

 Open access • Proceedings Article • DOI:10.1109/ICB.2013.6613022

Age estimation from face images: Human vs. machine performance

— [Source link](#) 

Hu Han, Charles Otto, Anil K. Jain

Institutions: Michigan State University

Published on: 04 Jun 2013 - International Conference on Biometrics

Topics: Facial recognition system

Related papers:

- [Automatic Age Estimation Based on Facial Aging Patterns](#)
- [Human age estimation using bio-inspired features](#)
- [MORPH: a longitudinal image database of normal adult age-progression](#)
- [Age Synthesis and Estimation via Faces: A Survey](#)
- [Image-Based Human Age Estimation by Manifold Learning and Locally Adjusted Robust Regression](#)

Share this paper:    

View more about this paper here: <https://typeset.io/papers/age-estimation-from-face-images-human-vs-machine-performance-1bcjvxjidf>

Age Estimation from Face Images: Human vs. Machine Performance

Hu Han, Charles Otto, and Anil K. Jain

Department of Computer Science and Engineering
Michigan State University, East Lansing, MI, U.S.A.

{hhan, ottochar, jain}@cse.msu.edu

Abstract

There has been a growing interest in automatic age estimation from facial images due to a variety of potential applications in law enforcement, security control, and human-computer interaction. However, despite advances in automatic age estimation, it remains a challenging problem. This is because the face aging process is determined not only by intrinsic factors, e.g. genetic factors, but also by extrinsic factors, e.g. lifestyle, expression, and environment. As a result, different people with the same age can have quite different appearances due to different rates of facial aging. We propose a hierarchical approach for automatic age estimation, and provide an analysis of how aging influences individual facial components. Experimental results on the FG-NET, MORPH Album2, and PCSO databases show that eyes and nose are more informative than the other facial components in automatic age estimation. We also study the ability of humans to estimate age using data collected via crowdsourcing, and show that the cumulative score (CS) within 5-year mean absolute error (MAE) of our method is better than the age estimates provided by humans.

1. Introduction

Humans can glean a wide variety of information from a face image, including identity, age, gender, and ethnicity (See Fig. 1). The identification characteristic of face images has been well explored in real-world applications [34], including passports and driver licenses. Face mugshot retrieval is also a powerful way for law enforcement agencies to identify potential suspects in criminal investigations. Despite the broad exploration of person identification from face images, there is only a limited amount of research [25] on how to accurately estimate and use the demographic information contained in face images such as age, gender, and ethnicity.

For many practical applications, relying on humans to supply demographic information from face images is not



Figure 1. A wide variety of information that can be extracted from a face image, such as identity, age, gender, ethnicity, and scars, marks and tattoos (SMT).

feasible. Hence, there has been a growing interest in automatic extraction of demographic information from face images. Here we focus on age estimation, whose objective is to determine the specific age or age range of a subject based on a facial image. Some of the potential applications of automatic age estimation are: (i) *Law enforcement*: Automatic age estimation systems can help to determine the potential suspects more efficiently and accurately by filtering the gallery database using the estimated age of the input mugshot. (ii) *Security control*: An automatic age estimation system can be used to prevent minors from purchasing alcohol or cigarette from vending machines or accessing inappropriate web pages. (iii) *Human-computer interaction (HCI)*: The system can adjust the contents presented to a user based on her age. For example, a smart shopping chart can be designed to provide recommendations according to the age of the customer.

Unlike other sources of variation in facial appearance (lighting, pose, and expression) which can be controlled during face image acquisition, face aging is an unavoidable natural process. Moreover, face aging is affected not only by internal factors, but external factors as well [1].

1.1. Background

A number of studies in the biological, psychological, and cognitive sciences areas, have reported on how the human

Table 1. A Comparison of Published Methods for Automatic Age Estimation.

