

Automatic Age Estimation: A Survey

Nabila Mansouri^{1,2}

¹ Ha'il University,
Saudi Arabia

² Sfax University,
ReDCAD Laboratory,
Tunisia

nabila.elmansouri@gmail.com

Abstract. Aging is non-reversible process. Human face and gait change with time which reflects major variations in appearance, the vast majority of people are able to easily recognize human traits like emotional states, where they can tell if the person is happy, sad or angry from the face, likewise, it is easy to determine the gender of the person. However, knowing person's age is a very challenge task. Hence, significant interest in the computer vision and pattern recognition research community is given to automatic age estimation. This paper, presents a thorough analysis of recent research in aging and age estimation. We discuss popular algorithms used in age estimation and existing models. Underline age estimation challenges especially using RGB images. Finally insights for future research based on depth map and Kinect camera.

Keywords. Age estimation, face aging, gait, survey, RGB-D, depth map.

1 Introduction

In the past few years manifold approaches have been proposed for studying age group recognition problem based on RGB images. There are two categories of approaches: Cranio-Facial approaches [1, 2, 3, 4, 5, 6] and Behavioral Approaches [7, 8]. Although, behavioral approaches are very promising research axes that give good performances when person is far from the camera, face remains the most affected body region by aging effect. Face appearance is widely described by 2D descriptors that aim to characterize the shape, texture or both of them.

Face-based age estimation problems have been mainly extensively studied using conventional RGB cameras at visible light. However, this makes some aging features extraction a challenging problem. Furthermore, face images acquired using such conventional sensors may have inherent restrictions that hinder the inference of some specific aging information in the face such as wrinkles' depth.

Microsoft Kinect is introduced in 2010. It is widely adopted by the computer vision research community in various applications [2] of face analysis such as face [9, 10, 11, 12, 13], gender [14, 15] and ethnicity [16] recognition.

It appears that most of the few attempts on using Kinect in face analysis are mainly devoted to the face recognition problem, gender recognition and ethnicity [17], hence overlooking and ignoring other face analysis tasks such as age estimation. Moreover, most of the proposed works focused on the fusion of Kinect depth information and RGB images but did not explicitly explore how much information Kinect facial depth data alone can reveal about the faces [17]. Some of the results are also reported on size-limited and/or private Kinect databases. Also problem of age labeled subjects in the available Kinect databases.

There has been enormous effort from both academia and industry dedicated towards modeling age estimation, designing of algorithms, aging face dataset collection, and protocols for evaluating system performance.

This paper summarizes the findings of recent studies in age estimation, evaluation protocol used, and feature extraction from both gait and face in section 2. Age recognition applications are presented in section 3. Section 4 presents age estimation purposes (regression, classification). Also different used dataset are presented. Finally a short discussion and proposition are presented.

2 Related Work

2.1 Approaches based on RGB Images

2.1.1 Facial Aging

Aging causes significant change in facial shape in formative years and relatively important texture variations with still minor change in shape in older age groups [20, 21, 22, 23, 24, 25, 26, 27].

In fact cranio-facial growth introduces Shape variations in younger age groups. Cranio-facial studies have shown that human faces change from circular to oval as one ages [28]. These changes lead to variations in the position of fiducial landmarks [29]. During cranio-facial development, the forehead slopes back releasing space on the cranium. The eyes, ears, mouth, and nose expand to cover interstitial space created. The chin becomes protrusive as cheeks extend. Facial skin remains moderately unchanged than shape. More literature on cranio-facial development is found in [29].

As one ages, facial blemishes like wrinkles, freckles, and age spots appear. Underneath the skin, melanin producing cells are damaged due to exposure to the sun's ultraviolet (UV) rays. Freckles and age spots appear due to overproduction of melanin. Consequently, light reflecting collagen not only decreases but also becomes non-uniformly distributed making facial skin tone nonuniform [30]. Parts adversely affected by sunlight are the upper cheek, nose, nose bridge, and forehead.

However, the most visible variations in adulthood to old age are skin variations exhibited in texture change. There is still minimal facial shape variation in these age groups.

Biologically, as the skin grows old, collagen underneath the skin is lost [20]. Loss of collagen and effect of gravity make the skin become darker, thinner, leathery, and less elastic. Facial spots and wrinkles appear gradually. The framework of bones beneath the skin may also start deteriorating leading to accelerated development of wrinkles and variations in skin texture.

