



UNIVERSITY  
OF WOLLONGONG  
AUSTRALIA

University of Wollongong  
Research Online

University of Wollongong Thesis Collection  
1954-2016

University of Wollongong Thesis Collections

2010

# Posture detection by kernel PCA-based manifold learning

Peng Cheng  
*University of Wollongong*

## UNIVERSITY OF WOLLONGONG

### COPYRIGHT WARNING

You may print or download ONE copy of this document for the purpose of your own research or study. The University does not authorise you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site. You are reminded of the following:

This work is copyright. Apart from any use permitted under the Copyright Act 1968, no part of this work may be reproduced by any process, nor may any other exclusive right be exercised, without the permission of the author.

Copyright owners are entitled to take legal action against persons who infringe their copyright. A reproduction of material that is protected by copyright may be a copyright infringement. A court may impose penalties and award damages in relation to offences and infringements relating to copyright material. Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.

## Recommended Citation

Cheng, Peng, Posture detection by kernel PCA-based manifold learning, Master of Computer Science thesis, School of Computer Science and Software Engineering - Faculty of Informatics, University of Wollongong, 2010. <http://ro.uow.edu.au/theses/3183>

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library: [research-pubs@uow.edu.au](mailto:research-pubs@uow.edu.au)

## **NOTE**

This online version of the thesis may have different page formatting and pagination from the paper copy held in the University of Wollongong Library.

## **UNIVERSITY OF WOLLONGONG**

### **COPYRIGHT WARNING**

You may print or download ONE copy of this document for the purpose of your own research or study. The University does not authorise you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site. You are reminded of the following:

Copyright owners are entitled to take legal action against persons who infringe their copyright. A reproduction of material that is protected by copyright may be a copyright infringement. A court may impose penalties and award damages in relation to offences and infringements relating to copyright material. Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.

# Posture Detection by Kernel PCA-based Manifold Learning

A thesis submitted in fulfillment of the  
requirements for the award of the degree

**Master of Computer Science**

from

UNIVERSITY OF WOLLONGONG

by

**Peng Cheng**

School of Computer Science and Software Engineering

September 2010

© Copyright 2010

by

Peng Cheng

All Rights Reserved

*Dedicated to*  
*Yuan-ying Cheng and Er-liang Zhang*

# Declaration

This is to certify that the work reported in this thesis was done by the author, unless specified otherwise, and that no part of it has been submitted in a thesis to any other university or similar institution.

---

Peng Cheng  
September 29, 2010

# Abstract

---

A Posture detection system aims to identify and localize any specific types of postures in images and video sequences. Unlike human or pedestrian detection where only one class of objects is required to be detected, posture detection is designed to detect multiple classes of postures. It remains a challenging problem because human bodies are complex and articulated with very diversified appearances. Posture detection often relies on a good generalization of the variations from large quantity of training examples that cover different situations. In this thesis, we devise a new posture detection framework that combines the histogram of gradient (HOG)-based feature with a novel manifold-based open-set classifier designed to achieve a better generalization. In this framework, each posture class is represented by a complex manifold that lies in the high-dimensional visual input space. The manifold is learned using Kernel PCA. Classification of a new observation is achieved by comparing it to each trained posture manifold. In addition, a new greedy Kernel PCA approximation algorithm is proposed to speed up the learning of the posture manifolds. The approximation algorithm seeks to remove the redundant training samples in the kernel space while best retaining the accuracy of kernel mapping, resulting in a new kernel PCA model that provides almost

identical learning and classification ability to the original kernel PCA with significantly lower computational cost. Both the detection framework and approximation algorithm were tested on 2D and 3D artificial datasets and real human and posture datasets. The results have shown that the approximation algorithm is effective and the proposed framework can provide accurate and efficient detection of different postures with a relatively small training set.



# Acknowledgments

---

Great thanks to my supervisors, Associate Professor Wanqing Li and Professor Philip Ogunbona, for their guidance and demonstration of the required attitudes and principles of a researcher, and their constant encouragement and motivation. Also great thanks to everyone in the Advanced Multimedia Research Lab: Ce Zhan, Nguyen Duc Thanh, Alister Coordinator, Shenbo Guo, Jun Hu, Zhenqiang Jiang and Zhou Sun for sharing their ideas and offering constructive discussions. This thesis would have not been completed without their help.