Publication	Face representation	Face aging database (#subjects, #images)	Human perception of age	Performance measure and accuracy
Lanitis <i>et al.</i> [26]	Holistic 2D shape and texture	Private (60, 500)	Studied on a subset with 32 images	MAE ¹ : 4.3 (Case2)
Hayashi <i>et al.</i> [18]	Texture and Wrinkle	Private (300, 300)	Studied	Hitting ratio ² of age group classification: 27%
Iga <i>et al.</i> [22]	Gabor, color, texture and intensity	Private (101, 101)	Studied	Hitting ratio of age group classification: 58.4%
Geng <i>et al.</i> [11]	Holistic appearance PCA	FG-NET (82, 1002) MORPH (NA, 433)	Studied on a subset of FG-NET with 51 images	FG-NET/MORPH MAE: 6.8 / 8.8 CS ³ : 65% / 46%
Fu and Huang [9]	Holistic appearance, Manifold	Private YGA (1600, 8000)	Not studied	MAE: 5~6 CS: F: 55%, M: 50%
Suo <i>et al.</i> [36]	Holistic and local topology, 2D shape, color, and gradient	FG-NET (82, 1002) Private (NA, 8000)	Studied with 500 images from the two databases	FG-NET/Private MAE: 6.0 / 4.7 CS: ~55% / ~66%
Guo <i>et al.</i> [13]	Holistic BIF	FG-NET (82, 1002) Private YGA (1600, 8000)	Not studied	FG-NET/YGA MAE: 4.8 / F: 3.9, M: 3.5 CS: 47% / F: 75%, M: 80%
Li <i>et al.</i> [27]	Local patch, code book	Private YGA (1600, 8000)	Not studied	MAE: 8.6 CS: 30.0%
Guo and Wang [12]	Holistic BIF, Kernel PLS	MORPH II (NA, ~55000)	Not studied	MAE: 4.2, CS: NA
Choi <i>et al.</i> [4]	Holistic appearance, Gabor, LBP	FG-NET (82, 1002) PAL (NA, 430) Private BERC (NA, 390)	Not studied	FG-NET/ PAL /BERC MAE: 4.7 / 4.3 / 4.7 CS: 73% / 70% / 65%
Luu <i>et al.</i> [28]	Holistic contourlet appearance model	FG-NET (82, 1002) PAL (NA, 443)	Not studied	FG-NET/PAL MAE: 4.1 / 6.0 CS: ~73% / NA
Chang <i>et al.</i> [2]	Ordinal hyperplane ranking	FG-NET (82, 1002) MORPH II (NA, 5492)	Not studied	FG-NET/MORPH II MAE: 4.5 / 6.1 CS: 74.7% / 56.5%
Guo and Wang [14]	Holistic BIF, PLS	PAL (590, 844) FACES (171, 1026)	Not studied	PAL/FACES MAE: 6.1 / 8.1
Wu <i>et al.</i> [40]	Grassmann manifold of Facial shape	FG-NET (82, 1002) Passport (109, 233)	Not studied	FG-NET/Passport MAE: 5.9 / 8.8 CS: 62% / 40%
Thukral <i>et al.</i> [37]	Landmark based hierarchical approach	FG-NET (82, 1002)	Not studied	MAE: 6.2
Chao <i>et al.</i> [3]	Label-sensitive relevant component analysis	FG-NET (82, 1002)	Not studied	MAE: 4.4
Proposed	Component and holistic BIF	FG-NET (82, 1002) MORPH II (20569, 78207) PCOS ⁴ (1802, 10036)	Studied with FG-NET and 2200 images from PCSO	FG-NET/MORPH II/PCSO MAE: 4.6 / 4.2 / 5.1 CS: 74.8% / 72.4% / 64.0%

¹MAE (mean absolute error) [26] is the average of the absolute difference between estimated and real ages. ²Hitting ratio [18] is the rank-1 age group classification accuracy. ³CS (cumulative score) [11] reflects the percentage of correct age estimations with different absolute errors. In the table, we only give the CS within 5-year absolute error. ⁴Interested researchers may contact the Pinellas County Sheriff's Office (PCSO) to access this database.

brain perceives, represents, and remembers faces. In particular, various aspects of human age estimation have been studied in the field of psychology (Rhodes provides a review [31]). Psychological studies often have the goal of examining the effects of a subject's age, gender, and race on the accuracy of age estimation that the subject provides [39]. These studies provide some context for the performance of automatic age estimation methods reported in the literature;

however, the accuracy of age estimation by human subjects on a large scale has not been reported for most databases used in automatic age estimation research.

Computational models based on the above research have also been proposed to provide insight into the problem of automatic demographic information estimation. Fu *et al.* provided a review of many of the existing methods until 2010 [8]; however, several novel approaches have recently

been proposed. In this paper, we provide a brief comparison of major approaches for automatic age estimation. Table 1 shows that existing approaches either represent faces with a holistic representation (e.g. [11, 13]) or use local features, e.g. hair, wrinkles, and mustache for age estimation. However, there is only a limited amount of analysis on how aging influences individual facial components [30, 36]; the study in [30] was limited to 0-10 year age gaps. Further, to our knowledge, no large scale studies on the human ability to estimate age have been conducted on public-domain face aging databases (e.g. FG-NET). Human performance can provide an interesting baseline for the age estimation task, since although we don't expect humans to predict ages with perfect accuracy, we do expect human estimates to at least fall within some broad age range of the true age (e.g. a human is unlikely to say a child is an adult).

1.2. Motivation

Face aging is determined by both intrinsic factors (e.g. human genes) and extrinsic factors (e.g. work environment, lifestyle, and health). The characteristics of facial aging include a gradual change in appearance which is not visibly apparent over a short age gap. Based on this observation, a hierarchical age estimator [4, 37] is proposed for automatic age estimation. Unlike [4, 37], where the whole age range is directly partitioned into multiple age groups, in this paper, we use a binary decision tree based on SVM (SVM-BDT) [38] to perform age group classification. Within each age group, a separate SVM age regressor is trained to predict the final age. In order to mitigate the age estimation errors due to incorrect age group classifications, we use overlapping age ranges while training the regression functions. With the hierarchical age estimator, we perform automatic age estimation using a component based representation (forehead, eyebrows, eyes, nose, mouth, shape, and holistic face), and analyze the influence of aging on individual facial components.

