Francisco J. Perales
Robert B. Fisher (Eds.)

LNCS 4069

Articulated Motion and Deformable Objects

4th International Conference, AMDO 2006
Port d'Andratx, Mallorca, Spain, July 2006
Proceedings

Springer

Francisco J. Perales
Robert B. Fisher (Eds.)

Articulated Motion and Deformable Objects

4th International Conference, AMDO 2006
Port d'Andratx, Mallorca, Spain, July 11-14, 2006
Proceedings

Springer

Volume Editors

Francisco José Perales
Universitat de les Illes Balears
Department of Mathematics and Computer Science
C/Valldemossa km 7.5, PC 07122, Palma de Mallorca, Spain
E-mail: paco.perales@uib.es

Robert B. Fisher
University of Edinburgh
School of Informatics
James Clerk Maxwell Building, The King's Buildings, Mayfield Road, Edinburgh
EH9 3JZ, UK
E-mail: rbf@inf.ed.ac.uk

Library of Congress Control Number: 2006928839

CR Subject Classification (1998): I.4, I.5, I.2.10, I.3, I.2.10, I.6

LNCS Sublibrary: SL 6 – Image Processing, Computer Vision, Pattern Recognition, and Graphics

ISSN	0302-9743
ISBN-10	3-540-36031-X Springer Berlin Heidelberg New York
ISBN-13	978-3-540-36031-5 Springer Berlin Heidelberg New York

Springer is a part of Springer Science+Business Media

springer.com

Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 11789239 06/3142 5 4 3 2 1 0

Preface

The AMDO-e 2006 conference took place at the Hotel Mon Port, Port d'Andratx (Mallorca), on July 11-14, 2006, sponsored by the International Association for Pattern Recognition (IAPR), the MEC (Ministerio de Educación y Ciencia, Spanish Government), the Conselleria d'Economia, Hisenda i Innovació (Balearic Islands Government), the AERFAI (Spanish Association in Pattern Recognition and Artificial Intelligence), the EG (Eurographics Association) and the Mathematics and Computer Science Department of the UIB. Important commercial sponsors also collaborated with practical demonstrations; the main contributions were from: VICOM Tech, ANDROME Iberica, GroupVision, Ndigital (NDI), CESA and TAGrv.

The subject of the conference was ongoing research in articulated motion on a sequence of images and sophisticated models for deformable objects. The goals of these areas are to understand and interpret the motion of complex objects that can be found in sequences of images in the real world. The main topics considered as priority were: geometric and physical deformable models, motion analysis, articulated models and animation, modelling and visualization of deformable models, deformable models applications, motion analysis applications, single or multiple human motion analysis and synthesis, face modelling, tracking, recovering and recognition models, virtual and augmented reality, haptics devices, biometrics techniques. These topics were grouped into four tracks: **Track 1:** Computer Graphics (Human Modelling and Animation), **Track 2:** Human Motion (Analysis, Tracking, 3D Reconstruction and Recognition), **Track 3:** Multimodal User Interaction (VR and AR, Speech, Biometrics) and **Track 4:** Advanced Multimedia Systems (Standards, Indexed Video Contents).

This conference was the natural evolution of the AMDO2004 workshop (Springer LNCS 3179). The goal of this conference was to promote interaction and collaboration among researchers working in the areas covered by the four tracks. New perceptual user interfaces and linked emerging technologies strengthen the relation between the conference themes and human–computer interaction. The perspective of the AMDO-e 2006 conference was to strengthen the relationship between the many areas that have as key point the study of the human body using computer technologies as the main tool. The response to the call of papers for this conference was very good. From 81 full papers submitted, 53 were accepted. The review process was carried out by the Program Committee, each paper being assessed by at least two reviewers. The conference included several parallel sessions of orally presented papers, poster sessions and three tutorials. Moreover, the conference benefited from the collaboration of the invited speakers covering various aspects of the main topics. These invited speakers were: Thomas Vetter from Basel University (Switzerland), José Santos-Victor from IST (Portugal), and Petia Radeva from Computer Vision Center (UAB, Spain).