2.1.2 Geometric Models

Geometric modeling of facial aging focuses on distance measurements between facial points. Face antropometry is the study of measuring sizes and proportions on human faces. As instance, Fu and Huang [35] developed a manifold embedding approach to the age estimation problem, whose purpose is to find a low-dimensional representation in the embedded subspace and capture geometric structure and data distribution. [36] explored a Grassmann manifold to model the facial shapes and considered the age estimation as regression and classification problems on this representation [37] for this work, both image and geometric features were taken into account for estimating ages from images. Concerning image features, two descriptors were extracted from the images: the first one is the Histogram of Oriented Gradients (HOG) (Dalal, Triggs, 2005), a very popular and robust descriptor used for detection and recognition of objects and faces (Deniz et al., 2011; Felzenszwalb et al., 2010; Suard et al., 2006; Zhu et al., 2006).

2.1.3 Active Shape Models

Active shape model (ASM) [34] is a statistical model that characterizes shape of an object. ASM builds a model by learning patterns of variability from a training set of correctly annotated images.

As instance, the Active Shape Models (ASM) were used for accurately localizing the facial region to extract features only from the eyes, nose, mouth and static wrinkles regions of the input image [31, 32, 33]. Only 68 points were used to cover the eyes, nose, mouth and static wrinkles regions. Finally the input image is cropped to just the area covered by the Active Shape Models fitted landmark points.

ASM is more similar to AAM but differs in the sense that instances in ASM can only deform according to variations found in the training set. ASM is not commonly used in age estimation; hence, more investigations adopting this modeling strategy are necessary [38].

2.1.4 Active Appearance Models

Active Appearance Models (AAM) [16, 17] based approaches consider both shape and texture rather than just the facial geometry as in the anthropometric model based methods. AAM uses a statistical model of object shape and appearance to synthesize a new image throughout a training stage which provides to the training supervisor a set of images and coordinates of landmarks existing in all of the images. AAMs represent a familiar group of algorithms for fitting shape models to images. Training a model requires labeling a database of images where a set of locations called landmarks typify the object group in question.

The formulation in [17] chooses a linear and generative model, i.e. an explicit model of the input data has to be provided. This leads to an iterative Gauss-Newton type procedure, where the error between the current image features and those synthesized using the current location of the model in the image are used to derive additive updates to the shape model parameters. Nonetheless, the computational load is heavy, since an explicit image feature model must be stated and evaluated at each algorithm iteration [16].

Lanitis et al. [18] extended AAMs for aging faces by proposing an aging function, $age=f(b)$ which explains the variation in age. But they have to deal with each aging face image separately. Kohli et al. [19] extracted feature vectors from images using AAMs and used ensemble of classifiers trained on different dissimilarities to distinguish between child/teen-hood and adulthood. By using the different aging functions, accurate age of the classified image is estimated. Chao et al. [20] proposed an age estimation method using AAM features. Their approach is based on label sensitive learning and age-oriented regression.

2.1.5 Appearance Models

Appearance models mainly model facial appearance using texture, shape, and wrinkle features for age estimation, face recognition, face verification, and gender estimation among other tasks. Image is represented by vectoring both shape and texture [73]. Appearance models are more like AAM [64] that builds a statistical model using the shape and texture of the face. Both global and local texture, shape and wrinkle features are extracted and modeled for age estimation. Texture and shape have been used for age and gender estimation [74,75].

Age estimation using appearance features can be improved by performing gender estimation prior since males and females exhibit varied aging patterns. Given a set of facial images $X = \{x_i : x_i \in \mathbb{R}^D\}_{i=1}^n$ and a vector of age labels $L = \{l_i : l_i \in \mathbb{N}^D\}_{i=1}^n$, facial features are extracted from vector x_i of images at a particular age. Every feature F_i has a one-to-one mapping with one of the age label l_i . After features are extracted and associated with age label, they are used for age estimation either using a regression model or classification.

Effectiveness of LBP [76] in texture characterization has made it popular in extraction of appearance features for age estimation. LBP has been used in [77] and achieved 80% accuracy in age estimation with nearest neighbor classifier and 80–90% accuracy with AdaBoost classifier [78].