The main contributions of this paper are as follows. (i) A hierarchical age estimation approach is proposed for automatic age estimation, and a component based representation is used to analyze the aging process of each facial component separately. (ii) Human perception ability in age estimation is studied using crowdsourcing which allows a comparison of the ability of machines and humans.

2. Face Aging Databases

Table 1 indicates that most face aging studies have evaluated their approaches on the public domain FG-NET database. However, the age distribution of FG-NET is biased significantly to children. Another public domain database that has been used for age estimation is the MORPH database [32]. A few studies have reported results on the MORPH Album2 data set, which contains 55,000

mugshot images. We performed experiments on the FG-NET and MORPH Album2 databases, and additionally on a 10,036 face image subset from a 1.5 million mugshot database available from the Pinellas County Sheriff's Office (PCSO), constructed so that faces in it have a uniform age distribution in the range [17, 68]. FG-NET is from personal photo collection, but MORPH Album2 and PCSO are databases from law enforcement agencies. Details of the three databases that we have used are provided below.

2.1. FG-NET

FG-NET¹ consists of 1,002 images of 82 individuals. The average number of images per individual is 12. Our experiments on FG-NET used the entire dataset, using a Leave-One-Person-Out (LOPO) protocol, similar to prior work on age estimation. Although the age of subjects in FG-NET ranges from 0-69 years, over 50% of the subjects in FG-NET are between the ages 0 and 13.

2.2. MORPH Album2

MORPH² is a database of mugshot images, with associated metadata giving the age, ethnicity, and gender of each subject in the database. We performed experiments on MORPH Album2, with a version of the database containing 78,207 images of 20,569 subjects. For our experiments, we randomly selected subjects (using all images of each subject selected) until we had a total of 10,001 images. We used these 10,001 images as a training set, and evaluated our method's performance on the remaining 68,206 images in the dataset. Subjects in MORPH Album2 range from 15-77 years old, although the number of images per age drops off above 50 years old. MORPH Album2 is a much larger database than FG-NET, and has a different age distribution since it is comprised primarily of adults, and contains no images of young children.

2.3. PCSO

This mugshot dataset was acquired from the Pinellas County Sheriff's Office (PCSO). Similar to MORPH Album2, this database contains mugshot images with associated metadata including a capture date for each image, and a date of birth for each subject. The complete PCSO data set contains some 1.5 million images, out of which we sampled a subset of 10,036 images covering an age range of 17-68, with exactly 193 images per age. Although both MORPH Album2 and FG-NET cover wide age ranges, their age distributions are far from uniform. Experiments on this subset of the PCSO dataset allow us to examine the effect of a uniform age distribution (a more challenging scenario, since in this case we cannot sacrifice performance on the upper end of the age range to improve overall performance) on our age

¹<http://www.prima.inrialpes.fr/FGnet/html/benchmarks.html>

²<http://www.faceaginggroup.com/projects.html>

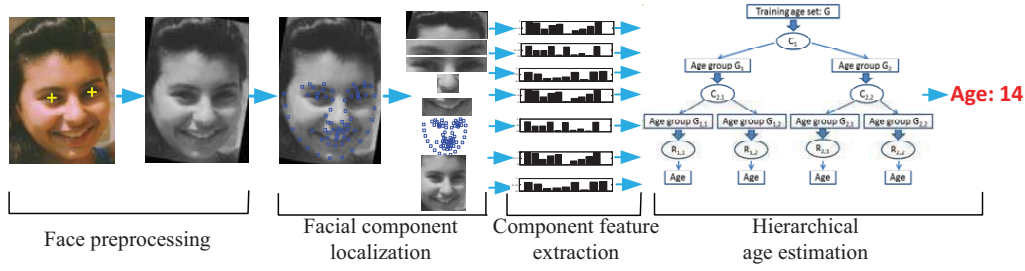


Figure 2. Overview of the proposed component based representation and hierarchical estimator for automatic age estimation.

estimation method while still maintaining a relatively large dataset. Results are reported for the entire PCSO subset using a leave-one-fold-out (totally 5 folds) protocol.

3. Proposed Age Estimation Method

We propose a component-based representation for age estimation from face images. As illustrated in Fig. 2, the proposed approach consists of four main steps: preprocessing, facial component localization, feature extraction, and hierarchical age estimation.

