

Automated video-based facial expression analysis of neuropsychiatric disorders

Peng Wang^a, Frederick Barrett^b, Elizabeth Martin^b, Marina Milonova^c,
Raquel E. Gur^{d,e,f}, Ruben C. Gur^{b,e}, Christian Kohler^f, Ragini Verma^{a,*}

^a Section of Biomedical Image Analysis, Department of Radiology, University of Pennsylvania, 3600 Market, Suite 380, Philadelphia, PA 19104, USA

^b Brain Behavior Center, Department of Psychiatry, University of Pennsylvania Medical Center, Hospital of the University of Pennsylvania,
3400 Spruce Street, 10th Floor Gates Building Philadelphia, PA 19104, USA

^c School of Arts and Sciences, University of Pennsylvania Medical Center, Hospital of the University of Pennsylvania,
3400 Spruce Street, 10th Floor Gates Building Philadelphia, PA 19104, USA

^d Department of Psychiatry, University of Pennsylvania Medical Center, Hospital of the University of Pennsylvania,
3400 Spruce Street, 10th Floor Gates Building Philadelphia, PA 19104, USA

^e Department of Neurology & Radiology, University of Pennsylvania Medical Center, Hospital of the University of Pennsylvania,
3400 Spruce Street, 10th Floor Gates Building Philadelphia, PA 19104, USA

^f Neuropsychiatry Section, University of Pennsylvania Medical Center, Hospital of the University of Pennsylvania,
3400 Spruce Street, 10th Floor Gates Building Philadelphia, PA 19104, USA

Received 16 July 2007; received in revised form 20 September 2007; accepted 20 September 2007

Abstract

Deficits in emotional expression are prominent in several neuropsychiatric disorders, including schizophrenia. Available clinical facial expression evaluations provide subjective and qualitative measurements, which are based on static 2D images that do not capture the temporal dynamics and subtleties of expression changes. Therefore, there is a need for automated, objective and quantitative measurements of facial expressions captured using videos. This paper presents a computational framework that creates probabilistic expression profiles for video data and can potentially help to automatically quantify emotional expression differences between patients with neuropsychiatric disorders and healthy controls. Our method automatically detects and tracks facial landmarks in videos, and then extracts geometric features to characterize facial expression changes. To analyze temporal facial expression changes, we employ probabilistic classifiers that analyze facial expressions in individual frames, and then propagate the probabilities throughout the video to capture the temporal characteristics of facial expressions. The applications of our method to healthy controls and case studies of patients with schizophrenia and Asperger's syndrome demonstrate the capability of the video-based expression analysis method in capturing subtleties of facial expression. Such results can pave the way for a video-based method for quantitative analysis of facial expressions in clinical research of disorders that cause affective deficits.

© 2007 Elsevier B.V. All rights reserved.

Keywords: Facial expression; Video analysis; Schizophrenia; Affective deficits; Pattern classification

1. Introduction

Facial expressions have been used in clinical research to study deficits in emotional expression and social cognition in neuropsychiatric disorders (Morrison et al., 1988; Berenbaum and

Oltmann, 1992; Kring et al., 1994; Mandal et al., 1998). Specifically, patients with schizophrenia often demonstrate two types of impairments in facial expressions: “flat affect” and “inappropriate affect” (Gur et al., 2006). However, most of the current clinical methods, such as the scale for assessment of negative symptoms (SANS (Andreasen, 1984)), are based on subjective ratings and therefore provide qualitative measurements. They also require extensive human expertise and interpretation. This underlines the need for automated, objective and quantitative measurements of facial expression. We previously reported a method for quantifying facial expressions based on static images

* Corresponding author. Tel.: +1 215 662 7471; fax: +1 215 614 0266.

E-mail addresses: peng.wang@uphs.upenn.edu (P. Wang),
fbarrett@bbl.med.upenn.edu (F. Barrett), raquel@bbl.med.upenn.edu
(R.E. Gur), gur@bbl.med.upenn.edu (R.C. Gur), kohler@bbl.med.upenn.edu
(C. Kohler), Ragini.Verma@uphs.upenn.edu (R. Verma).

(Verma et al., 2005; Alvino et al., 2007). However, temporal information plays an important role in understanding facial expressions because emotion processing is naturally a temporal procedure. Therefore, facial expression analysis from static 2D images lacks the temporal component, which is essential to capture subtle changes in expression. Although video-based acquisition has been employed in the examination of facial emotion expression (Kring and Sloan, 2007), currently there is no objective and automated way of facial expression analysis for the study of neuropsychiatric disorders, particularly due to the large volume of data that makes human analysis prohibitive. In this paper, we present a computational framework that uses videos to automatically analyze facial expressions and can be used to characterize impairments in such neuropsychiatric disorders.

