Feature Extraction and Localisation using Scale-Invariant Feature Transform on 2.5D Image

[1]SukTing Pui, [2]Jacey-Lynn Minoi, [3]Terrin Lim
Faculty of Computer Science and Information Technology
Universiti Malaysia Sarawak
94300 Kota Samarahan, Sarawak, Malaysia
[1]polly_sukting@hotmail.com, [2]jacey@fit.unimas.my, [3]terrin@fit.unimas.my

João Fradinho Oliveira
C3i/Instituto Politécnico de Portalegre, Portalegre
Portugal
jfoliveira@estgp.pt

Duncan Fyfe Gillies
Department of Computing
Imperial College London
United Kingdom
d.gillies@imperial.ac.uk

ABSTRACT

The standard starting point for the extraction of information from human face image data is the detection of key anatomical landmarks, which is a vital initial stage for several applications, such as face recognition, facial analysis and synthesis. Locating facial landmarks in images is an important task in image processing and detecting it automatically still remains challenging. The appearance of facial landmarks may vary tremendously due to facial variations. Detecting and extracting landmarks from raw face data is usually done manually by trained and experienced scientists or clinicians, and the landmarking is a laborious process. Hence, we aim to develop methods to automate as much as possible the process of landmarking facial features. In this paper, we present and discuss our new automatic landmarking method on face data using 2.5-dimensional (2.5D) range images. We applied the Scale-invariant Feature Transform (SIFT) method to extract feature vectors and the Otsu's method to obtain a general threshold value for landmark localisation. We have also developed an interactive tool to ease the visualisation of the overall landmarking process. The interactive visualization tool has a function which allows users to adjust and explore the threshold values for further analysis, thus enabling one to determine the threshold values for the detection and extraction of important keypoints or/and regions of facial features that are suitable to be used later automatically with new datasets with the same controlled lighting and pose restrictions. We measured the accuracy of the automatic landmarking versus manual landmarking and found the differences to be marginal. This paper describes our own implementation of the SIFT and Otsu's algorithms, analyzes the results of the landmark detection, and highlights future work.

Keywords

Feature extraction, localization, landmarking, Otsu's algorithm

1. INTRODUCTION

The human vision system can perceive features such as the edges, tips or corners of an object, without any difficulties. For example, a human is able to detect and recognize the eyes, the tip of the nose and/or the mouth of a person at first glance. However, a computer is unable to do such task easily and effortlessly [CUS13]. The human vision system and brain mechanisms that are responsible for the detection of features are so complex that despite the work of neurobiologists, mathematicians and computer scientists, it is still not possible to replicate facial detection accurately.

In this paper, we propose methods to obtain distinct features on a face and label them automatically by placing landmarks. Our objective is to first empirically find specific landmarking threshold values that are valid for a given set of example datasets. Later, we use these found values in an automatic landmark setting with new example datasets and compare the accuracy versus manual landmarking.

Extracting facial information automatically is a challenging process, due to linear and nonlinear transformations, and therefore may not give a valid representation of the objects. The extents of spatial features differ significantly from different scales, sizes and resolutions. Therefore, the feature extraction methods ought to be invariant of scale and orientation. The SIFT approach takes a face image and transforms it into a collection of feature vectors. These features can either be global or local, defining the whole or a part of the image respectively. Usually the local interest descriptors which can be used to find distinct features on the face are selected manually. In our project, we make