

Chapter 18

Detecting Eyes and Lips Using Neural Networks and SURF Features

Artem A. Lenskiy

Korea University of Technology and Education, Korea

Jong-Soo Lee

University of Ulsan, Korea

ABSTRACT

In this chapter, the authors elaborate on the facial image segmentation and the detection of eyes and lips using two neural networks. The first neural network is applied to segment skin-colors and the second to detect facial features. As for input vectors, for the second network the authors apply speed-up robust features (SURF) that are not subject to scale and brightness variations. The authors carried out the detection of eyes and lips on two well-known facial feature databases, Caltech. and PICS. Caltech gave a success rate of 92.4% and 92.2% for left and right eyes and 85% for lips, whereas the PCIS database gave 96.9% and 95.3% for left and right eyes and 97.3% for lips. Using videos captured in real environment, among all videos, the authors achieved an average detection rate of 94.7% for the right eye and 95.5% for the left eye with a 86.9% rate for the lips

INTRODUCTION

A number of approaches have been proposed to detect facial features. Among all facial features, eyes have become a maverick for their variety of applications and implications. In the case of car driver eye detection, color cameras and cameras with IR illumination have been widely used as

a video capturing device. The use of cameras with IR illumination significantly simplifies the problem of eye detection. When an infrared led is located at the camera axis, the eye irises appear as two bright spots caused by the reflection of the blood-rich retina. Thus, IR illumination based approaches gained high popularity in eye detection (Zhao & Grigat, 2006) and consequently driver attention monitoring tasks. In (Ji & Yang, 2002) a system for monitoring driver vigilance is pro-

DOI: 10.4018/978-1-61350-429-1.ch018

posed. The main idea is to place two cameras at different angles. The camera axis coincides with two coplanar concentric rings where along their circumferences a number of IR LEDs are evenly and symmetrically distributed. One camera has a wide range view for head tracking, and the other has a narrow view for eye detection. A similar hardware setup was applied by Batista et al. (Batista, 2005), the difference from (Ji & Yang, 2002) consists in face and eye detection algorithms. Hammoud et al. (Hammoud, Witt, Dufour, Wilhelm, & Newman, 2008) proposed a complete driver drowsiness detection system that detects irises in the near infrared spectrum. Although IR based approaches perform reasonably at night time, it was noted (Hartley, Horberry, & Mabbott, 2000) that those methods often malfunctioned during daytime under the presence of sunlight. Moreover, when eyes are closed the reflection in IR range disappears, making eye detection a difficult task. Another disadvantage of IR based approaches is the necessity of installing an IR LEDs setup.

In comparison with IR cameras, CMOS and CCD cameras are passive, meaning there is no IR radiation. The effect of long term IR radiation should be studied to guarantee that there is no danger to eye health (Pitts, Cullen, & Dayhew-Barker, 1980). CMOS cameras are relatively inexpensive and ergonomic. Furthermore, according to (Hartley, et al., 2000) 52% of drivers nodded off while driving between the times of 6:00 a.m. and 9:00 p.m. comprising of the majority of bright daylight on any given day. As a consequence, using IR cameras during those hours is impractical since IR cameras are inefficient under direct sunlight.

In the case of a color camera, there is a possibility to take color information into account for skin-color segmentation purposes (Rong-ben, Ke-you, Shu-ming, & Jiang-wei, 2003). The skin-color segmentation process is commonly done in RGB, HSI or YCbCr color spaces. Some authors have heuristically found an RGB to 2-dimensional color space transform and then approximated the skin color domain in 2D space (Butler, Sridharan,

& Chandran, 2002; Hamdy, Elmahdy, & Elsbroudy, 2007; Naseem & Deriche, 2005; Phil & Christos, 2005; Tariq, Jamal, Shahid, & Malik, 2004). These color spaces along with manual skin-domain approximation methods (Hernandez & Kleiman, 2005; Rong-ben, et al., 2003) are not capable of finding complex boundaries of skin-color domains. One approach to define complex boundaries of skin-color domains is to train an artificial neural network to separate skin and non-skin colors. Several applications based on neural networks have been proposed for skin color filtering. A two layer multi-layer perceptron (MLP) with two inputs and three hidden neurons was applied by Chen et al. (Chen & Chiang, 1997). In their work the RGB color space was transformed into normalized CIE XYZ color space, then the values of X and Y coordinates served as inputs for MLP. Another attempt using the MLP skin-color filter was proposed by Seow (Seow, Valaparla, & Asari, 2003), where one additional neural network was used at the learning stage to interpolate spatial skin region in each of the training images. The two hidden layers MLP then learns to distinguish skin-colors obtained from interpolated regions and non-skin colors from the rest of the image. Sahbi and Boujemaa (Sahbi & Boujemaa, 2000) applied a two layer MLP with two hidden neurons and three inputs and outputs. This structure allowed them to extract principle skin color components. Lenskiy and Lee (A. Lenskiy & J.-S. Lee, 2010) applied one layer feed forward neural network to segment skin-colors. They suggest an interesting and simple approach to train the network with negative samples uniformly distributed in the color space. As soon as the segmentation process is over, the remaining skin-color segments are analyzed to select a face candidate. For this purpose facial proportions and morphological operators are usually applied.

For the detection of eyes and lips a number of methods have been suggested. Some approaches are based on raw eye images which are fed into a classifier, such as support vector machines (Jee,

15 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/detecting-eyes-lips-using-neural/62698

Related Content

Structure Analysis of General Type-2 Fuzzy Controller and Its Application

Jianzhong Shi (2023). *International Journal of Fuzzy System Applications* (pp. 1-20).

www.irma-international.org/article/structure-analysis-of-general-type-2-fuzzy-controller-and-its-application/319813

The Race Between Cognitive and Artificial Intelligence: Examining Socio-Ethical Collaborative Robots Through Anthropomorphism and Xenocentrism in Human-Robot Interaction

Anshu Saxena Arora and Amit Arora (2020). *International Journal of Intelligent Information Technologies* (pp. 1-16).

www.irma-international.org/article/the-race-between-cognitive-and-artificial-intelligence/243367

An Agent-Based Approach for Sourcing Business Rules in Supply Chain Management

Sudha Ram and Jun Liu (2005). *International Journal of Intelligent Information Technologies* (pp. 1-16).

www.irma-international.org/article/agent-based-approach-sourcing-business/2376

Design and Simulation of an Automated Sanding Machine for the Roughing of Plastic Parts

Noemi González León and Abraham Hernández Santiago (2023). *Streamlining Organizational Processes Through AI, IoT, Blockchain, and Virtual Environments* (pp. 46-59).

www.irma-international.org/chapter/design-and-simulation-of-an-automated-sanding-machine-for-the-roughing-of-plastic-parts/325336

Human Health Risk Assessment via Amalgamation of Probability and Fuzzy Numbers

Palash Dutta (2019). *Emerging Trends and Applications in Cognitive Computing* (pp. 99-123).

www.irma-international.org/chapter/human-health-risk-assessment-via-amalgamation-of-probability-and-fuzzy-numbers/218997