การหาตำแหน่งของลักษณะเด่นบนใบหน้าโดยปราศจากการพิจารณาการปรากฏของเนื้อภาพ

นายศุภกานต์ พิมลธเรศ

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรดุษฎีบัณฑิต

สาขาวิชาวิทยาการคอมพิวเตอร์ ภาควิชาคณิตศาสตร์

คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2547

ISBN 974-17-6472-3

FACIAL FEATURE LOCALIZATION WITHOUT CONSIDERING THE APPEARANCE OF IMAGE CONTEXT

Mr.Suphakant Phimoltares

A Dissertation Submitted in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy in Computer Science

Department of Mathematics
Faculty of Science
Chulalongkorn University

Academic year 2004
ISBN 974-17-6472-3

Thesis Title	FACIAL FEATURE LOCALIZATION WITHOUT CONSIDERING THE APPEARANCE OF IMAGE CONTEXT
By	Mr. Suphakant Phimoltares
Field of study	Computer Science
Thesis Advisor	Professor Chidchanok Lursinsap, Ph.D.
Thesis Co-advisor	Associate Professor Kosin Chamnongthai, Ph.D.

Accepted by the Faculty of Science, Chulalongkorn University in Partial Fulfillment of the Requirements for the Doctor's Degree

.... Deputy Dean for Administrative Affairs
Acting Dean, The Faculty of Science
(Associate Professor Tharapong Vitidsant, Ph.D.)

THESIS COMMITTEE

.... Chairman
(Associate Professor Peraphon Sophatsathit, Ph.D.)

.... Thesis Advisor
(Professor Chidchanok Lursinsap, Ph.D.)

.... Thesis Co-advisor
(Associate Professor Kosin Chamnongthai, Ph.D.)

.... Member
(Associate Professor Boonserm Kijsirikul, Ph.D.)

.... Member
(Assistant Professor Rajalida Lipikorn, Ph.D.)

.... Member
(Pirawat Watanapongse, Ph.D.)

นายศุภกานต์ พิมลธเรศ : การหาตำแหน่งของลักษณะเด่นบนใบหน้าโดยปราศจากการพิจารณาการปรากฏของเนื้อภาพ. (FACIAL FEATURE LOCALIZATION WITHOUT CONSIDERING THE APPEARANCE OF IMAGE CONTEXT) อ. ที่ปรึกษา : ศาสตราจารย์ ดร. ชิดชนก เหลือสินทรัพย์, อ.ที่ปรึกษาร่วม : รองศาสตราจารย์ ดร. โกสินทร์ จำนงไทย จำนวนหน้า 92 หน้า. ISBN 974-17-6472-3.

การหาตำแหน่งของลักษณะเด่นบนใบหน้ามีบทบาทสำคัญในการประยุกต์ใช้งานต่างๆ เช่น การโต้ตอบของคนและคอมพิวเตอร์ ความระวังระไวทางภาพ (video surveillance) การติดตามใบหน้า และ การรู้จำใบหน้า ดังนั้นขั้นตอนวิธีเพื่อการหาตำแหน่งของใบหน้าและลักษณะเด่นบนใบหน้าจึงจำเป็นสำหรับการประยุกต์ใช้งานดังกล่าว วิทยานิพนธ์นี้ได้นำเสนอขั้นตอนวิธีสำหรับภาพใบหน้าทุกชนิดในสภาวะทางภาพที่หลากหลายซึ่งประกอบไปด้วยสองขั้นตอนที่สำคัญ ขั้นตอนที่หนึ่งเป็นการหาตำแหน่งของหน้าจากภาพต้นฉบับ การหาขอบภาพแคนนี่ (Canny edge detection) ถูกนำมาใช้เพื่อหาขอบของภาพ จากนั้นจะคัดเลือกบริเวณที่น่าจะเป็นใบหน้าจากบริเวณที่มีจำนวนของจุดภาพ (pixel) สอดคล้องกับแผ่นแบบของใบหน้า (face template) หลังจากนั้นใช้ค่าความสอดคล้องที่คำนวนได้ในการหาบริเวณที่เป็นใบหน้าจริง ขั้นตอนที่สองคือการหาตำแหน่งของลักษณะเด่นบนใบหน้าซึ่งได้นำมาใช้กับใบหน้าที่ได้จากขั้นตอนแรกโดยนำเสนอแบบจำลองภาพทางระบบประสาท (neural visual model) ในการรู้จำตำแหน่งของลักษณะเด่นบนใบหน้าที่เป็นไปได้ทั้งหมด พารามิเตอร์ที่นำมาใช้กับแบบจำลองหาได้จากลักษณะพิเศษของใบหน้าและตำแหน่งของลักษณะเด่นบนใบหน้าที่เป็นอิสระจากข้อมูลของความเข้มสี หลังจากนั้นได้นำเอาวิธีการประมวลผลภาพที่เรียกว่าการพองภาพ (image dilation) มาใช้เพื่อเอาส่วนที่ไม่เกี่ยวข้องที่หลงเหลือออกเพื่อให้ได้ผลที่ดีขึ้น นอกจากนั้นแล้วขั้นตอนวิธีดังกล่าวยังสามารถนำไปใช้กับปัญหาการไม่แปรเปลี่ยนแบบหมุน (rotational invariance) โดยใช้การแปลงเรดอน (Radon transform) เพื่อหามุมหลักของใบหน้า ขั้นตอนวิธีดังกล่าวได้นำมาใช้ทดสอบกับภาพใบหน้าหลายชนิดกว่า 1,000 ภาพเช่น ใบหน้าสี, ใบหน้าเทา, ใบหน้าขาวดำ, ใบหน้าที่มีการสวมแว่นตากันแดด, ใบหน้าที่มีการสวมผ้าพันคอ, ใบหน้าที่มีการแสดงสีหน้า, ใบหน้าที่มีผลกระทบจากแสง, ใบหน้าพร่ามัวหรือมีสัญญาณรบกวน, ใบหน้าสีและใบหน้าที่เขียนแบบหวัดจากภาพเปรียบเปรย (animated cartoon) ด้วยขั้นตอนวิธีดังกล่าว อัตราการค้นหา (detection rate) มีมากกว่าร้อยละ 94 โดยเฉลี่ย

