A Computational Approach to Finding Facial Patterns of a Babyface

Zi-Yi Ke
Department of Computer Science and Information Engineering
National Taiwan Normal University
40147024s@ntnu.edu.tw

Mei-Chen Yeh
Department of Computer Science and Information Engineering
National Taiwan Normal University
myeh@csie.ntnu.edu.tw

ABSTRACT

Facial babyishness has a strong impact on social perceptions and interactions; however, the components constituting a babyface remain unclear. In this paper, we present a computational approach for identifying important but less apparent facial patterns of a babyface, using voluminous face images on the web. The proposed approach is built upon computationally efficient data mining techniques. A new image set with ground truth data collected from users and an evaluation approach based on age estimation are presented in the experiment. The results show that the mined patterns are effective for understanding and determining babyfaces. The findings of this study should provide information for future investigations on the prediction and analysis of trait impressions using the patterns.

CCS Concepts

•**Information systems** → *Image search;* •**Computing methodologies** → *Object recognition;*

Keywords

Babyface recognition, facial landmark, visual mining

1. INTRODUCTION

A wealth of literature has revealed that our impressions of unknown people can be influenced by a variety of non-behavioral cues. In particular, the physical appearance has a strong impact on social perceptions and interactions, which may have important implications for personality development [6]. Facial appearance especially matters because some facial qualities are so useful in guiding adaptive behavior that even a trace of those qualities can elicit a response [12]. One of such appearance qualities is facial babyishness. Baby-faced people are perceived as more socially dependent, intellectually naïve, physically weak, honest, and warm than their mature-faced peers [2]. Moreover, baby-faced people of all ages and both sexes experience social outcomes that are consistent with their perceived traits [7]. A recent news report reveals that they even live longer, regardless of gender and environment.

According to the Oxford English Dictionary, a babyface refers to a smooth round face like a baby's. To understand the perceptions of baby-faced individuals from infancy to older adulthood, an interesting question is: What facial features are associated with babyface ratings? Psychological studies have found that large, round eyes, high eyebrows, and a small chin each yields the perception of a babyish facial appearance [2, 7, 12]. The findings were obtained through the conduction of small-scale experiments. For example, Berry et al. used twenty stimulus slides of face images of white male college students and invited forty male and forty female undergraduates to rate the slides [2]. A face image was characterized by only eleven features, such as eye size, eye shape, eyebrow height, distance between eye centers and etc. Subjects rated the babyishness of the faces using a 7-point bipolar scale. The relations between babyfacedness ratings and each physiognomic measure (i.e., facial feature) were analyzed.

However, combinations of babyish facial cues often influence impressions more than individual cues. Modifying one facial feature can alter the spatial arrangements among all features, which may in turn impact impressions. Therefore, it is important to consider not only individual facial features but also their arrangements. Because of the limited amount of data available in the psychological experiments, the existence of such a configuration remains unclear.

With the number of web images continuing to grow, the images can be an important and valuable source for investigating further the babyface recognition problem. In this paper, we propose a computational, data-driven approach utilizing voluminous face images on the web to discover and identify facial patterns (e.g., facial features, landmark configurations, golden ratios) of a babyface. In particular, we construct a babyface image dataset, from which frequent and unique facial patterns associated with babyfaces are discovered. We describe a face image by a set of facial landmarks and design physiognomic measures in a way suitable for data mining. We adopt a frequent itemset mining algorithm [1] that is capable of efficiently analyzing the large amount of face images and returning spatial configurations of facial fiducial points frequently re-occurring over the training images. Based on the mined patterns, we can automatically determine if a person has a babyface or to be more mature-looking, and can adjust the spatial layout of facial components to increase the babyishness of a face image.

ACM ISBN 978-1-4503-2138-9.
DOI: 10.1145/1235

The remainder of the paper is organized as follows. We review related work in section 2. Next, section 3 describes the construction of a babyface image dataset. In section 4 we present the approach to mining frequent spatial configurations of facial landmarks from babyface images. A user study designed for experimental evaluation is presented in section 5. Finally, we conclude the paper with a short discussion summarizing our findings.

2. RELATED WORK

Particular facial qualities contain rich social contextual information, which have been extensively studied in social psychology but not fully explored in computer science. In this section we briefly review the literature that uses psychological experiments to analyze the facial quality of a babyface. Then we present some computational approaches using mid-level attributes for representing images, and in particular, facial appearance.