# Contents

---

<b>Abstract</b>	<b>v</b>
<b>Acknowledgments</b>	<b>vii</b>
0.1 Glossary . . . . .	1
0.2 Notations . . . . .	4
<b>1 Introduction</b>	<b>7</b>
1.1 Motivation . . . . .	7
1.2 Contributions . . . . .	9
1.3 Publication List . . . . .	9
1.4 Outline of the thesis . . . . .	9
<b>2 Literature Review</b>	<b>11</b>
2.1 Overview . . . . .	11
2.2 Detection by patterns . . . . .	15
2.2.1 Extraction of ROIs . . . . .	17
2.2.2 Feature extraction . . . . .	19

2.2.3	Classification . . . . .	28
2.2.4	Merging . . . . .	36
2.2.5	Summary . . . . .	36
2.3	Detection by local descriptors . . . . .	37
2.3.1	Local descriptors . . . . .	38
2.3.2	Combination of local cues . . . . .	41
2.3.3	Summary . . . . .	46
2.4	Performance Evaluation . . . . .	47
2.4.1	Datasets . . . . .	47
2.4.2	Evaluation criteria . . . . .	50
2.5	Major Challenges . . . . .	53
2.5.1	Occlusion . . . . .	53
2.5.2	Articulation . . . . .	56
2.5.3	Clothing . . . . .	56
2.5.4	Illumination . . . . .	57
2.5.5	Image degradation . . . . .	59
2.5.6	Cluttered background . . . . .	59
2.6	Summary and Discussion . . . . .	60
<b>3</b>	<b>Kernel PCA for open-set classification</b>	<b>61</b>
3.1	Introduction . . . . .	61
3.2	Overview of manifold learning . . . . .	64
3.3	Kernel PCA . . . . .	65
3.4	Open-set Classification based on Reconstruction Error . . . . .	69

3.5	Greedy Approximation by Minimizing the Mapping Error . . . . .	70
3.5.1	The existant methods . . . . .	71
3.5.2	The proposed method . . . . .	73
3.6	Performance Evaluation . . . . .	77
3.6.1	Mapping and Reconstruction . . . . .	77
3.6.2	Open-set Classification . . . . .	79
3.7	Discussion . . . . .	83
<b>4</b>	<b>Posture Detection by Kernel PCA</b>	<b>86</b>
4.1	System Description . . . . .	86
4.2	Extraction of HOG . . . . .	87
4.3	Training . . . . .	88
4.4	Detection . . . . .	89
4.5	Experimental Results . . . . .	93
4.5.1	Detection of humans . . . . .	93
4.5.2	Detection of postures . . . . .	95
4.5.3	Posture detection with KPCA approximation . . . . .	97
4.5.4	Detection of posture in videos . . . . .	97
4.6	Discussion . . . . .	98
<b>5</b>	<b>Conclusion</b>	<b>99</b>
5.1	Summary . . . . .	99
5.2	Future Work . . . . .	100
	<b>Bibliography</b>	<b>102</b>

# List of Tables

---

2.1	The datasets that may be used for evaluating posture detection algorithms.	49
2.2	Summary of the key literature . . . . .	53
3.1	Number of samples chosen and dimensions of the six test datasets . . .	78
3.2	Specification of 6 datasets and corresponding parameters for the reduced KPCA . . . . .	81

# List of Figures

---

2.1	Top view of a walking person, its shape can only be recognized if the viewpoint is known. . . . .	13
2.2	The framework of detection by patterns . . . . .	16
2.3	From left to right: input image, silhouette, contour, and histograms of x and y axes[36] . . . . .	21
2.4	Examples of rectangular features used in [75] . . . . .	23
2.5	Generation of the edgelet feature (a) [80] and the HOG feature (b) [19]	24
2.6	Hierarchical structure of the decision tree used in [33] and [46] . . . . .	30
2.7	The Graph indicating causality of a 2D markov random field . . . . .	35
2.8	Detection by local descriptors CITE . . . . .	38
2.9	Comparison between the combination process of the implicit shape model in [44],[65] and [66] . . . . .	45
2.10	The four cases of occlusion . . . . .	55
2.11	The four cases of clothing variations . . . . .	58