July 2006

F. J. Perales and B. Fisher
General Co-chairs
AMDO-e 2006

Organization

AMDO-e 2006 was organized by the Computer Graphics and Vision team of the Department of Mathematics and Computer Science, Universitat de les Illes Balears (UIB) in cooperation with IAPR (International Association for Pattern Recognition), AERFAI (Spanish Association for Pattern Recognition and Image Analysis) and EG (Eurographics Association).

Executive Committee

General Conference Co-chairs:	F. J. Perales, Mathematics and Computer Science Department, UIB (Spain) B. Fisher, University of Edinburgh (UK)
Organizing Chairs:	M. J. Abásolo, J. M. Buades, A. Delgado, G. Fiol, M. González, A. Igelmo, A. Jaume-Capó, C. Manresa-Yee, R. Mas, M. Mascaró P., M. Miró, P. Palmer, J. Varona Mathematics and Computer Science Department (UIB, Spain)
Tutorial Chairs:	M. González, J. Varona, F.J. Perales UIB (Spain)

Program Committee

Abásolo, M. J.	Universitat Illes Balears, Spain
Aloimonos, Y.	University of Maryland, USA
Aggarwal, J. K.	University of Texas, USA
Ayache, N.	INRIA, Sophia-Antipolis, France
Badler, N. I.	University of Pennsylvania, USA
Bibiloni, A.	Universitat Illes Balears, Spain
Blake, A.	Microsoft Rs, Cambridge
Boulic, R.	EPFL, Switzerland
Bowden, R.	University of Surrey, UK
Brunet, P.	UPC, Spain
Campilho, A.	University of Oporto, Portugal
Carlsson, S.	CVAP, Sweden
Cerezo, E.	Universidad de Zaragoza, Spain
Cipolla, R.	Univ. of Cambridge, UK
Cohen, I.	University of Southern California, USA
Costa-Orvalho, V.	UPC, Spain
Cristani, M.	University of Verona, Italy

Dalal, N.	INRIA, Alpes, France
Davis, L. S.	University of Maryland, USA
Del Bimbo, A.	Univ. di Firenze, Italy
Di Fiore, F.	Hasselt University, Belgium
Fengjun, Lv	University of Southern California, USA
Fernández-Caballero, A	Univ. Castilla-La Mancha, Spain
Fiol, G.	Universitat Illes Balears, Spain
Fleet, D.	University of Toronto, Canada
Flerackers, E.	Hasselt University, Belgium
Fua, P	EPFL, Switzerland
Gong, S.	QM Westfield College, UK
González, M.	Universitat Illes Balears, Spain
González, J.	CVC-UAB, Spain
Guerra-Filho, G.	University of Maryland, USA
Hilton, A.	University of Surrey, UK
Horaud, R.	INRIA, Alpes, France
Huang, T. S.	Univ. Urbana-Champaign, USA
Igelmo, A.	Universitat Illes Balears, Spain
Iñesta, J. M.	Universitat d'Alacant, Spain
Jain, A.	Michigan State University, USA
Kittler, J.	University of Surrey, UK
Kunii, T. L.	University of Hosei, Japan
Long, Q.	Hong Kong University of S& T, China
Marchand, E.	IRISA-INRIA Rennes, France
Mas, R.	Universitat Illes Balears, Spain
Mascaró-Portells, M.	Universitat Illes Balears, Spain
Medioni, G.	University of Southern California, USA
Mir, A.	Universitat Illes Balears, Spain
Miró, M.	Universitat Illes Balears, Spain
Moeslund, T.	University of Aalborg, Denmark
Moscini, D.	University of Verona, Italy
Murino, V.	University of Verona, Italy
Obrenovic, Z.	Univ. of Belgrade, Serbia
Oncina, J.	Universitat d'Alacant, Spain
Pentland, A.	Media Lab, MIT, USA
Perales, F.J.	Universitat Illes Balears, Spain
Pérez de la Blanca, N.	University of Granada, Spain
Pla, F.	University of Jaume I, Spain
Poggio, T.	MIT, USA
Qin, H.	Stony Brook University, New York, USA
Radeva, P	UAB-CVC, Spain
Raducanu, B.	CVC-UAB, Spain
Raymaekers, C.	Hasselt University, Belgium
Reid, I.	University of Oxford, UK