Gao and Ai [79] used Gabor filter [67] appearance feature extraction technique for age estimation and reported better results compared to LBP technique. BIF [80, 81] is also used in appearance-based models as used in [82].

Using age manifold, BIF and SVM classifier, MAE of 2.61 and 2.58 years for females and males, respectively, can be achieved on YGA database [11]. This shows BIFs' superior performance in age estimation. Spatially flexible patch (SFP) proposed in [83, 84] is another feature descriptor that can be used for characterizing appearance for age estimation.

Other techniques that can be used to build appearance models for age estimation are linear discriminant analysis (LDA) and principal component analysis (PCA).

2.1.6 Hybrid Models

What is the best modeling approach for age estimation? It is hard to certainly answer this question since each of the modeling approaches discussed have their inherent strengths and limitations. To get the answer to the question, one may try different modeling approaches on the representative images and compare their performance.

By comparing different modeling approaches, strengths and limitations of each of the models can be found. Modeling approaches that are complementary of each other can be combined to form a hybrid modeling approach. Hybrid age estimation modeling combines several modeling techniques to take advantage of the strengths of each technique used. By combining different modeling techniques, age estimation accuracies are expected to not only improve but also be robust. These models could be combined in a hierarchical manner or parallel and results from different models combined for final age estimation.

2.2 Gait Aging

Since face-based human age estimation approaches require people's collaboration in order to extract face features, they seem to be inapplicable when people are far from the camera even in the one random situation without the cooperation of the subject. Gait-based approaches can solve this problem and ensure a reliable features extraction at a great distance from the camera, as well as to deal with point of view variation and image resolution when the face is not available. As a biometric information, gait is the most common human activity which represents an individual's way of walking and/or posture.

A number of gait-based techniques have existed for quite some times [8, 9, 10, 11], but most of them are used for human identification [8, 9, 12, 13] and gender recognition [14, 15, 16, 17, 18, 11]. Gait-based approaches are usually free-model [8,19], that aim to extract gait features directly from the silhouette without using any model to represent gait structure or motion, or Model-based [20, 10, 9], that deploy some structural or motion models

to model either the entire human body or some specific body parts.

Furthermore, when dealing with human gaits, the term of pose manifold was often used to represent the sequential and cyclic pattern of the human gaits. Several publications have appeared in recent years documenting this issue [20, 21, 22, 23]. These works represent the variability of different walking styles from multiple individuals, where dual gait generative models were proposed, one for visual data and one for kinematic data.

Nevertheless, medical and psychological studies have shown that gait feature also contains age discriminative information. In fact, according to several medical researches [24, 25], human gait performance starts evolving from early age to reach maturity at the age of seven.

However, this performance decreases significantly after the age of 60. In addition, gait velocity, stride size and walking posture are significantly different for young and elder persons [26, 27]. It was reported also, that during even surface walking there were differences in kinetic and kinematic gait patterns between the young and the elderly [29, 30, 31].

A quick glance at the literature shows that few approaches have addressed age estimation problem based on gait features [28]. Makihara et al. [32] proposed a gait analysis of gender and age. Such as study serves as an introduction to describes gait differences between age classes and genders, namely Young, Adults, Elderly, Males and Females. They represented the silhouette in frequency domain to extract gait features and provided various classification experiences.

This study presents an important analysis of the uniqueness of gait for each class to acquire insight into gait differences among genders and age classes. Rely on this work an important changes in gait parameters have been reported because of the aging. Compared to young, elder have low gait velocity, step length, stride length and single leg support time. Moreover, age estimation based on gait features approaches can be divided into two categories: (i) Contour-based approaches and (ii) Silhouette-based ones. As an instance of contour based approaches, we cite Zhang et al. [27] work.

Authors addressed age classification problem using a contour-based descriptor to extract human gait features, namely Feature to Exemplar Distance (FED), and Hidden Markov Model (HMM). FED descriptor consists of measuring the distance between the silhouette's centroid and contour points situated at centroid-derived segment, with an angle augmenting by 6° for each point, and contour overlap. Classification performance reached high rates, however, only a small data-set of 14-persons was used to test the descriptor.