3.1	Manifold representation (red solid curve) versus cluster representation (green ellipses) of a scattering dataset, this picture illustrates the simplicity and fidelity of the manifold representation comparing to its opponent. . . . .	64
3.2	Mapping errors (y axis) w.r.t. $m$ (x axis), the results obtained by Franc's algorithm and the proposed algorithm are represented by red and blue curves respectively . . . . .	80
3.3	Reconstruction errors (y axis) w.r.t. $m$ (x axis), the results obtained by Franc's algorithm and the proposed algorithm are represented by red and blue curves respectively . . . . .	80
3.4	Decision manifolds of the open-set classifiers trained for the six synthesized objects. Column (a) shows training data in 3D data space, for 2D data this column is missing because the data are illustrated in other columns. Column (b), (c) and (d) show the classification boundaries obtained by the original KPCA, the proposed KPCA approximation method and the Franc's approximation method respectively. . . . .	82
3.5	The areas under ROC curves (AUC) of the three open-set classifiers w.r.t. $m$ (horizontal axis): Red-Original KPCA, Green-Proposed approximation method, Blue-Franc's method . . . . .	84
4.1	Schematic of the proposed posture detection system . . . . .	87
4.2	The ROC curves of the proposed detector and Dalal's detector [19] on human detection task with different kernel width $w$ . . . . .	94

4.3	ROC curves of the proposed detector in human detection task with different kernel width $w$ when negative examples are introduced, the result is compared to Dalal's detector. . . . .	95
4.4	Typical images for the 12 postures from the Weizmann action dataset [7]	95
4.5	(a) ROC curve of the proposed detector on Weizmann action database with different kernel width $w$ (b) The confusion matrix of the proposed detector. P1-P12 represents the 12 postures and NG represents the negative samples . . . . .	96
4.6	ROC curve of the proposed detector on Weizmann action database with different size of the reduced set $m$ in KPCA approximation . . . . .	97
4.7	Some detection results on the images from ETRI and INRIA videos . .	98



## 0.1 Glossary

**Adaboost** : Adaptive boosting, a boosting algorithm for ensemble learning.

**auROC** : Area under Receiver's Operating Characteristic curve, a scalar equal to the integration of an ROC curve from 0 to 1, used to measure the performance of classifiers.

**DET** : Decision Error Trade-off curve, a curve showing the missing rate versus the false positive per window (FPPW).

**DR** : Detection Rate, a scalar denoting the percentage of successful detection per window.

**DT** : Distance Transform, a 2D pixel map showing the distance from pixels to a given contour.

**DoG** : Derivative of Gaussian.

**FPPW** : False Positive per Window, a scalar denoting the percentage of false detection per window.

**GMM** : Gaussian Mixture Model, a statistic model for probabilistic estimation of a multivariate distribution.

**HMM** : Hidden Markov Model, a random process model primarily used in modeling temporary events.

**HOG** : Histogram of Gradient, a visual feature mainly used in human detection and body representation.

---

**ICA** : Independent Component Analysis: a toolset consisting of several matrix decomposition and factorization techniques that aim to isolate statistically mutually independent bases from a set of data.

**ISM** : Implicit Shape Model, an object detection framework by synthesizing local detection results through spatial voting.

**Isomap** : Isometric mapping, a nonlinear manifold embedding technique to reconstruct a mapping space that preserves the graph-distance.

**KPCA** : Kernel Principal Component Analysis, a generalized nonlinear manifold learning framework that performs the principal component analysis in a kernelized feature space.

**KNN** : K-Nearest Neighbors, a simple example-based classifier that labels each unknown datum to the majority of its K-Nearest Neighbors.

**LBP** : Local Binary Pattern, a visual descriptor that is invariant to illumination.

**LLE** : Locally Linear Embedding: a nonlinear manifold embedding technique to reconstruct a mapping space that minimizes the change of distances between adjacent data.

**LPP** : Locally Preserving Projection, a linear dimensionality reduction technique that aims to find a linear mapping to a lower dimension space that minimizes the same objective function with Laplacian eigenmap.

**MDS** : Multidimensional Scaling, a toolset that aims to reconstruct a equivalent dataset in an explicit Euclidean space from a similarity or dissimilarity matrix.

---

**MHI** : Motion History Image, a 2D greyscale image characterising motion information of a binary video.

**MRF** : Markov Random Field, a 2D statistic graphic model used to model spatially correlated random variables.

**NMF** : Non-negative Matrix Factorization, a matrix decomposition and factorization technique that aims to isolate non-negative bases from a dataset.