Serón, F.	University of Zaragoza, Spain
Shimada, N.	University of Osaka, Japan
Shirai, Y.	University of Osaka, Japan
Skala, V.	University of Plzen, Czech Republic
Susin, A.	UPC, Spain
Tavares, J.	Univ. do Porto, Portugal
Terzopoulos, D.	University of Toronto, Canada
Thalmann, D.	EPFL, Switzerland
Varona, J.	Universitat Illes Balears, Spain
Vetter, Th.	Univ. Basel, Switzerland
Villanueva, J.	UAB-CVC, Spain
Weinland, D.	INRIA, Alpes, France

Sponsoring Institutions

IAPR (International Association for Pattern Recognition)
AERFAI (Spanish Association for Pattern Recognition and Image Analysis)
EG (Eurographics Association)
MEC (Ministerio de Educación y Ciencia, Spanish Government)
Conselleria d'Economia, Hisenda i Innovació (Balearic Islands Government)
Mathematics and Computer Science Department, Universitat de les Illes Balears (UIB)
Ajuntament d' Andratx, Sol de Ponent
Consell de Mallorca

Commercial Sponsoring Enterprises

VICOM-Tech S.A., www.vicomtech.es
ANDROME Iberica S.A, www.androme.es
GroupVision, www.groupvision.es
Ndigital (NDI), www.ndigital.com
CESA S.A, www.cesa.es
TAGrv S.L., www.tagrv.com

Table of Contents

Articulated Motion and Deformable Objects AMDO^{-e} 2006

A Study on Human Gaze Detection Based on 3D Eye Model

Kang Ryoung Park

Division of Media Technology, Sangmyung University,
7 Hongji-Dong, Jongro-ku, Seoul, Republic of Korea,
Biometrics Engineering Research Center
parkgr@smu.ac.kr

Abstract. Human gaze can give valuable tips for human computer interaction, but it is very difficult to detect human gaze position with one or two camera systems. Conventional method has the limitation of inaccurate gaze detection performance or not being able to track the fast motion of user's face and eye.

To overcome such problem, in this paper, we propose a new gaze detection method. Compared to previous works, our method has following three advantages. First, our method uses three camera systems, such as a wide and narrow view stereo cameras and allows user's natural head and eye movement. Second, to obtain gaze position on a monitor, we detect the 3D eye position and gaze vector of eyeball. Third, to enhance the eye detection performance, we use AdaBoost eye detector and PCA algorithm.

Experimental results showed that our method could be used for real-time gaze detection system.

Keywords: Gaze Detection, a Wide and Stereo Narrow View Cameras, 3D Eye Position and Gaze Vector of Eyeball.

1 Introduction

Human gaze can provide important tips for many applications such as view controlling in 3D simulation programs, virtual reality, video conferencing and special human-machine interface/controls. Most previous researches were focused on 2D/3D head motion estimation [2][11], the facial gaze detection (allowing for only head movement)[3-9][12][13][15] and the eye gaze detection (allowing for only eye movement)[10][14]. Wang et al.[1]'s method provides the advanced approaches that combines head pose and eye gaze estimation by a wide view camera and a panning/tilting narrow camera. However, in order to compute the gaze position, their method supposes that they know the 3D distances between two eyes, eye corners, both lip corners and the 3D diameter of eye ball. Also, they suppose that there is no individual variation for the 3D distances and diameter. However, our preliminary experiments show that there are much individual variations for the 3D distances/3D diameter and such cases can increase gaze errors. Moreover, the accuracy of their method rapidly drops down according as the distance between the camera and the user's face increases. More advanced method using narrow and wide view stereo cameras were shown [15]. However, in that method, user should gaze at 5 known

F.J. Perales and R.B. Fisher (Eds.): AMDO 2006, LNCS 4069, pp. 1 – 9, 2006.