In addition, HMM requires an off-line learning phase to construct models which need extra computing-time when compared to other classification methods (i.e. SVM). Moreover, some basic geometric parameters such as stride frequency and head-to-body ratio are introduced in [28] in order to discriminate between adult and child. The contour of the pedestrian image is firstly drawn. Secondly body skeleton points are localized and the geometric parameters (stride frequency and head-to-body ratio) are computed relying on head and ankles points. This study is evaluated on a small portion of OU-ISIR dataset using a linear separator.

The good 2 reached result may due to the small number of subjects in the experiment. It is so, conjectured that the performance might drop thorough test using the complete OU-ISIR gait database with the same approach but much larger number of subjects [28].

However as a silhouette-based approach, the most interesting one [33] has provides a base-line algorithm for gait-based age estimation using Gaussian process regression. In their work they used Gait Energy Image (GEI), a FREQUENCY domain silhouette representation (FREQ) and Gait Period (GP) as gait descriptors. Tests were conducted using the most huge gait dataset, OU-ISIR [34] that includes a great variability of the subjects' ages. The three proposed descriptors (GEI, FREQ, GP) are compared and the best performances are achieved by the GEI.

Even though, GEI was widely used in the literature for gait analysis and age estimation purposes, it has several shortcoming such as manipulating the entire silhouette data without focusing on the most discriminating interest regions.

Already presented limitation makes GEI a heavy descriptor which encapsulates very high detailed (useless in some cases) data which decrease the quality of discriminating features.

In addition, the work presented in [35] proves that soft biometric characteristics, such as age can also be derived from gait patterns. Authors use an enhancement Gabor filter and maximization of mutual information to extract low-dimensional features. Gabor wavelets are applied for feature extraction, which decompose body shape into local orientations and scales. The experimental study conducted with a supervised learning phase and HMM classification shows very encouraging results even if are performed on a small dataset, seven people for each age group.

Although, the outstanding aging characteristics that appear on the elderly gait, it is not well used as the age classifier [27, 32] as is the case for gender recognition and people identification. Hence, this paper introduces a new type of gait-based human age descriptor relying on medical and bio-mechanical researches [36, 37, 38] that confirm that arms swing, head pitch, hunched posture and stride length are among the most outstanding aging characteristics. Proposed descriptor represents spatio/temporal gait variations and take advantage of these features for better characterization of each age class.

2.3 Approaches based on RGB-D Images

Microsoft Kinect is introduced in 2010. It is widely adopted by the computer vision research community in various applications [2] of face analysis such as face [19, 20, 21, 22, 23, 7], gender [24, 25] and ethnicity [26] recognition. Recently, a pioneer work explores the RGB-D image efficiency in age estimation [27].

It appears that most of the few attempts on using Kinect in face analysis are mainly devoted to the face recognition problem, gender recognition and ethnicity [1], hence overlooking and ignoring other face analysis tasks such as age estimation. Moreover, most of the proposed works focused on the fusion of Kinect depth information and RGB images but did not explicitly explore how much information Kinect facial depth data alone can

reveal about the faces [1]. Some of the results are also reported on size-limited and/or private Kinect databases.

3 Age Estimation Application

3.1 Video Surveillance

Automatic age estimation provides an important tools to tack care of elders leaving alone. An active research area is interested of the smart city and home.

3.2 Electronic Customer Relationship Management (ECRM)

The goal of electronic customer relationship management (eCRM) systems is to improve customer service, retain valuable customers, and to aid in providing analytical capabilities. Furthermore, it is the infrastructure that enables the delineation of and increases in customer value, and the correct means by which to motivate valuable customers to remain loyal [39]. Hence, automatic age estimation canwedly improves this topic.

3.3 Biometrics

Age estimation via faces is a soft biometric [32] that can be used to compliment biometric techniques like face recognition, fingerprints, or iris in order to improve recognition, verification, or authentication accuracies. Age estimation can be applied in age-invariant face recognition [10], iris recognition, hand geometry recognition, and fingerprint recognition in order to improve accuracy of hard (primary) biometric system [11].

3.4 Human and Machine Interaction: HMI

If computers could determine the age of the user both the computing environment and the type of interaction could be adjusted according to the age of the user. Apart from standard HCI, such a system could be used in combination with secure internet access control in order to ensure that under-aged persons are not granted access to internet pages with unsuitable material.

3.5 Researchers in Dermal and Cosmetic

The approach proposed in [40] presents a new model called Human Injected by Botox Age Estimation (HIBAE) model, a human age estimator based on active shape models, speed up robust feature, and support vector machine to accurately estimate the age of people that are exposed to Botox injections. Human Injected by Botox Age Estimation proposed model was trained by a crossover of Productive Aging Lab.