3.1. Face Preprocessing

Images in some of the aging face databases, e.g. FG-NET, may have been captured using different methods, including scanned photographs, digitized film, or digital cameras. As shown in Fig. 6, face images in FG-NET can be either in gray-scale or color, and some of the color images even have a color cast. To mitigate the influence of inconsistent colors, we first convert all color face images into gray-scale³. There are also in-plane and out-of-plane face rotations. In order to improve the accuracy of facial component localization, a nonreflective similarity transformation is applied to normalize each face image based on two eyes. In the FG-NET database, 68 landmarks (including the eye centers) are provided for each face image. For the PCSO and MORPH Album2 face databases which do not have labeled landmarks, we automatically detect the eye centers using the FaceVACS SDK [5]. The normalized face images are then cropped to the same size and interpupillary distance (IPD) (see Figs. 2 and 3).

3.2. Facial Component Localization

Component-based face recognition methods have been reported in [15, 19, 20, 23]. However, these algorithms were mainly proposed to resolve the misalignment problem in face matching, and did not address the age estimation problem. We propose an efficient method to localize individual facial components which is based on ASM [6] to automat-

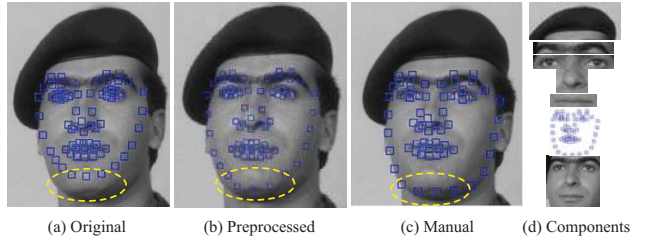


Figure 3. Facial component localization for one image from the FG-NET database. (a) Facial landmarks detected in the original face image; (b) Facial landmarks detected in the preprocessed face image, which are closer to the true landmark locations; (c) Manual facial landmarks; and (d) Localized facial components based on the landmarks in (c).

ically detect a set of facial landmarks⁴, followed by component localization based on subsets of landmarks corresponding to the forehead, eyebrows, eyes, nose, and mouth. We also utilize the vectorized face shape coordinates⁵ and the whole face image as two additional components in our experiments. Fig. 3 shows that face preprocessing improves not only the accuracy of ASM in detecting facial landmarks, but also the localization accuracy of facial components. Per-component scaling is applied to compensate for differences in absolute component sizes.

3.3. Feature Extraction

Following the theory of a feedforward path in the cortex [33], Serre *et al.* [35] proposed a biologically inspired model (BIM) for robust object recognition, which has also been found to be effective in automatic age estimation [13]. Instead of extracting biologically inspired features (BIF) from a holistic face like in [13], we extract BIF features from individual facial components.

In its simplest form, the BIM consists of two layers of computational units, where simple S_1 units are followed by complex C_1 units. The S_1 units correspond to the classical

³Additional methods for handling illumination variations, e.g. [16, 17], will be investigated in our future work.

⁴The open source software Stasm [29] with 76 predefined facial landmarks is utilized as the landmark detector. For the FG-NET database, we directly use the manual facial landmarks provided with each face image.

⁵The coordinates of facial keypoints have already been normalized to the same scale in the face preprocessing step.

simple cells of Hubel and Wiesel found in the primary visual cortex (V1) [21]. They are usually implemented with the real components of Gabor filters [10] which have been shown to provide a good model of cortical simple cell receptive fields [24]. We determine the parameters, e.g. aspect ratio γ , the effective width σ , and the wavelength λ , the orientation θ , and filter scale s as in [35], but we use 8 orientations and 12 scales for the Gabor filters in our automatic age estimation task.

The C_1 units correspond to cortical complex cells which are robust to shift and scale variations. The C_1 units can be calculated by pooling over the preceding S_1 units with the same orientation but at two neighboring scales. We apply “MAX” pooling operator and “STD” normalization (MAX-STD) to extract C_1 features from the S_1 layer.

3.4. Hierarchical Age Estimation

We propose a hierarchical age estimation method as shown in Fig. 2. Each facial component is first classified into one of four disjoint age groups using a binary decision tree based on SVM (SVM-BDT) [38]. Within each age group, a separate SVM age regressor is trained to predict the final age. However, in contrast to the age group classification stage with SVM-BDT, we use all available training samples with ages up to 5 years outside of the age range for each group. In this way we mitigate error due to incorrect age group classification of face images with ages that are close to the group boundaries, and since we only use overlapped age ranges in the regression stage this does not introduce class label ambiguity during the classification stage. We use the RBF kernel for all SVM classifiers and regressors in this framework. For each data set, we select the parameters C and γ of the RBF kernel using a 5-fold cross-validation on the training set.

3.5. Age Estimate Fusion

We train separate hierarchical age estimators for the different feature types we extract from each face image. In practice we have found that an average of the ages predicted by several individual estimators can improve the overall performance, as long as the estimators have comparable performance. Further, we have found that since the best performing feature may depend on the age range, age estimation performance can be improved by fusing different features for different age groups.