ภาควิชา คณิตศาสตร์	ลายมือชื่อนิสิต Suphakant Phimoltares
สาขาวิชา วิทยาการคอมพิวเตอร์	ลายมือชื่ออาจารย์ที่ปรึกษา
ปีการศึกษา 2547	ลายมือชื่ออาจารย์ที่ปรึกษาร่วม Kosin Chamnongthai

4373839723 : MAJOR COMPUTER SCIENCE

KEY WORD: FACIAL FEATURE LOCALIZATION, FACE DETECTION, FACE LOCATION, NEURAL VISUAL MODEL, NEURAL NETWORKS, CANNY EDGE DETECTION, IMAGE DILATION, RADON TRANSFORM

SUPHAKANT PHIMOLTARES : FACIAL FEATURE LOCALIZATION WITHOUT CONSIDERING THE APPEARANCE OF IMAGE CONTEXT. THESIS ADVISOR : PROF. CHIDCHANOK LURSINSAP, Ph.D., THESIS CO-ADVISOR : ASSOC. PROF. KOSIN CHAMNONGTHAI, Ph.D., 92 pp. ISBN 974-17-6472-3.

Facial feature detection plays an important role in various applications such as human computer interaction, video surveillance, face tracking, and face recognition. Efficient face and facial feature detection algorithms are required for applying to those tasks. This dissertation presents the algorithms for all types of face images in the presence of several image conditions. There are two main steps corresponding to face and facial feature detection algorithms. First, the faces are detected from an original image. Canny edge detection is applied to find the edge of the image. A candidate face region can be found from the region having the number of pixels corresponding to average face template. Then, the matching value is calculated and applied to find the actual face. Second, facial feature detection is applied to the actual face obtained from the previous step. A proposed neural visual model (NVM) is used to recognize all possibilities of facial feature locations. The input parameters are obtained from the face characteristics and the locations of facial features which are independent of the intensity information. For the better result, an image processing technique called dilation is applied to remove some irrelevant feature regions. In addition, the algorithms can be extended to cover rotational invariance problem by using Radon transform to extract the main angle of the face. With more than 1,000 experimental images, the algorithms are successfully tested on various types of faces with color intensity, gray intensity, binary intensity, object occlusion such as sunglasses, scarf, and hand, facial expression, lighting effect, noise and blurry images, as well as color and sketchy images from animated cartoon. In particular, the method achieves more than 94% detection rate on the average.

Department	Mathematics	Student's signature..........................
Field of study	Computer Science	Advisor's signature..........................
Academic year	2004	Co-advisor's signature..........................

Acknowledgments

I wish to thank the teachers who helped shape my scientific viewpoint and educational style. Two of them stand out. My thesis advisor, Prof. Dr. Chidchanok Lursinsap, is due a special note for his unique way of inspiring students through clarity of thought, enthusiasm, and caring. He is a model for me and many others. He provided a lot of the ideas for the dissertation. I am grateful to my thesis co-advisor, Assoc. Prof. Dr. Kosin Chamnongthai for introducing me to digital image processing and the art of research.

I would like to thank Assoc. Prof. Dr. Petr Musilek for providing some computing facilities and the pleasant atmosphere he created since I visited FACIA (the Facility for Advanced Computational Intelligence and Applications), University of Alberta, Canada.