**PCA** : Principal Component Analysis.

**RANSAC** : Random Sample Consensus, a fast model estimation meta-algorithm.

**RBF** : Radial Basis Function, a bivariate function in the form of  $f(\|x - y\|^2)$ .

**RMI** : Recurrent Motion Image, a 2D greyscale image characterising recurrent motion information of a binary video.

**ROC** : Receiver's Operating Characteristic curve, a curve showing the detection rate versus the false positive detection.

**ROI** : Region of Interest.

**SIFT** : Scale Invariant feature transform.

**SVM** : Support Vector Machine.

## 0.2 Notations

$E(\cdot)$  : expectation of a variable or a set of random vectors.

$cov(\cdot)$  : covariance matrix of a variable or a set of random vectors.

$p(X)$  : marginal probability or likelihood of a random variable  $X$ .

$p(X|Y)$  : conditional probability or likelihood of  $X$  given  $Y$ .

$X$  : the column matrix of a training dataset.

$x_i$  : the  $i^{th}$  vector in the training dataset.

$c_i$  : the class of the  $i^{th}$  example.

$\phi(\cdot)$  : the implicit non-linear mapping that maps  $[\cdot]$  into an infinite-dimensional feature space.

$\Phi$  : the implicit column matrix of  $\phi(x_i)$ .

$k(\cdot, \cdot)$  : the positive semidefinite bivariate kernel function that defines  $\phi(\cdot)$  by its inner-product.

$K$  : the inner-product matrix of  $\phi(x_i)$  where  $K_{ij} = k(x_i, x_j)$ .

$H$  : the constant matrix for centralizing data in the feature space.  $H = I_n - 1_n$ ,  $I_n$  is an  $n \times n$  identity matrix and  $1_n$  denotes a  $n \times n$  matrix in which each element takes the value of  $1/n$ .

$\hat{\Phi}$  : the implicit column matrix of  $\phi(x_i)$  centered at zero mean.

$\hat{\phi}(x_i)$  : the  $i^{th}$  column vector of  $\hat{\Phi}$ .

$\hat{K}$  : the inner-product matrix of  $\hat{\Phi}$ ,  $\hat{K} = HKH$ .

$\lambda_i$  : the  $i^{\text{th}}$  eigenvalue.

$D$  : the diagonal matrix with each diagonal element  $D_{ii} = \lambda_i$ .

$P$  : the implicit column matrix of eigenvectors of  $\text{cov}(\hat{\Phi})$  (also known as principal components).

$A$  : the column matrix of eigenvectors of  $\hat{K}$ .

$z$  : an arbitrary datum sample in the test dataset to be classified.

$c$  : the ground truth of  $z$ .

$y(\cdot)$  : the function that maps  $z$  into the subspace of principal components in the feature space.

$H_A$  : the abbreviation for  $HA$ , it is the most important matrix in defining the KPCA model and the projection  $y(\cdot)$ .

$w(\cdot)$  : polynomial part of  $y(\cdot)$  defined by  $w(\cdot) = (H_A)^T k(X, \cdot)$ , where  $k(X, \cdot) = [k(x_1, \cdot), k(x_2, \cdot) \dots k(x_n, \cdot)]^T$

$b$  : the constant part of  $y(\cdot)$  defined by  $b = (H_A)^T K 1_n$

$U$  : the multidimensional scaling result of  $K$ .

$\tilde{X}$  : the column matrix of a subset of the training dataset selected by a kernel approximation algorithm.

$\tilde{P}$  : shortened  $P$  for faster KPCA mapping  $y(\cdot)$ .

---

$\tilde{H}_A$  : shortened  $H_A$  for faster KPCA mapping  $y(\cdot)$ .

$W_{\tilde{X}}$  : the  $w$ -mapping of the subset  $\tilde{X}$  into the KPCA mapping space (the column space of the original principal components  $P$ ).

$Q$  and  $R$  : QR decomposition of  $D_n^{\frac{1}{2}}W_{\tilde{X}}$ .

$V$  : the set of  $D_n^{\frac{1}{2}}W_{\tilde{X}}$  orthonormalized with  $W_{\tilde{X}}$ .