Berry and McArthur presented physical measurements and subjective ratings of various facial features for 20 adult male stimulus faces [2]. The faces were rated on five personality dimensions, physical attractiveness, age, and babyfacedness. The study suggested that the physical measurements of large, round eyes, high eyebrows, and a small chin each yielded the perception of a babyish facial appearance, and a weighted linear composite derived from the measures of eye size and chin width accounted for 57% of the variance in the ratings of babyfacedness. A similar finding was presented in [11], in which the experiments included a slightly extensive set of targets representing one of six age groups: infants, pre-schoolers, fifth graders, eighth graders, young adults, and older adults. The authors concluded that large eyes, a round face, thin eyebrows, and a small nose bridge characterized a babyface for male and female targets across the life span. Because of the limited amount of data available in the psychological experiments, the discovered physical measurements are simple, involving only a few individual facial features.

Over the past few years, several studies in computer science concerned the representation of objects using meaningful intermediate features, i.e., attributes. Several papers have applied the idea for describing facial appearance [5]. Describing people by semantic facial attributes is intuitive; however, in order to train a set of models for recognizing individual facial attributes, a large training set must be manually labeled for each attribute at high cost. In this preliminary study, we treat facial qualities as attributes of faces, and the representation may jointly model age and other unnamed appearance attributes associated with the people having similar facial qualities.

3. DATASET COLLECTION

This section details the construction of a babyface image dataset, from which facial patterns associated with babyfaces are identified. The dataset is only used for training the approach. The evaluation uses a completely different dataset (discussed in section 5).

First, we compiled a list of 40 generally recognized baby-faced celebrities (i.e., each occurred frequently on news and magazine articles describing celebrities with a permanent baby face), including both western and eastern baby-faced people. A preliminary dataset was collected from several image search websites using the baby-faced celebrity names as query. Subsequently, we applied a geometric method [3] to estimate head pose orientations characterized by pitch, roll, and yaw angles, and retained only the face images in frontal or near frontal view (10° rotation off the image plane both for pitch and yaw angles). The removal of non-frontal images helps eliminate the facial pattern variations caused by pose.

Figure 1: A few samples in our babyface image dataset, collected by image search engines using names of baby-faced celebrities as query.

Finally, we have a clean babyface image dataset containing 2178 frontal or near-frontal images of 40 identities—21 males (52.5%) and 19 females (47.5%)—between the age group of 22-45 years. A few examples in the dataset are shown in Fig. 1.

4. APPROACH

Recall that the objective of this study is to identify facial patterns that constitute a babyface from a collection of babyface images. This problem can be considered a type of frequent itemset mining problem, where an image characterized by a set of facial patterns is analogous to a transaction of items in market basket analysis. The technical details are explained in the following subsections.

4.1 Pattern Formation

A facial pattern (a scalar) is formed by describing the spatial configuration of facial landmarks on a face, which contains at least the following types: (1) the distance of two points, (2) the ratio of two lines, (3) the area of a polygon, and (4) the ratio of two areas. To transfer patterns into items, each pattern is quantized into M bins, using a uniform scalar quantizer [9] where the decision boundaries are determined by the pattern distribution derived from the babyface dataset. We observe that the facial patterns, in general, follow a normal distribution. Therefore, we use the mean and standard deviation of the distribution to determine the decision boundaries.

In the implementation, we used the facial landmark detector [13] to extract 83 facial fiducial points from a face image. Concerning a face representation using meaningful interme-

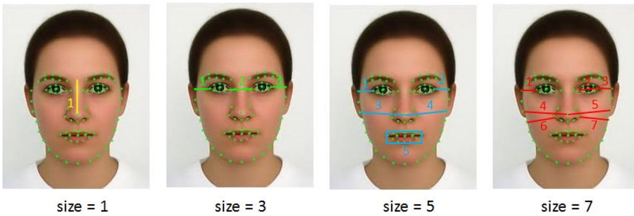

Figure 2: Frequent and distinctive facial patterns discovered from the babyface image dataset and LFW. The mined itemsets have semantic qualities.

diate features, we designed 50 physiognomic facial patterns referring to [2], each explicitly or implicitly described the size, shape, relative positions, relative areas of the main facial features, e.g., eye, nose, and etc. A pattern was quantized into seven bins ($M = 7$). As a result, we extracted 50 facial patterns from an image and each was indexed by one of the 350 (50×7) possible items.

