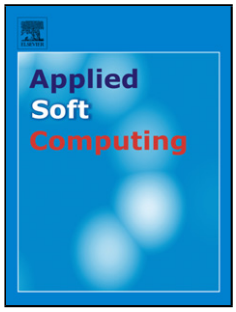

Title: A robust incremental clustering-based facial feature tracking

Author: Md. Nazrul Islam Manjeevan Seera Chu Kiong Loo

PII: S1568-4946(16)30658-5
DOI: http://dx.doi.org/doi:10.1016/j.asoc.2016.12.033
Reference: ASOC 3972

To appear in: *Applied Soft Computing*

Received date: 4-1-2015
Revised date: 2-12-2016
Accepted date: 19-12-2016

Please cite this article as: Md.Nazrul Islam, Manjeevan Seera, Chu Kiong Loo, A robust incremental clustering-based facial feature tracking, Applied Soft Computing Journal http://dx.doi.org/10.1016/j.asoc.2016.12.033

A robust incremental clustering-based facial feature tracking

Md. Nazrul Islam [a], Manjeevan Seera [b*], Chu Kiong Loo [c]

[a] *Department of Computer Science & Engineering, Dhaka University of Engineering & Technology, Gazipur, Bangladesh*
[b] *Faculty of Engineering, Computing and Science, Swinburne University of Technology (Sarawak Campus), Malaysia*
[c] *Faculty of Computer Science and Information Technology, University of Malaya, Kuala Lumpur, Malaysia*
Email address: [a] *nazrul.cse@duet.ac.bd,* [b] *mseera@gmail.com,* [c] *ckloo.um@um.edu.my*

Highlights

- RICFFT offers a person-independent, automatic facial feature tracking system.
- A 17-point feature model enables emotion-specific facial feature on frontal face regions to capture dynamic facial muscle movement.
- Two incremental clustering algorithms are developed to build andupdate the appearance and model incrementally through online learning of dynamic facial features.
- Results show capability of tracking facial features online, captured dynamic facial muscle movement with increased accuracy.

Abstract Emerging significance of person-independent, emotion specific facial feature tracking has been actively tracked in the machine vision society for decades. Among distinct methods, the Constrained Local Model (CLM) has shown significant results in person-independent feature tracking. In this paper, we propose an automatic, efficient, and robust method for emotion specific facial feature detection and tracking from image sequences. A novel tracking system along with 17-point feature model on the frontal face region has also been proposed to facilitate the tracking of human basic facial expressions. The proposed feature tracking system keeps patch images and face shapes till certain number of key frames incorporating CLM-based tracker. After that, incremental patch and shape clustering algorithms is applied to build appearance model and structure model of similar patches and similar shapes respectively. The clusters in each model are built and updated incrementally and online, controlled by amount of facial muscle movement. The overall performance of the proposed Robust Incremental Clustering-based Facial Feature Tracking (RICFFT) is evaluated on the FGnet database and the Extended Cohn-Kanade (CK+) database. RICFFT demonstrates mean tracking accuracy of 97.45% and 96.64% for FGnet and CK+ database respectively. Also, RICFFT is more robust by minimizing average shape distortion error of 0.20% and 1.86% for FGnet and CK+ (apex frame) database, as compared with classic method CLM.

Keywords Facial feature tracking; incremental clustering; feature tracking framework; facial feature model; constrained local model.

1. Introduction

State-of-the-art computing systems are anticipated to establish interactions with humans in a harmonious and natural way that emulates face-to-face encounters. With relevant communication tools and techniques, computing systems are well equipped and smart enough to interact with the users. Not only treated as imitating user's doings, it also has self-learning capabilities to respond with user's dynamic behaviour [1]. Computing system has gained popularity [2] in various aspects such as motion tracking [3], telemonitoring of elderly people, video-conferencing with family members, customer satisfaction studies, e-health [4], and in education and learning [5].

Using visual cues and embedding emotional intelligence in computing system, future generations of human-computer interaction (HCI) will be more sophisticated and human-like [6]. In a human-human communication (HHC), implicit and non-verbal information such as facial expressions, transmit their emotions to other people in a non-ambiguous way [7]. Several researchers in the field of psychology have reported facial expressions and emotions are closely related to each other, and it can be regarded as one of the most essential methods to represent one's emotions [8]. It is an important tool for analysing abnormality of neuropsychiatric disorders to evaluate their emotional impairment level [2, 9]. Utilizing this tool with facial features, several significant researches have been carried out to criminal investigation [10], identify speaker [11], detect fatigue level of driver [12], and warn driver [13] in real time. In this regard, precise localization of facial feature points surrounding facial components such as the mouth, nose, eyes, and eyebrows would greatly benefit to both HCI and HHC. At this point,