4. Age Estimation by Humans

As a baseline for our automatic age estimation results, we gathered age estimates made by human workers using the Amazon Mechanical Turk service (AMT) crowdsourcing service⁶. We collected human age estimates for the en-

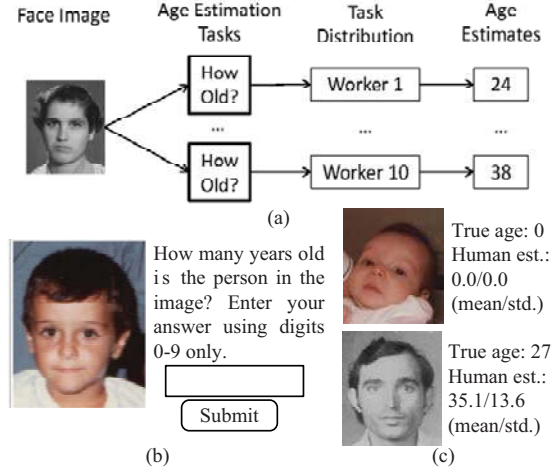


Figure 4. Age estimation using crowdsourcing experiment using Amazon Mechanical Turk. (a) Human age estimation process; (b) Overview of one HIT task. (c) Examples of consistent and inconsistent estimates by 10 workers.

tire FG-NET database, and a subset of the PCSO database consisting of 2,200 face images (with a similar age distribution to the larger subset used in our experiments). Mechanical Turk workers are anonymous, and we did not collect any personally identifying or demographic information from workers in these experiments. In order to keep the individual tasks assigned to workers (referred to as HITs or human intelligence tasks) as fine-grain as possible, a single task consisted of displaying a face image with the prompt string “How many years old is the person in the image? Enter your answer using digits 0-9 only.”, along with a text input box (See Fig. 4). For each face image, we asked ten workers to provide the age estimates. We posted a total of 32,020 such HITs on AMT. Out of the 32,020 HITs, 10,020 HITs were for the FG-NET database (1,002 face images \times 10 workers) and 22,000 HITs were for the PCSO database (2,200 face images \times 10 workers). So, for each face image 10 workers provided age estimates. The payment for each HIT was 3 cents, for a total cost of \$1,120.70 (including service fees).

As is typical of crowdsourcing experiments, some of the response data was noisy and unusable. For example, some workers submitted an empty text box, some entered a 3 digit long age, some entered an age range (e.g. 52-60), and at least one worker typed out the age in words e.g. “forty”. When workers entered an age range, we took the middle of the range as their estimate, and we manually converted the few word entries to integers. We rejected the remaining problematic cases (empty text boxes, age estimates larger than 100), and in these cases obtained replacement age estimates from additional workers. Out of the 10,020 initial FG-NET age estimates we found 16 of them to be unusable,

⁶<https://www.mturk.com/>

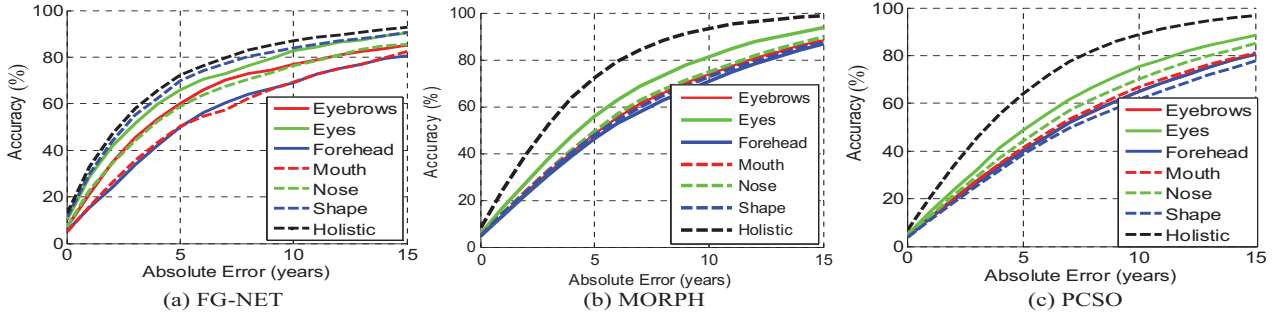


Figure 5. Per-component age estimation performance on (a) the FG-NET, (b) the MORPH Album2, and (c) the PCOS databases.

and out of the initial 22,000 PCSO age estimates, we found 604 of them to be unusable⁷.

Given 10 age estimates of each face provided by human workers, one needs to select a method for coming up with a single age estimate. Given the noisy nature of the crowd-sourced data, a simple strategy is to discard the highest and lowest age estimates for each face image, take the mean of the remaining 8 estimates and finally compute the MAE as we would for automatic age estimates. However, taking the mean of all age estimates for individual face images allows too-high and too-low age estimates to counteract each other [7, 39]. Hence, the MAE based on the mean age does not represent the performance of an individual human. We, therefore, discard the highest and lowest estimates for each image, then calculate the mean absolute error of all remaining individual age estimates.

