How to predict the global instantaneous feeling induced by a facial picture?

Arnaud Lienhard [a,b,*], Patricia Ladret [a,b], Alice Caplier [a,b]

[a] *Univ. Grenoble Alpes, GIPSA-Lab, F-38000 Grenoble, France*
[b] *CNRS, GIPSA-Lab, F-38000 Grenoble, France*

ARTICLE INFO

ABSTRACT

Picture selection is a time-consuming task for humans and a real challenge for machines, which have to retrieve complex and subjective information from image pixels. An automated system that infers human feelings from digital portraits would be of great help for profile picture selection, photo album creation or photo editing. In this work, two models of facial pictures evaluation are defined. The first one predicts the overall aesthetic quality of a facial image, and the second one answers the question "Among a set of facial pictures of a given person, on which picture does the person look like the most friendly?". Aesthetic quality is evaluated by the computation of 15 features that encode low-level statistics in different image regions (face, eyes, and mouth). Relevant features are automatically selected by a feature ranking technique, and the outputs of 4 learning algorithms are fused in order to make a robust and accurate prediction of the image quality. Results are compared with recent works and the proposed algorithm obtains the best performance. The same pipeline is considered to evaluate the likability of a facial picture, with the difference that the estimation is based on high-level attributes such as gender, age, and smile. Performance of these attributes is compared with previous techniques that mostly rely on facial keypoint positions, and it is shown that it is possible to obtain likability predictions that are close to human perception. Finally, a combination of both models that selects a likable facial image of good quality for a given person is described.

1. Introduction

1.1. Context

Social psychological studies have shown that people form impressions from facial appearance very quickly [1]. With the widespread use of digital cameras and photo sharing applications, selecting the best picture of a particular person for a given application is a time-consuming task for humans. Thus, an automated system providing a feedback about facial images would be an interesting and useful tool. Sorting images automatically, editing images to enhance their visual aspect or selecting a few images among an entire collection would be simplified for home users. Generally, images with low aesthetic quality are manually rejected whereas appealing images are selected.

In the particular case of facial pictures, features have to be adapted to the considered use: profile pictures on social networks are different from pictures presented in a pro-fessional purpose (resumes and visiting cards). To this end,

* Corresponding author.
E-mail addresses: arnaud.lienhard@gipsa-lab.grenoble-inp.fr
(A. Lienhard), Patricia.Ladret@gipsa-lab.grenoble-inp.fr (P. Ladret),
Alice.Caplier@gipsa-lab.grenoble-inp.fr (A. Caplier).

this work focuses both on predicting the overall aesthetic quality of a facial image and selecting images that infer a feeling of likability. Ideally, the models developed in this work should encode relevant information about the global image aesthetics adapted to facial pictures as well as information related to facial expressions and high-level attributes (smile, age, gender, etc.). Facial beauty is not considered at all: the main idea is to estimate the feeling induced by a given facial picture (especially in terms of likability), which is not necessarily correlated with the reality.

1.2. Previous work

1.2.1. Aesthetic photo quality evaluation

Automated aesthetic evaluation of facial pictures is a challenging task that requires to understand subjective notions that are implicitly encoded in the image. To solve this problem, different approaches exist. In most of recent works, a large number of features describing the image aesthetics are extracted and machine learning algorithms are applied to fit the feature values to ground truth images obtained from human evaluation. Features can either be explored at pixel level (e.g. Fisher vectors) [2] or by estimation of high-level attributes (smiles and eyes closeness) [3,4] that are closer to human interpretation. To encode both local and global information into the models, the main approach for evaluating portraits aesthetic quality is characterized by computing a set of features about the subject (face) and background (non-face) regions. Often, low-level image statistics such as contrast, sharpness or color distribution are computed in addition to features that describe subject–background relationship [5,6].

To the best of our knowledge, few researches have been done on the particular case of pictures containing a unique and centered frontal face [7]. Plus, there are no publicly available datasets containing facial images and human aesthetic ratings, which makes comparison with previous models difficult. In previous work [8], we developed a method that segments precisely a portrait (hair, shoulders, skin, and background) and computes features in each region. The main result of this previous work is that facial area is almost sufficient to describe efficiently the global aesthetic of the entire facial picture. This idea is exploited in the proposed work, where features are extracted in small and informative facial areas (eyes and mouth).

1.2.2. Likability evaluation

The feeling induced by a facial picture depends on facial expression, face shape and other cues such as make-up or face adornments. However, state-of-the-art face evaluation systems do not consider many of these attributes. A first attempt to create a data-driven model of several evaluation traits is discussed in [9], in which 300 faces are generated by the Facegen Modeller software (http://www.facegen.com) with different shape parameters. A subjective experiment is conducted, where participants evaluate each face with respect to a particular trait: aggressiveness, attractiveness, threat, etc. Finally, shape parameters are fitted to the ground truth scores provided by participants to build a regression model for each social judgment.

Besides, behavioral studies have shown that facial image quality estimation does not only rely in face shape and that reflectance (cues such as skin illumination and texture) also plays an important role in face perception [10]. A more complete model including reflectance parameters is elaborated and validated in [11]. However, the faces considered in all their experiments are synthetic and without facial hair, make-up or accessories. Real 3-D scanned faces have been used in [12] to identify relevant shape and reflectance features. Even in recent attempts of automated face expression evaluation in videos [13], the use of facial keypoints is still predominant. The disadvantage of these models is that it only takes into account the position of facial keypoints and reflectance parameters. Plus, facial keypoints are heavily related to the face shape whereas our goal is to predict the most likable image of a given person which is not entirely defined by the face shape.

High-level attributes are defined as abstract and global concepts describing an image. They correspond to descriptors that cannot directly be obtained by extracting visual data due to the semantic gap between information contained in pixels and human analysis. Many attributes (age, gender, presence of glasses, beard, smile, etc.) have already been successfully used in various research domains such as face recognition or verification [14] and portraiture aesthetics [3]. A small set of such attributes provides more significant information than the relative positions of many facial keypoints.

1.3. Objectives

The main contribution of this paper is to propose the first model that combines both aesthetic quality assessment and likability estimation for frontal facial pictures, in order to perform automatic picture selection. For each criterion (aesthetic quality and likability), a model that outperforms state-of-the-art methods is presented, and the most relevant features are described.

Aesthetic quality of facial pictures is evaluated using the same feature set than in our recent work [15]. The difference relies on the use of 4 learning algorithms that are combined to provide a more accurate and robust prediction, which outperforms our previous results. This work also focuses on demonstrating the advantages of using high-level attributes in order to build likability evaluation models. 3 tools are considered to compute the attributes: Betaface (http://betaface.com), SkyBiometry (http://skybiometry.com) and SHORE [16]. It is shown that for real images, these features are significantly more efficient to predict likability.

This work is organized as follows. Section 2 describes the main steps of the proposed method, including feature computation and selection and learning algorithms. Section 3 demonstrates the relevance of the algorithm combination and compares the results with previous recent works. In Section 4, the same pipeline is applied to perform likability evaluation. The major difference between aesthetic quality and likability evaluation is the feature extraction process, which rely either on low-level statistics or on high-level attributes. Finally, both predictions are combined in Section 5 to perform automatic