Investigating the Dynamics of Facial Expression

Jane Reilly, John Ghent, and John McDonald

National University of Ireland, Maynooth

Abstract. This paper is concerned with capturing the dynamics of facial expression. The dynamics of facial expression can be described as the intensity and timing of a facial expression and its formation. To achieve this we developed a technique that can accurately classify and differentiate between subtle and similar expressions, involving the lower face. This is achieved by using *Local Linear Embedding* (LLE) to reduce the dimensionality of the dataset and applying *Support Vector Machines* (SVMs) to classify expressions. We then extended this technique to estimate the dynamics of facial expression formation in terms of intensity and timing.

1 Introduction

Darwin, in 1872, was the first to recognize the importance of facial expressions and the role that they play in human communication [1]. During the subsequent years as behavioral scientists sought a means to objectively measure facial expressions, many different techniques and methodologies for describing facial expressions were developed (see [2] for a comprehensive review). The *Facial Action Coding System* (FACS), created by Ekman and Friesen, in 1978, is the most comprehensive of these standards and is widely used in research. The FACS provides an unambiguous quantitative means of describing all movements of the face in terms of 46 *Action Units* (AUs) [3].

Recent research has shown that it is not just the expression itself, but also its dynamics that are important when attempting to decipher its meaning [4,5,6,7, 8,9,10]. The dynamics of facial expression can be defined as the intensity of the AUs coupled with the timing of their formation. Ekman et al. suggest that the dynamics of facial expression provides unique information about emotion that is not available in static images, namely the temporal information itself [11]. There is now a growing body of psychological research that argues that temporal dynamics of facial behavior - such as the timing, duration and intensity of facial activity - is a critical factor for the interpretation of the observed behavior [4]. This paper aims to not only classify expression but also capture the dynamics of facial expression formation.

In this paper we extend our previous work [12] to classify subtle changes in expression and capture the dynamics of facial expression. To achieve this we use *Locally Linear Embedding* (LLE). LLE was originally proposed in 2000 as a non-linear dimensionality reduction technique that computes low-dimensional neighborhood preserving embeddings of high-dimensional data by unfolding the underlying manifold [13]. We show that this non-linear dimensionality reduction

G. Bebis et al. (Eds.): ISVC 2006, LNCS 4292, pp. 334–343, 2006.

technique provides a means for classifying subtle changes in expression and allows for analysis of the dynamics of facial expression. In other words, if LLE can capture the underlying manifold of facial expression independent of identity, it would then provide information about the timing and intensity of that expression, hence capturing the dynamics of facial expression.

This paper is organized as follows. In Section 2, recent research on the importance of the dynamics of facial expressions is introduced from a behavioral science perspective. Following this we survey a number of computer vision based automated techniques for the analysis of both the appearance and dynamics of facial expression. Our proposed methodology is presented in Section 3. Experiments are presented in Section 4 . Conclusions and a discussion on the results of these experiments is provided in Section 5.

2 Related Work

In this section we first review the state-of-the-art in expression analysis from a behavioral science point of view. Subsequent to this we detail computer vision based approaches for the automated analysis of facial expression.

2.1 Behavioral Science

Since the importance of facial expressions was first established in 1872 [1], many studies have been carried out attempting to interpret their meaning. However, according to Ambadar et al. [5], only a few investigators have examined the impact of dynamics in deciphering faces. These studies were largely unsuccessful due to their reliance on intense facial expressions. While the primary reasoning behind the research carried out by Ambadar et al. was to establish the importance of dynamics in deciphering facial expressions, it also highlighted the fact that facial expressions are frequently subtle. They found that subtle expressions that were not identifiable in static presentations suddenly became apparent in dynamic display.

Schmidt et al. studied the normative characteristics and individual differences that occur in spontaneous smiles. They found that spontaneous smiles, in contrast to posed smiles, have a relatively fast onset, averaging slightly over 0.5 seconds. Approximately one third of smiles included in the study consisted of AU 12 (Lip Corner Puller)[1] on its own, while the remaining smiles had one or more co-occurring AUs appearing within one second. [14]

According to Cohn, further development within the area of Automatic Facial Action Unit Recognition will eliminate the need for manual FACS coding [15]. This would prove to be a significant step forward as currently it takes approximately 100 hours of training to become even marginally proficient with FACS, and it takes 1 hour for an experienced FACS coder to encode one minute of video footage [3].

[1] for more information on AUs see the FACS manual [11].