5. Experimental Results

5.1. Aging of Individual Facial Components

The per-component age estimation performance of the proposed method on FG-NET, MORPH Album2, and PCSO are shown in Figs. 5 (a-c)⁸. On the FG-NET database, the holistic BIF features perform best overall, followed closely by BIF features extracted from the eye region, and the shape based features. On the other hand, performance on the MORPH Album2 and PCSO databases shows that the holistic BIF features significantly outperform all others, followed by the eyes, with the remaining feature types clustered relatively close together. A major consideration here is that the FG-NET database provides manually selected keypoint locations, while on the other two databases the keypoint locations were detected automatically. Since the shape features are based directly on the keypoint locations, their performance necessarily suffers when keypoint

localization is not very accurate. Our component localization method also relies on the keypoint locations, so the per component features do not perform as well as on FG-NET on automatically detected keypoints in the MORPH Album2 and PCSO databases.

5.2. Overall Performance

On the FG-NET database, we found that the best performance was attained by a fusion of the three best performing features, namely holistic BIF, shape, and eye region BIF. We attained a MAE of 4.6 by combining these three features for the lowest age group (0, 7), just combining holistic and shape based features for the 2nd age group (8, 17) and 3rd age group (18, 25), and just combining eye and holistic estimates for the oldest age group (26, 69).

On the MORPH Album2 and PCSO datasets, we found that the holistic BIF features outperformed the shape and per-component features to the extent that simple fusion methods showed no improvement in age estimation over the holistic BIF based estimate alone. For these two datasets, we therefore report only the performance of the holistic BIF based age estimator, which is as follows: an MAE of 4.2 on MORPH Album2, and an MAE of 5.1 on PCSO. While the MORPH Album2 and PCSO datasets both consist of mugshots of adults, we attribute this performance difference to the different age distributions of these two databases. The uniform age distribution of PCSO means that there are no cases where performance on one part of the age range can be sacrificed to improve performance on the overall dataset. Examples of face images where our algorithm gave good and poor age estimates are shown in Fig. 6.

Age estimation results from state-of-the-art methods on FG-NET and MORPH Album2 are included in Table 1. Better results than ours have been reported on FG-NET in [2], [3], and [28] with MAE of 4.5, 4.5, and 4.1 respectively. The best known performance on MORPH Album2 is 4.18 years MAE reported by Guo and Mu [12], which is slightly better than our MAE of 4.2. Compared to [12], we are using a larger version of MORPH Album2 (78,207 vs 55,000 images), but more significantly we handle ethnicity

⁷The human age estimates for the FG-NET database is available to interested researchers through our lab's website: <http://www.cse.msu.edu/biometrics/pubs/databases.html>

⁸There is no variation bars for the CS curves on the FG-NET and PCSO databases, because the CS curves are calculated on the whole dataset, not the average of individual folds.

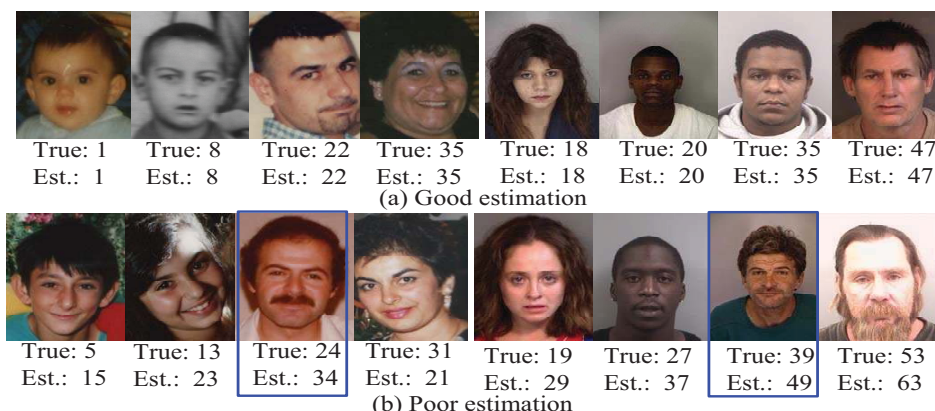


Figure 6. Examples of age estimation using the proposed approach. (a) Good age estimates; (b) Poor age estimates.

and gender variations differently. Guo and Mu constructed 10,000 image training sets with relatively balanced gender and ethnic groups, and trained a model that simultaneously predicts age, gender, and ethnicity. We randomly sample a 10,000 image training set which follows the race and ethnicity distribution of the complete dataset, and construct a model that does not use either the race or ethnicity ground truth during training. Although our method cannot predict gender or ethnicity, we do achieve comparable age estimation accuracy to [12] without directly modeling the effects of gender and ethnicity variation in the dataset.