4.2 Pattern Mining and Scoring

Next, frequent and distinctive patterns are discovered as follows. First, a transaction database D is built from the babyface image dataset (section 3). Each pattern is considered as an item; therefore, a transaction (i.e., a babyface image) is simply an unordered collection of 50 items (i.e., facial patterns). An itemset A is called frequent in D if the support of A is larger than or equal to a threshold. Since frequent itemsets are subject to the monotonicity property, the APriori algorithm [1] can be used to efficiently find the frequent itemsets.

Frequent facial patterns may not be distinctive. A distinctive pattern should appear frequently on babyfaces and rarely on mature faces. Therefore, we refer to the mining approach [8] and add transactions from a more general training set (containing not necessarily babyfaces) to the database. All transactions originating from the babyface database are assigned a label *babyface* as an additional item, while we append the item *others* to the alternative transactions. We filter the association rules and keep only those which infer babyface with high confidence:

$$conf(P \rightarrow babyface) > conf_{min}, \quad (1)$$

where P is a frequent facial pattern set and $conf_{min}$ is a confidence threshold. Rules do not have a high confidence if they appear frequently on both training sets. We further remove the itemsets if they are entirely contained in another one. In the implementation, we used a subset of frontal or near-frontal faces from the LFW dataset [4] to be the general training set. The support and confidence thresholds were empirically set at 0.58 and 0.6, respectively.

Each discovered itemset contains one or a few facial patterns, and each pattern is associated with a specific range. For example, the left-most image in Fig. 2 shows one of the discovered itemset, which contains only a facial pattern—the ratio of face height and nose length—, with a range of 3.33 ± 0.26. It indicates that the face length tends to be 3.33 times the nose length or, alternatively, the nose of a babyface is approximately one third of the face's length. Figure 2 shows a few more examples of the discovered itemsets with various sizes. As additional advantage, many of the mined itemsets have semantic qualities. For example, the third image (size = 5) in Fig. 2 specifies the eye locations, cheek width and mouth size (normalized by the face size).

Given a test image $\mathcal{I}$, we can now match the mined facial patterns to it and derive a measure of the babyfacedness of the image. Let P_i be the i-th discovered itemset and $|P_i|$ be the number of items in P_i. Every mined facial pattern P_i is matched to the test image $\mathcal{I}$ and the babyface measure is defined as:

$$f(\mathcal{I}) = \sum_i |P_i| \cdot m(\mathcal{I}, P_i), \quad (2)$$

where $m(\mathcal{I}, P_i) = 1$ indicates a match. The effectiveness of the measure is evaluated in the experimental section.

5. EXPERIMENTS

In this section we validate the usage of the mined patterns. In particular, we develop an evaluation approach based on age estimation and demonstrate an application of altering the babyishness of a face by adjusting the locations of facial parts guided by the mined patterns. The evaluation involves a user study as well because determining a babyface is subjective, and more importantly, ground truth data are unavailable.

5.1 Setup

We firstly applied our babyfacedness scoring function on 100 randomly selected images from the LFW dataset [4] (images used for training were excluded). The images were ranked based on its babyfacedness score. The top 15% and the bottom 15% ranked images (each is labeled with the name of the person pictured) were used for evaluation, containing 14 males and 16 females.

Next, a user study was conducted to build the ground truth data for the evaluation set. The user study involved an online survey that required participants to estimate the age of each selected person (30 people in total). We displayed a few clear images representing the person's recent appearance to a participant. A participant was requested to type in the estimated age of the persons. We computed the mean and standard deviation of the estimated age, which was subsequently compared with the chronological age assessed by recording date of birth. The survey was posted for a week and 57 responses were received.

5.2 Results

Figure 3 shows the comparison of the estimated (blue circles) and chronological (red dots) ages, where the x-axis represents the selected persons and the y-axis displays the ages. First, the age estimation task is difficult for humans, because the received estimated ages have large variations. Second, we observe that the celebrities used in the evaluation set have, in general, a smaller estimated age. A possible reason is that the celebrities wear their makeup when being pictured, and a successful makeup can make people look younger. Another possible reason is that not all of the test images are neutral faces. The age of faces displaying happy expressions is most likely underestimated [10]. However, if we examine individually those having a high babyfacedness score (bright area) and those with a low babyfacedness score (dark area), the automatically identified babyfaced people tend to look much younger than their current age. The av-

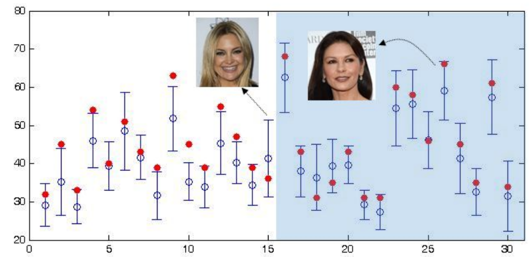

Figure 3: Individual age estimation results. The automatically identified baby-faced celebrities (with high babyfacedness score) are perceived younger than their current age.