(pre-determined) positions on a monitor to obtain the 3D position information of facial and eye features in calibration stage. Also, it uses the method of mapping 2D feature information (in narrow view eye image) to the gaze position on a monitor, directly, without considering the 3D information of eye feature and it can be main factor to increase gaze error.

To overcome such problems of previous researches and systems, we propose the new method for detecting gaze position with three cameras composed of one wide and stereo narrow view cameras. To exclude the large specular reflection on glasses surface, we use dual (left and right) IR-LED illuminators for wide and narrow view camera.

In section II, we present the method of detecting facial and eye features. In section III, the method of computing 3D eye feature position is shown. In section IV, the method of calculating final gaze position on a monitor is explained. Experimental results and conclusion are included in section V and VI, respectively.

2 Detecting Eye Region in Wide View Image by AdaBoost Algorithm and PCA

In order to detect gaze position on a monitor, facial features in wide view images should be obtained. To detect facial features robustly, we implement a gaze detection system.

To detect the eye features in wide view camera image, we use AdaBoost eye detector and PCA (Principal Component Analysis). In previous work, AdaBoost algorithm is used for face detection and we adopt it for eye detection. Original AdaBoost classifier uses a boosted cascade of simple classifiers using Haar-like features capable of detecting faces in real-time with both high detection rate and very low false positive rates, which is considered to be one of the fastest systems [3][4]. For that, we trained 90 eye images and 90 non-eye images. Then, with trained AdaBoost eye detection classifier, we detect eye region from input test image.

However, some case of FAR (False Acceptance error which accept non-eye region as eye region) happen and to reduce it, we also use PCA eye detection to verify the

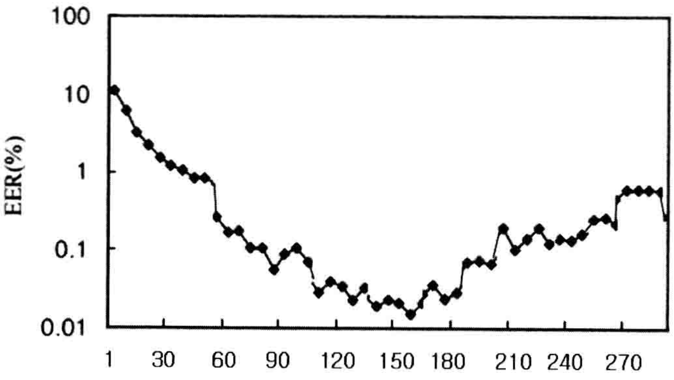

Fig. 1. The EER versus the number of the PCA eigenvector

detected eye region. For PCA training, we used 90 eye images with the size of 60*30 pixels. From that, we obtained total 1800 eigenvectors and found the optimal number of eigenvectors with which the EER (Equal Error Rate) for find eye region was minimized. Experimental results showed that the EER was smallest in case of 160 eigenvectors as shown in Fig. 1. Finally, experimental results showed that correct rate of eye detection was 99.8%.

In the detected eye region, we locate the accurate left and right eye (iris) center by the circular edge detection method [12]. Some examples of the detected eye regions are shown in Fig.2. Experimental results show that RMS error between the detected eye center positions and the actual ones are 1 pixel in 640×480 pixels image. Also, because the eye center detection is performed only in the detected eye region, it takes little time as 1ms in Pentium-IV 1.8 GHz PC.