5.3. Comparisons with Human Age Estimation

On FG-NET the MAE of human age estimates over the entire database is 4.7 with a variance of 24.8, while on the PCSO data the MAE of human age estimates is 7.2 with a variance of 32.0. On FG-NET (Fig. 7 (a)), the cumulative scores of human age estimation and the proposed automatic method are fairly close to each other. However, on the PCSO data ((Fig. 7 (b))) the automatic age estimation scores are consistently better than the human age estimation scores. To understand these results, we consider the per age range errors. On the FG-NET data, the human age estimates are significantly more accurate on the lower age ranges (0-15) than higher age ranges. This explains the difference in the overall performance on FG-NET and PCSO, since PCSO data does not include any face images of 0-15 year olds. In fact, the human MAE on FG-NET excluding 0-15 year olds is 7.4 with a variance of 32.8, which is fairly consistent with the human MAE on PCSO database.

6. Conclusions and Future Work

We have proposed a hierarchical approach for automatic age estimation, and analyzed the influence of aging on individual facial components using a component based representation. Human perception ability to estimate age is evaluated using crowdsourced data obtained via the Ama-

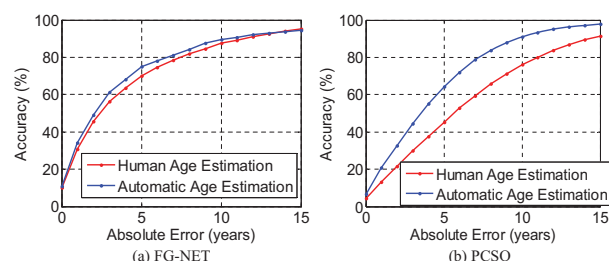


Figure 7. Human vs. automatic age estimation on (a) the FG-NET database, and (b) a subset of the PCSO database with 2,200 images.

zon Mechanical Turk service, and compared with the performance of the proposed automatic age estimation. Experimental results on the FG-NET, MORPH Album2, and PCSO databases show that eyes and nose are more informative in age estimation than the other facial components (forehead, eyebrows, mouth, and shape). We also show that the performance of the proposed age estimation method is better than or comparable to the age estimates provided by humans on FG-NET and a small subset of PCSO database.

Although the fusion of per-component age estimates provided no benefit on the MORPH Album2 and PCSO databases, we nevertheless report comparable performance to the best known result on MORPH Album2, and do so without taking advantage of the ground truth demographic information provided with the database. For future work, we plan to investigate methods to improve our automatic keypoint detection accuracy, additional strategies for age fusion, as well as the possibility of further improving our results by incorporating demographic information in our age estimation method.

References

- [1] Aging Skin Net. Causes of aging skin. <http://www.skincarephysicians.com/agingskinnet/basicfacts.html>.