The merits of automated facial expression analysis (AFEA) are two-fold: using it can avoid intensive human efforts, and can provide unified quantitative results. There are already many AFEA methods being presented in both clinical and computer vision communities (Gaebel and Wölwer, 1992; Hellewell et al., 1994; Schneider et al., 1990; Pantic and Rothkrantz, 2000; Fasel and Luetttin, 2003; Tian et al., 2005). Most of the current AFEA methods focus on the recognition of posed facial expressions with application to human computer interaction tasks, and only a few of them have been applied to clinical studies (Verma et al., 2005; Alvino et al., 2007). In previous work on expression quantification (Verma et al., 2005; Alvino et al., 2007), the expression changes were modeled using elastic shape transformations between the face of a neutral template and the corresponding emotionally expressive face. Again, as most of the current AFEA methods, this approach is based on static 2D images without any temporal component.

In this paper, we present a computational framework that uses videos for the analysis of facial expression changes. This framework explores the dynamic information that is not captured by static images during emotion processing, and provides computationally robust results with potential clinical applicability. Broadly, our computational framework includes the detection of faces in videos, which are then tracked through the video, incorporating shape changes. Based on tracking results, features are extracted from faces to create probabilistic facial expression classifiers. The probabilistic outputs of facial expression classifiers are propagated throughout the video, to create probabilistic profiles of facial expressions. Probabilistic profiles contain dynamic information of facial expressions, based on which quantitative measures are extracted for analysis. As an application of this framework, such quantitative measurements for facial expressions could be correlated with clinical ratings to study the facial expression deficits in neuropsychiatric disorders. To our knowledge, the presented framework is the first to apply video-based automated facial expression analysis in neuropsychiatric research.

The rest of the paper is organized as follows: In Section 2, previous related work is reviewed. Our computational framework is presented in Section 3. The experimental results are provided in Section 4. We discuss the results and conclude in Section 5.

2. Related work

2.1. Clinical facial expression analysis

In clinical research, facial expressions are usually studied using 2D images that are described in two ways: either as a combination of muscular movements or as universal global expressions. The Facial Action Coding System (FACS) has been developed to describe facial expressions using a combination of action units (AU) (Ekman and Friesen, 1978). Each action unit corresponds to a specific muscular activity that produces momentary changes in facial appearance. The global facial expression handles the expressions as a whole without breaking up into AUs. The most commonly studied universal expressions include happiness, sadness, anger and fear, which are referred to as universal emotions. While most of the work has been on static 2D images, the Facial Expression Coding System (FACES) (Kring and Sloan, 2007) has been designed to analyze videos of facial expressions, in terms of the duration, content and valence of universal expressions. However, these methods need intensive human intervention to rate the images and videos of facial expressions. Such rating methods are prone to subjective errors, and have difficulties in providing unified quantitative measurements. There is need for automated, objective and quantitative measurements of facial expressions.

2.2. Automated facial expression analysis

Automated facial expression analysis (AFEA) allows computers to automatically provide quantitative measurements of facial expressions. Several factors have contributed towards making AFEA challenging. First, facial expressions vary across individuals due to the differences of the facial appearance, degree of facial plasticity, morphology and frequency of facial expressions (Tian et al., 2005). Second, it is difficult to quantify the intensity of facial expressions, especially when they are subtle. In FACS, a set of rules are used to score AU intensities (Ekman and Friesen, 1978). However, such criteria are subjective to the rater; therefore, it is difficult to extend the measurements to computer-based facial expression analysis, although there have been methods to automatically detect AUs (Pantic and Rothkrantz, 2000). Many AFEA methods have been developed recently to address such problems (Pantic and Rothkrantz, 2000; Fasel and Luetttin, 2003). These methods can be categorized as image-based, video-based and 3D surface-based, according to the data used. Below we summarize some typical image-based and video-based facial expression analysis methods.

2.2.1. Image-based methods

Image-based methods extract features from individual images, and create classifiers to recognize facial expressions. Commonly used are geometric features, texture features, and their combinations. Geometric features represent the spatial information of facial expressions, such as positions of eyes and mouth, the distance between two eyebrows. The geometric features used by Tian et al. (2001) are grouped into permanent

Download English Version:

<https://daneshyari.com/en/article/4336635>

Download Persian Version:

<https://daneshyari.com/article/4336635>

[Daneshyari.com](https://daneshyari.com)