Vision Research 127 (2016) 1–10

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Contributions of feature shapes and surface cues to the recognition of facial expressions

Mladen Sormaz, Andrew W. Young, Timothy J. Andrews *

Department of Psychology, University of York, York YO10 5DD, UK

A R T I C L E I N F O

Article history:
Received 11 December 2015
Received in revised form 7 June 2016
Accepted 7 July 2016
Available online 18 July 2016

Keywords:
Face
Expression
Shape
Surface
Texture
Image

A B S T R A C T

Theoretical accounts of face processing often emphasise feature shapes as the primary visual cue to the recognition of facial expressions. However, changes in facial expression also affect the surface properties of the face. In this study, we investigated whether this surface information can also be used in the recognition of facial expression. First, participants identified facial expressions (fear, anger, disgust, sadness, happiness) from images that were manipulated such that they varied mainly in shape or mainly in surface properties. We found that the categorization of facial expression is possible in either type of image, but that different expressions are relatively dependent on surface or shape properties. Next, we investigated the relative contributions of shape and surface information to the categorization of facial expressions. This employed a complementary method that involved combining the surface properties of one expression with the shape properties from a different expression. Our results showed that the categorization of facial expressions in these hybrid images was equally dependent on the surface and shape properties of the image. Together, these findings provide a direct demonstration that both feature shape and surface information make significant contributions to the recognition of facial expressions.

1. Introduction

The human face has a complex musculature that allows it to create a remarkable variety of facial expressions (Du, Tao, & Martinez, 2014). Although there are individual differences between people in the precise anatomical arrangement of the facial muscles, those muscles involved in producing facial expressions of what are considered to be basic emotions (which include happiness, sadness fear, anger, and disgust) are highly consistent across individuals (Waller, Cray, & Burrows, 2008). These muscles allow a person to move critical expressive features such as the eyebrows, eyes, nose and mouth in ways that can change their shapes (e.g. raising or lowering the corners of the lips, widening or narrowing the eyes), their positions (raising or lowering the eyebrows), or often both (wrinkling the nose, or lowering the jaw to open the mouth).

Despite this well-known anatomical background, the nature of the visual information underlying recognition of facial expressions is poorly understood. While an obvious place to begin looking for critical visual cues might seem to be in the patterns of movement themselves, these are difficult to define and the good recognition of photographs of normal intensity basic emotions shows that the apex of a set of muscle contractions often creates an easily recognisable expressive configuration of the facial features. Moreover, notational systems such as the Facial Action Coding System (FACS: Ekman & Friesen, 1978) depend on the fact that the underlying pattern of muscle contractions that create an expression is evident even in a static image. Many studies therefore begin by exploiting the recognisability of well-validated photographs of facial expressions such as the Ekman and Friesen (1976) series, as we do here.

There are many ways of thinking about the visual information conveyed by a photograph of a face, but one that has proved very useful is in terms of its shape and surface properties. Any facial image consists of a set of edges created by abrupt changes in reflectance due to the shapes and positions of facial features and a broader pattern of reflectance based on the surface properties of the face – also known as texture or albedo (Bruce & Young, 1998, 2012). Shape properties can be operationally defined by the spatial locations of fiducial points that correspond to facial features; note that in this sense 'shape' properties will include both the feature shapes and their positions. In contrast, surface properties result from the pattern of reflectance of light due to the combination of ambient illumination, the face's pigmentation, and shape from shading cues.

* Corresponding author.
E-mail address: timothy.andrews@york.ac.uk (T.J. Andrews).

http://dx.doi.org/10.1016/j.visres.2016.07.002
0042-6989/

The distinction of shape from surface properties is widely used in face perception research (Bruce & Young, 1998, 2012) and is implicit in standard approaches to computer image manipulation (Tiddeman, Burt, & Perrett, 2001). These image manipulation techniques allow quasi-independent changes to a face's shape or surface properties. Such changes cannot be fully independent, of course, because many of the shape and surface properties of images will necessarily covary. For example, the surface property of shading is clearly affected in part by the face's shape. None the less, such methods allow us to hold face shape fixed as closely as possible (by using the same fiducial positions for a set of images) or to hold the surface properties fixed as closely as possible (by using the same surface brightness patterns in a set of images). This then allows a direct test of the relative contributions of shape and surface information. Studies based on this approach have demonstrated independent contributions of shape and surface properties to the perception of a range of facial characteristics including gender, age, attractiveness and dominance (Burt & Perrett, 1995; Russell, 2003; Torrance, Wincenciak, Hahn, DeBruine, & Jones, 2014).

Thinking of facial images as broadly consisting of shape (feature positions) and surface (pigmentation, shading patterns) properties has also helped our understanding of facial identity recognition, where it is clear that both shape and surface cues can contribute (Russell, Sinha, Biederman, & Nederhouser, 2006; Troje & Bülthoff, 1996), but that the role of surface cues becomes more salient for familiar faces (Burton, Jenkins, Hancock, & White, 2005; Russell & Sinha, 2007).