- [2] K.-Y. Chang, C.-S. Chen, and Y.-P. Hung. Ordinal hyperplanes ranker with cost sensitivities for age estimation. In *Proc. IEEE CVPR*, pages 585–592, 2011.
- [3] W.-L. Chao, J.-Z. Liu, and J.-J. Ding. Facial age estimation based on label-sensitive learning and age-oriented regression. *Pattern Recogn.*, 46(3):628 – 641, 2013.
- [4] S. E. Choi, Y. J. Lee, S. J. Lee, K. R. Park, and J. Kim. Age estimation using a hierarchical classifier based on global and local facial features. *Pattern Recogn.*, 44(6):1262–1281, Jun. 2011.
- [5] Cognitec Systems GmbH. Facevac software developer kit. <http://www.cognitec-systems.de>, 2010.
- [6] T. Cootes, C. Taylor, D. Cooper, and J. Graham. Active Shape Models - Their Training and Application. *Comp. Vis. Img. Und.*, 61(1):38 – 59, 1995.
- [7] N. Ebner, M. Riediger, and U. Lindenberger. FACES - a database of facial expressions in young, middle-aged, and older women and men: Development and validation. *Behavior Research Methods*, 42(1):351–362, 2010.
- [8] Y. Fu, G. Guo, and T. Huang. Age synthesis and estimation via faces: A survey. *IEEE Trans. PAMI*, 32(11):1955–1976, Nov. 2010.
- [9] Y. Fu and T. Huang. Human age estimation with regression on discriminative aging manifold. *IEEE Trans. Multimedia*, 10(4):578–584, Jun. 2008.
- [10] D. Gabor. Theory of communication. *J. of the Institution of Electrical Engineers*, 93(26):429–441, Nov. 1946.
- [11] X. Geng, Z.-H. Zhou, and K. Smith-Miles. Automatic age estimation based on facial aging patterns. *IEEE Trans. PAMI*, 29(12):2234–2240, Dec. 2007.
- [12] G. Guo and G. Mu. Simultaneous dimensionality reduction and human age estimation via kernel partial least squares regression. In *Proc. IEEE CVPR*, pages 657–664, 2011.
- [13] G. Guo, G. Mu, Y. Fu, and T. S. Huang. Human age estimation using bio-inspired features. In *Proc. IEEE CVPR*, pages 112–119, 2009.
- [14] G. Guo and X. Wang. A study on human age estimation under facial expression changes. In *Proc. IEEE CVPR*, pages 2547–2553, 2012.
- [15] H. Han, B. Klare, K. Bonnen, and A. K. Jain. Matching composite sketches to face photos: A component-based approach. *IEEE Trans. IFS*, 8(1):191–204, 2013.
- [16] H. Han, S. Shan, X. Chen, and W. Gao. A comparative study on illumination preprocessing in face recognition. *Pattern Recognition*, 46(6):1691–1699, 2013.
- [17] H. Han, S. Shan, X. Chen, S. Lao, and W. Gao. Separability oriented preprocessing for illumination-insensitive face recognition. In *Proc. ECCV*, pages 307–320, 2012.
- [18] J. Hayashi, M. Yasumoto, H. Ito, and H. Koshimizu. Age and gender estimation based on wrinkle texture and color of facial images. In *Proc. ICPR*, pages 405–408, 2002.
- [19] B. Heisele, P. Ho, J. Wu, and T. Poggio. Face recognition: component-based versus global approaches. *Comp. Vis. Img. Und.*, 91(1-2):6–21, Jul. 2003.
- [20] J. Huang, V. Blanz, and B. Heisele. Face recognition using component-based svm classification and morphable models. In *Proc. Int'l Workshop Pattern Recogn. with SVM*, pages 334–341, 2002.
- [21] D. H. Hubel and T. N. Wiesel. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiology*, 160(1):106–154, Jan. 1962.
- [22] R. Iga, K. Izumi, H. Hayashi, G. Fukano, and T. Ohtani. A gender and age estimation system from face images. In *Proc. SICE Annual Conference*, pages 756–761, 2003.
- [23] Y. Ivanov, B. Heisele, and T. Serre. Using component features for face recognition. In *Proc. FGR*, pages 421 – 426, May 2004.
- [24] J. P. Jones and L. A. Palmer. An evaluation of the two-dimensional gabor filter model of simple receptive fields in cat striate cortex. *J. Neurophysiology*, 58(6):1233–1258, Nov. 1987.
- [25] N. Kumar, A. Berg, P. Belhumeur, and S. Nayar. Describable visual attributes for face verification and image search. *IEEE Trans. PAMI*, 33(10):1962–1977, 2011.
- [26] A. Lanitis, C. Taylor, and T. Cootes. Toward automatic simulation of aging effects on face images. *IEEE Trans. PAMI*, 24(4):442–455, Apr. 2002.
- [27] Z. Li, Y. Fu, and T. Huang. A robust framework for multi-view age estimation. In *Proc. IEEE CVPR Workshops*, pages 9–16, 2010.
- [28] K. Luu, K. Seshadri, M. Savvides, T. Bui, and C. Suen. Contourlet appearance model for facial age estimation. In *Proc. IJCB*, pages 1–8, 2011.
- [29] S. Milborrow and F. Nicolls. Locating facial features with an extended active shape model. In *Proc. ECCV*, pages 504–513, 2008.
- [30] C. Otto, H. Han, and A. Jain. How does aging affect facial components? In *Proc. ECCV*, pages 189–198, 2012.
- [31] M. Rhodes. Age estimation of faces: A review. *Applied Cognitive Psychology*, 23(1):38 – 59, 1995.
- [32] K. Ricanek and T. Tesafaye. MORPH: a longitudinal image database of normal adult age-progression. In *Proc. FGR*, pages 341–345, 2006.
- [33] M. Riesenhuber and T. Poggio. Hierarchical models of object recognition in cortex. *Nature Neuroscience*, 2(11):1019–1025, Nov. 1999.
- [34] S. Z. Li and A. K. Jain (eds.). Handbook of face recognition, 2nd edition. *Springer-Verlag, London*, 2011.
- [35] T. Serre, L. Wolf, and T. Poggio. Object recognition with features inspired by visual cortex. In *Proc. IEEE CVPR*, pages 994 – 1000, 2005.
- [36] J. Suo, S.-C. Zhu, S. Shan, and X. Chen. A compositional and dynamic model for face aging. *IEEE Trans. PAMI*, 32(3):385–401, Mar. 2010.
- [37] P. Thukral, K. Mitra, and R. Chellappa. A hierarchical approach for human age estimation. In *Proc. IEEE ICASSP*, pages 1529–1532, 2012.
- [38] V. N. Vapnik. *Statistical Learning Theory*. John Wiley, New York, 1998.
- [39] M. C. Voelkle, U. L. N. C. Ebner, 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.
- [40] T. Wu, P. Turaga, and R. Chellappa. Age estimation and face verification across aging using landmarks. *IEEE Trans. IFS*, 7(6):1780–1788, Dec. 2012.