural jenis MLP. Kaedah histogram dengan teknik "threshold" telah digunakan untuk menentukan sifat kulit (komponen "chrominance" dalam sistem warna yang diberikan) yang memberikan peratusan tertinggi untuk pengesanan kulit yang tepat. Sifat kulit tersebut akan digunakan sebagai input kepada sistem pengesanan kulit berasaskan MLP. Satu modifikasi algoritma yang berasaskan algoritma "Growing" untuk menentukan bilangan neuron dalam lapisan tersembunyi "hidden layer" rangkaian neural turut diperkenalkan dan ianya berkesan terutamanya dalam konteks penjimatan masa berbanding algoritma yang lazim digunakan iaitu algoritma "Growing". Dalam kajian ini, strategi penggabungan merangkumi strategi menggabungkan sifat kulit dan sistem pengesanan kulit yang berbeza. Tiga sifat kulit (komponen "chrominance" daripada sistem warna yang dipilih) yang memberikan keputusan terbaik dalam sistem pengesan kulitnya dipilih untuk strategi penggabungan yang dicadangkan. Strategi menggabungkan sifat kulit dilakukan dengan menggabungkan dua dan tiga sifat kulit. Manakala untuk strategi menggabungkan sistem pengesanan kulit, operator AND dan OR digunakan untuk menggabungkan dua dan tiga sistem pengesan kulit, sementara penggabung "Voting", "Sum of Weights" dan "New Neural Network" digunakan untuk menggabungkan tiga sistem pengesan kulit. Strategi "Sum of Weights" dan "New Neural Network" adalah dua strategi penggabungan yang baru dicadangkan dalam tesis ini. Pangkalan data imej yang dikenali sebagai "Compaq database" telah digunakan untuk menilai kecekapan sistem pengesan kulit yang dicadangkan. Untuk strategi menggabungkan sifat kulit, penggabungan dua sifat kulit (C_b/C_r & C_r) memberikan peratusan pengesanan kulit tepat yang tertinggi iaitu 82.61% dengan peningkatan sebanyak 3.01% lebih pengesanan kulit berbanding sistem pengesanan kulit terbaik yang diberikan oleh C_b-C_r untuk sistem pengesan kulit yang hanya menggunakan satu sifat kulit. Manakala, strategi menggabungkan sistem pengesan kulit menggunakan penggabung "Sum of Weights" memberikan prestasi terbaik, iaitu 83.98% dengan peningkatan sebanyak 4.38% lebih pengesanan kulit berbanding sistem pengesan kulit C_b-C_r. Strategi tersebut juga memberikan 1.37% lebih pengesanan kulit berbanding strategi menggabungkan dua sifat kulit. Strategi baru "New Neural Network" berupaya mengesan 82.21% kulit dengan tepat. Hasil yang diperolehi dalam tesis ini boleh dikatakan amat memberangsangkan memandangkan sifat-sifat gambar yang sukar diklasifikasikan daripada "Compaq database".

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

CCD	Charge Coupled Device
CDR	Correct Detection Rate
CIE	Commission Internationale de l'Eclairage
CIF	Common Intermediate Format
EM	Expectation-Maximization
FAR	False Acceptance Rate
FRR	False Rejection Rate
GMM	Gaussian Mixture Model
GW	Gray World
LUT	Lookup Table Method
MAP	Maximum a Posteriori
ML	Maximum Likelihood
MLP	Multilayer Perceptron
MSE	Mean Squared Error
NTSC	National Television System Committee
PAL	Phase Alternating Line
PDF	Probability Distribution Function
RBF	Radial Basis Functions
SCNS	Skin Corrected by a Non-Skin
SGM	Single Gaussian Model
SOM	Self-Organizing Map
SPM	Skin Probability Map
STCB	Selftunable Colour Balancing
SVM	Support Vector Machine
TAR	True Acceptance Rate
TDSD	Test Database for Skin Detection

WP	White Patch
WPt	White Point
WWW	World Wide Web

LIST OF SYMBOLS

$\sum$	Diagonal covariance matrix
$b_{x,y}$	Backprojected image at (x,y) coordinates
B	The blue chrominance component of the RGB colour space
c	Colour vector
C	A given chrominance component
c_j	Colour sample
$C_{x,y}$	Colour value at (x,y) coordinates
C1	Class 1
C2	Class 2
C_b	The blue chrominance of the YC_bC_r colour space
$C_b.C_r$	The blue chrominance multiplied with the red chrominance of YC_bC_r colour space
C_b/C_r	The ratio of the blue and red chrominance components of YC_bC_r colour space
C_b+C_r	The blue chrominance added to the red chrominance of YC_bC_r colour space
C_b-C_r	The blue chrominance minus the red chrominance of YC_bC_r colour space
CDR	Correct detection rate
C_r	The red chrominance of the YC_bC_r colour space
EI	Error index
G	The green chrominance component of the RGB colour space
G/B	The ratio of green and blue chrominance components of the RGB colour space
H	Hue
$h(C_{x,y})$	Histogram bin corresponding to $C_{x,y}$
HN	Number of neurons in the hidden layer of a neural network