In contrast to the established role of surface cues in the perception of facial identity, judgements of expression are often thought to be based primarily on the shapes and positions of critical expressive features such as the eyebrows, eyes, nose and mouth. This makes sense because these shape changes are a direct consequence of facial muscle movements. Evidence for the primary importance of shape cues in facial expression recognition comes from contrast reversal (as in a photo negative). In a contrast-reversed image the edges that define feature shape properties remain in the same positions, despite the huge change in overall surface properties. Although contrast negation is well-known to be very disruptive of facial identity recognition (Bruce & Young, 1998, 2012), it turns out that judgements based on facial expression are still possible in contrast-reversed images (Bruce & Young, 1998; Harris, Young, & Andrews, 2014a; Magnussen, Sunde, & Dyrnes, 1994; Pallett & Meng, 2013; White, 2001). Similarly image manipulations that completely remove surface information, such as line drawings of faces, also show relatively preserved expression perception (Etcoff & Magee, 1992; McKelvie, 1973). Using such evidence, most current accounts posit shape information to be the most important cue in the perception and recognition of expression (Bruce & Young, 2012; Calder, Young, Perrett, Etcoff, & Rowland, 1996).

Although previous studies have suggested that feature shape is the dominant cue for the perception and recognition of facial expressions, there are grounds for thinking that surface information might also play a role (Benton, 2009; Calder, Burton, Miller, Young, & Akamatsu, 2001). For example, Benton (2009) found a decrease in the emotional expression aftereffect to facial expressions when images were negated, suggesting that the perception of facial expression can be affected by changes in surface information. Using Principal Component Analysis (PCA), Calder et al. (2001) found that principal components (PCs) that convey variation in surface information could be used to categorize different facial expressions, albeit to a lesser extent than PCs that convey variation in shape. However, while these findings show a potential role for surface cues, they do not provide a direct test of whether surface properties are actually used for the recognition of facial expression. None the less, there are obvious ways in which surface properties might be useful to facial expression recognition. For example the feature shape change of opening the mouth will be accompanied by a bright region if the teeth are bared or a relatively dark region if the teeth are retracted; these different surface brightnesses are a direct reflection of muscle movements that clearly convey different expressions. Moreover, there are also indirect effects of underlying muscle movements such as the skin folding around the mouth and eyes resulting from smiling. These changes do not correspond to specific facial features, and are largely evident from their impact on surface shading patterns.

The aim of the current study was therefore to investigate the contribution of changes in the shapes of key expressive features (such as the eyebrows, eyes, nose and mouth) and changes in surface brightness patterns (such as those resulting from showing the teeth, or furrowing the brow) to the categorization of facial expression. In Experiment 1, we manipulated images to create facial expressions that varied primarily in shape or primarily in surface cues. This was achieved by reshaping images of different expressions to standardise the locations of the fiducial positions across the images, or by standardising the surface properties as far as possible by overlaying the same averaged surface onto the fiducials that characterise each expression. Because many of the shape and surface properties of images will necessarily covary, this method does not orthogonally manipulate shape and surface information, but it does allow us to hold the shape fixed as closely as possible (by using the same fiducial positions for all images) or to hold the surface properties fixed as closely as possible (by using the same surface brightness patterns in all images). This then allows a direct test of whether the information that remains free to vary across images can actually be used for the categorization of facial expression. In Experiment 2, we used contrast-reversed versions of the images used in Experiment 1 to further probe the role of shape and surface properties in the recognition of facial expressions. In Experiment 3, we then created hybrid images that combined the surface properties from one expression with the shape of a different expression. This approach offers a complementary method to that used in Experiment 1 and 2 for determining the relative contribution of surface and shape cues to the categorization of facial expressions.

2. Experiment 1

2.1. Method

2.1.1. Participants

Participants (n = 20, female = 10, mean age = 24.8 years, SD = 3.8) were drawn from an opportunity sample of students and staff at the University of York. Participants gave informed consent and were paid or given course credit for their participation. All data were collected in accordance with the ethical guidelines determined by the Psychology Department of the University of York and were in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki).

2.1.2. Stimuli

Fig. 1 and Supplementary Fig. 1 shows the stimuli for the three conditions used in Experiment 1 (original, shape varying, and surface varying; these are the 5 $\times$ 5 image matrices that form the left-most columns in Fig. 1). Static images of expressions were presented as these are well-recognised as long as they represent the apex of the pattern of muscle movements involved in producing the expression (see Bruce & Young, 2012). Five models (females F5, F6, F8, males M1, M6) were selected from the FEEST set (Young, Perrett, Calder, Sprengelmeyer, & Ekman, 2002) of Ekman and