Figure 4: The perceived babyishness can be affected by the facial feature arrangement; top row: original images, bottom row: altered images guided by the discovered patterns.

erage gap between the real and estimate ages for these two groups are 5.24 and 2.41, respectively. One exception is Kate Hudson. She is perceived five years older than her real age in our experiment, which we cannot explain.

Finally, we demonstrate an application on altering the babyishness of a human face. Figure 4 demonstrates three face pairs. The top row shows the original images. We change the spatial arrangement of the facial landmarks so that more mined facial patterns are present. The altered images are shown in the bottom row. Note that the size of each facial part remains unchanged. Although the difference between the face pairs is minor, the perceived babyishness is somehow different. Changing facial feature arrangement could alter the babyishness of a human face.

6. CONCLUSION

The paper presents a computational approach for identifying facial patterns of a babyface. Several important but less visible facial patterns are discovered through mining frequent and distinctive patterns from babyface images. We collect a new babyface image dataset that serves to train our approach, and an evaluation set with ground truth, which should both be broadly useful to the community studying related problems. The experiment show that the discovered patterns are effective for identifying baby-faced people. These patterns are very different from those previously discovered [2, 12], and should be useful for predicting and analyzing trait impressions in future psychological studies.

Future research can involve enlarging the pattern pool to include more sophisticated patterns describing the facial structure. The study primarily focuses on static appearance cues. Another research direction is the investigation of multiple modalities, including facial movement, voice and gesture, which may also convey the babyishness of a person.

7. REFERENCES

[1] R. Agrawal, T. Imielinski, and A. N. Swami. Mining association rules between sets of items in large databases. In *Proc. of ACM Intfl Conf. SIGMOD*, 1993.

[2] D. S. Berry and L. Z. McArthur. Some components and consequences of a babyface. *Journal of Personality and Social Psychology*, 48:312–323, 1985.

[3] T. Horprasert, Y. Yacoob, and L. Davis. Computing 3-d head orientation from a monocular image sequence. In *Proc. of IEEE Intfl Conf. Automatic Face and Gesture Recognition*, 1996.

[4] Gary B. Huang, Manu Ramesh, Tamara Berg, and Erik Learned-Miller. Labeled faces in the wild: A database for studying face recognition in unconstrained environments. Technical Report 07-49, University of Massachusetts, Amherst, October 2007.

[5] N. Kumar, A.C. Berg, P.N. Belhumeur, and S.K. Nayar. Describable visual attributes for face verification and image search. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 33(10):1962–1977, 2011.

[6] L. Z. McArthur. Judging a book by its cover: A cognitive analysis of the relationship between physical appearance and stereotyping. *Cognitive Social Psychology*, pages 149–211, 1982.

[7] J. M. Montepare and L. A. Zebrowitz. Person perception comes of age: The salience and significance of age in social judgments. *Advances in Experimental Social Psychology*, 30:93–161, 1998.

[8] T. Quack, V. Ferrari, B. Leibe, and L. V. Gool. Efficient mining of frequent and distinctive feature configurations. In *Proc. of IEEE Intfl Conf. Computer Vision*, 2007.

[9] K. Sayood. *Introduction to Data Compression, 4th Edition*. Morgan Kaufmann, 2012.

[10] M. C. Voelkle, N. C. Ebner, U. Lindenberger, and M. Riediger. Let me guess how old you are: Effects of age, gender, and facial expression on perceptions of age. *Psychology and Aging*, 27(2):265–277, 2012.

[11] L. A. Zebrowitz and J. M. Montepare. Impressions of babyfaced individuals across the life span. *Developmental Psychology*, 28(6):1143–1152, 1992.

[12] L. A. Zebrowitz and J. M. Montepare. Social psychological face perception: Why appearance matters. *Social and Personality Psychology Compass*, 2(3):1497–1517, 2008.

[13] E. Zhou, H. Fan, Z. Cao, Y. Jiang, and Q. Yin. Extensive facial landmark localization with coarse-to-fine convolutional neural network. In *Proc. of ICCV workshop on 300 Faces in-the-Wild Challenge*, 2013.