HN_B	The number of hidden neurons obtained from binary search method
HN_{BH}	The nearest higher value HN_B
HN_{BL}	The nearest lower value HN_B
I	Intensity
i	Index bin
M	Model histogram
M(x,y)	Pixel at the (x,y) Coordinates in the masked image
$M_{h(C_{x,y})}$	Model histogram with bin corresponding to $C_{x,y}$
N	Number of Gaussians
NC1	Number of pixels in class 1
NC2	Number of pixels in class 2
NCC1	Number of pixels in class 1 correctly classified as class 1
NCC2	Number of pixels in class 1 correctly classified as class 2
NEC1	Number of class 1 misclassified as class 2
NEC2	Number of class 2 pixels misclassified as class 1
$N_M(x,y)$	Pixel at the (x,y) Coordinates in the modified normalised RGB Image
$N_Y(x,y)$	Pixel at the (x,y) Coordinates in the YC_bC_r image
O(x,y)	The output pixel at (x,y) coordinates
P(x,y)	The input pixel at (x,y) coordinates
PEC1	Percentage of class 1 error
PEC2	Percentage of class 2 error
PSE	Percentage of segmentation error
R	Ratio histogram
R	The red chrominance component of the RGB colour space
R_i	Ratio histogram with i index of bin
r.b	The red chrominance multiplied with the blue chrominance of the rgb colour space

r.g	The red chrominance multiplied with the green chrominance of the rgb colour space
R/B	The ratio of red and blue chrominance components of the RGB colour space
r/b	The ratio of the red and blue chrominance components of the rgb colour space
R/G	The ratio of red and green chrominance components of the RGB colour space
r/g	The ratio of the red and green chrominance components of the rgb colour space
R/G+R/B	The ratio of red and green chrominance components added to the ratio of red and blue chrominance components of the RGB colour space
r-b	The red chrominance minus the blue chrominance of the rgb colour space
r-g	The red chrominance minus the green chrominance of the rgb colour space
S	The saturation component of the TSL colour space
T	The tint component of the TSL colour space
T+S	The tint component added to the saturation component of the TSL colour space
T^C	Threshold for Chrominance Component C
T_H^C	Higher Threshold for Chrominance Component C
T_L^C	Lower Threshold for Chrominance Component C

KEYWORDS

Skin Detection, Combining Neural Networks, Multilayer Perceptron, Histogram Thresholding Technique, Feature Extraction.

CHAPTER 1

INTRODUCTION

1.1 Introduction to Skin Detection

Skin detection can be defined as the decision whether a pixel belongs to a skin region or a non-skin region based on that pixel colour value. Albiol et al. (2001) defined skin detection as the process of selecting which pixels of a given colour image corresponds to human skin. Skin detection is often used as the first step for subsequent feature extraction in a wide range of image processing applications such as face detection and recognition, face tracking, content-based image filtering, gesture analysis and person identification. Skin detection techniques can be classified into two categories namely pixel-based methods and region-based methods. Pixel-based methods classify each pixel as skin or non-skin individually based on human skin colour. Most researchers have used pixel-based methods compared to the region-based methods because skin colour information can be used for detecting human skin in various computer vision applications since skin colour allows high processing speed due to its low-level processing and is highly robust against rotations, scaling and partial occlusions.

1.2 Challenges of Skin Detection

Skin detection using the skin colour is considered a challenging task due to the sensitivity of skin appearance in images to various factors. Kakumanu et al. (2007) have identified several factors, which are:

Illumination: A change in the lighting condition produces a change in the apparent colour of the skin in the image.

Camera characteristics: Even under the same lighting conditions, the skin colour distribution for the same person differs from one camera to another depending on the camera sensor characteristics.

Ethnicity: Skin colour varies from person to another person belonging to different ethnic group and from persons across different regions.

Individual characteristics: An individual characteristic such as age, sex and body parts also affect the skin colour appearance.

Other factors: Different factors such as subject appearance (make-up, hairstyle and glasses), background colours, shadows and motion also influence the appearance of the skin colour.

1.3 State of the Art in Skin Detection

Phung et al. (2001) categorized the techniques used for skin detection according to the way the skin colour distribution is modelled into parametric, non-parametric and semi-parametric. The example of a parametric model is Gaussian classifier. Gaussian classifier is a classifier based on normal distribution. In probability theory, the normal (or Gaussian) distribution is a continuous probability distribution that is often used as a first approximation to describe real-valued random variables that tend to cluster around a single mean value. It has the ability to generalize data well even with a small size of training data and it requires small storage space. However, parametric modelling techniques are affected by the colour space representation and by the amount and the quality of the training data available. The non-parametric methods such as histogram-based methods are not affected by the choice of colour space and it is also fast in training and is independent of the shape of the skin distributions (Kakumanu et al., 2007). Nevertheless, due to its incapability to interpolate data, this method requires a very large training dataset in order to obtain a good classification rate. An example of semi-parametric model is the Multilayer Perceptron (MLP) neural network. Multilayer Perceptron is an example of an artificial neural network that is used extensively for the solution of a number of different problems, including pattern recognition and interpolation. The Multilayer Perceptron is able to learn complex non-linear input-output relationships as well as to generalize to any given data. However, the performance of the network is dependent on the network properties such as the number of hidden layers, the number of neurons in the hidden layer and the learning rates. Despite