Fig. 2. The examples of detected eye region by wide view camera

3 Locating Eye Features in Narrow View Camera

Based on the detected 2D eye center positions in wide view camera, we try to pan and tilt the stereo narrow view cameras to capture the magnified eye image. However, with the detected 2D eye center positions in wide view camera, only the line of sight through the user's eye center point can be obtained by single wide view camera. To determine the accurate panning/tilting angles of stereo narrow view cameras, we should know the 3D Z distance between the user's eye and the wide view camera and it is infeasible with mono wide view camera. In addition, because we do not know the accurate panning/tilting angles, we cannot also determine the accurate viewing angle of narrow view cameras, with which they can capture the user's magnified eye image. However, we can preliminarily determine the initial viewing angle of narrow view camera based on following conditions without knowing the accurate panning & tilting angles by Z distance; 1) Most users tend to sit in the Z distance range of 50 ~ 70 cm in front of monitor. 2) The sitting heights (we define it as the Y distance between the origin in monitor coordinate ($\mathbf{X_m}$, $\mathbf{Y_m}$, $\mathbf{Z_m}$) and the facial center of most user) are about -20 ~ +30 cm on average (we measured them from 95 persons test data by Polhemus position tracker sensor [11]). 3) The stereo narrow view cameras capture the user's eye on the slant (on the wide view camera).

From the above conditions, we can restrict the initial viewing angle of narrow view camera as 4.3 degree (-2.15 ~ +2.15 degree, vertically) and obtain the magnified eye image by narrow view cameras (In this case, the diameter of iris is about 135 pixels at the Z distance of 50 cm) without knowing the accurate panning & tilting angle of narrow view camera. From the captured eye image, we can detect more accurate 2D pupil center and boundary points (we detect 6 boundary points with the angular interval of 60 degrees as shown in Fig. 3) by locating corneal specular reflection and circular pupil edge detection [12]. Experimental results show that RMS error between the detected feature positions and the actual ones are 1.2 pixels (of pupil center) and 1.1 pixels (of pupil boundary points) in 640×480 pixels image. Some examples of the detected features are shown in Fig.3.

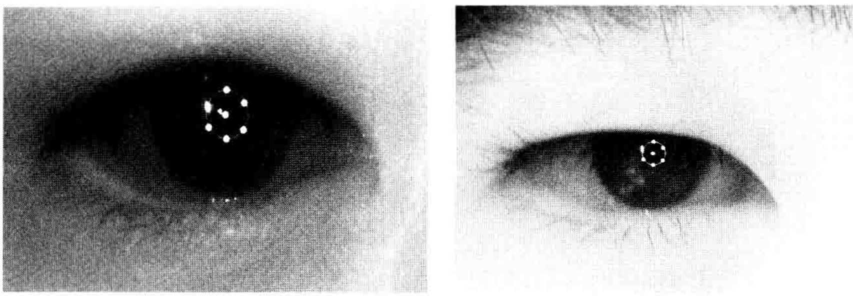

Fig. 3. The Examples of the detected eye features by narrow view camera

However, in some case of users with glasses, large specular reflection on the glasses surface happens by our illuminator. In such a case, our algorithm may detect the erroneous pupil center and boundary points and calculate inaccurate gaze position, consequently. To prevent such problems, our algorithm checks the average gray level of detected pupil region and if it exceeds in the predetermined threshold (we define it as 30), our algorithm commands to turn on the other illuminator in opposite side (from left to right, or from right to left) of narrow view camera and capture the clear eye image which does not include the occlusion of large specular reflection.

4 Computing 3D Eye Feature Positions

In this section, we explain the method of computing 3D eye positions by stereo narrow view cameras. Supposing that the point "M" (which is the pupil center of right eye) is observed by the left and right narrow view cameras, then we can obtain 3D positions of "M" by conventional stereo camera theory [8].

Considering the coordinate conversion between the left and right narrow view camera [8], we can obtain the relationship between two camera coordinates by affine transformation. Using them, we can obtain the Z distance (Z_l) of "M" point in the left narrow view camera coordinate. With the obtained Z distance (Z_l), we can obtain the 3D positions(X_l, Y_l, Z_l) of the feature point (M) in the left narrow view camera coordinate ($\mathbf{X_{N1}}$, $\mathbf{Y_{N1}}$, $\mathbf{Z_{N1}}$).