Also, it is my pleasure to acknowledge the financial support from the Thailand Research Fund through the Royal Golden Jubilee Ph.D. Program.

It has been a pleasure to work with the AVIC members. They provided additional comments that help improve this dissertation and contribute to the environment of the AVIC.

Finally, I am deeply grateful to my parents, grandparents, brother and sister, and Miss Atchara Mahaweerawat for their love, support and patience.

Table of Contents

Chapter page

List of Tables

List of Figures

Figure page

Figure page

Figure page

List of Abbreviations

NVM	Neural Visual Model
MLP	Multilayer Perceptron
BP	Back-Propagation

List of Symbols

I_p	image p
$\widetilde{I}$	average image
$I_q^r(i,j)$	red(r) component of image I_q at position (i,j)
$I_q^g(i,j)$	green(g) component of image I_q at position (i,j)
$I_q^b(i,j)$	blue(b) component of image I_q at position (i,j)
$\widetilde{I}^r(i,j)$	red(r) component of average image $\widetilde{I}$ at position (i,j)
$\widetilde{I}^g(i,j)$	green(g) component of average image $\widetilde{I}$ at position (i,j)
$\widetilde{I}^b(i,j)$	blue(b) component of average image $\widetilde{I}$ at position (i,j)
m	image vertical length
n	image horizontal length
q	image index
Q	the number of images in the databases
γ	the number of white pixels in the edge face template
$\phi_{\mathbf{a}}$	the number of white pixels of region **a**
$\epsilon_{\mathbf{a}}$	the number of feature pixels appearing in the black holes of the black-and-white face template from candidate region **a**
$l \times l$	mean face template size
$\phi_{i,j}$	the number of white pixels in the region bounded by rows $i-l+1$ to i and columns $j-l+1$ to j
$\epsilon_{i,j}$	the number of feature pixels in the region bounded by rows $i-l+1$ to i and columns $j-l+1$ to j
$\hat{\phi}_{i,j}$	normalized $\phi_{i,j}$
$\hat{\epsilon}_{i,j}$	normalized $\epsilon_{i,j}$
$\lambda_{i,j}$	matching value in the region bounded by rows $i-l+1$ to i

	and columns $j-l+1$ to j
$width_{max} \times length_{max}$	maximum size of detected faces
(x_o, y_o)	coordinate of the center point O
(r, θ)	polar coordinate at any point (x, y) with respect to the center point
L_w	face width
L_A	distance between face center and far upper left corner of face
L_B	distance between face center and far upper right corner of face
L_C	distance between face center and far lower left corner of face
L_D	distance between face center and far lower right corner of face
(x_l, y_l)	left boundary point of face
(x_r, y_r)	right boundary point of face
$x_1, x_2, \ldots, x_m$	input signals
$w_{k1}, w_{k2}, \ldots, w_{km}$	synaptic weights connecting all input nodes to node k
u_k	linear combiner output of node k due to the input signals
b_k	bias of node k
$\psi(\cdot)$	activation function
y_k	output signal of node k
$input_j$	input signal of node j
out_i	output signal of node i

w_{ji}	synaptic weights connecting node i to node j
η	learning rate parameter
δ_k	rate of change of error with respect to the input of node k
d_k	desired output of node k
E_{msq}	mean squared error
p	index of input pattern
P	the number of all input patterns
E	objective function (error function) of the training algorithm
$E_{tr}(t)$	average error per example over the training set, measured after epoch t
$E_{va}(t)$	average error per example over the validation set, measured after epoch t
$E_{te}(t)$	average error per example over the test set, measured after epoch t
$E_{opt}(t)$	the lowest validation set error obtained in epochs up to t
$GL(t)$	generalization loss at epoch t
GL_ς	generalization loss exceeding a stopping threshold ς
ς	stopping threshold
k	length of training strip
TP_k	training progress (in per thousand) measured after k epochs
PQ_ς	quotient of generalization loss and training progress exceeding a stopping threshold ς
UP_s	notion of stopping when generalization errors increase in successive steps of strip s
s	strip index

X	set of Euclidean coordinates corresponding to the input binary image
K	set of coordinates for the kernel
Kx	translation of K so that its origin is at x
(x, y)	position of pixel in rectangular coordinate
$I_{BW}(x, y)$	black-and-white face image
$I_{Dilated}(x, y)$	dilated image
$I_{NVM}(x, y)$	detected image from NVM
$I_{Detected}(x, y)$	detected image after eliminating the irrelevant regions
ρ	smallest distance between a line and the center of image
θ	angle of a line
$I_{edge}(x, y)$	edge image
α	radius parameter for rotational invariant facial feature detection
β	face angle