4 Age Estimation Algorithms

In recent years many methods have been applied to automatic recognize the person's age . We can distinguish two approaches within age recognition: age regression and age classification.

4.1 Age Regression

For age estimation different types of regression methods have been used. The most recent are based on the 50 raw model parameters, [41] investigated linear, quadratic, and cubic formulation of aging function. Genetic algorithm is used to learn optimal model parameters from training face images of different ages. Quadratic and cubic aging function achieved better MAE 0.86 and 0.75, respectively, compared to 1.39 of linear function. This suggests that quadratic function offers the best alternative since its MAE was not significantly different from that of cubic function and it is not computationally intensive as cubic function.

Approaches proposed in [42, 44] use linear support vector regression (SVR) on age manifold for age estimation. They reported MAE of 7.47 and 7.00 years for males and females, respectively, on YGA dataset and MAE of 5.16 on FG-NET dataset. Yan et al. [45] formulated a regression problem for age estimation using semi-definite programming (SDP). The regressor was learned from uncertain non-negative labels.

They reported MAE of 10.36 and 9.79 years for males and females, respectively, on YGA. They further demonstrated that age estimation by SDP formulation achieves better results compared to ANN.

The limitation of SDP is that it is computationally expensive especially when the training set is large. A regression model for age estimation is used in [46]. The face image was represented by a multi-level local binary pattern (MLBP). This study achieves a MAE of 6.6.

In addition, the approach presented in [43] achieves a MAE of 4.0 by using BIF to model a regression model for age estimation. Using manifold of raw pixel intensities to represent face image, a regression model evaluated in [47] on MORPH II dataset obtains a MAE of 5.2 for White ethnic group and 4.2 for Black ethnic group. The approach in [48] applies a boosted regressor on age rank local binary patterns (arLBP). They reported a MAE of 2.34 on FG-NET using LOPO validation protocol. Their approach demonstrated that age ranking with correlation of aging patterns across age groups improves performance of age estimation.

4.2 Age Classification

The age classification problem was previously examined in [49] that explored the performance of nearest neighbor, artificial neural network (ANN), and quadratic function in age estimation tasks. Although the quadratic function used to relate face representations to face labels is a regression function, the authors referred to it as a quadratic function classifier [49]. The quadratic function reported MAE of 5.04, which was superior to MAEs reported by nearest neighbor. ANN and self-organizing maps (SOMs) reported better performance compared to quadratic function. The authors proposed clustering and hierarchical age estimation for improving performance.

Furthermore a Comparison between humans and computers in age estimation was also done and found that computers can estimate age almost as reliable as humans. Ueki et al. [50] built 11 Gaussian models in low dimensional 2DLDA and LDA feature space using expectation maximization (EM). Age-group estimation was determined by fitting probe image to each cluster and comparing the probabilities.

They reported a higher accuracy, 82% male and 74% female, with wide age groups of 15 years



Fig. 1. Typical aging face sequence in FG-NET Aging Database



Fig. 2. Typical aging face sequence in FG-NET Aging Database

as compared to 50% male and 43% female in age groups of a 5-year range. This demonstrates that this approach can only post better accuracies where age groups have wide ranges and hence not applicable in a narrow-range age-group estimation. Fusing texture and local appearance, Huerta et al. [51] used a deep learning classification for age estimation.

Using speeded-up robust features (SURF) [52], and histogram of oriented gradients (HOG) [?], he evaluated the performance of deep learning on two large datasets and achieved MAE of 3.31. Hu et al. [54] used Kullback-Leibler/raw intensities for face representation before using convolutional neural network (CNN) for age estimation. Their approach achieved MAE of 2.8 on FGNET and 2.78 on MORPH II. This demonstrates that deep learning (deep neural networks or CNN) achieves better MAE compared to traditional classification methods.

5 Age Estimation Databases

5.1 RGB Face Databases

5.1.1 FG-NET-DB

The FG-NET Aging Database [55] contains 1,002 face images from 82 subjects with approximately 10 images per subject. The ages in the database are distributed in a wide range from 0 to 69. The age distribution of the FGNET database is given in Table 1. One can see from the table that the images are not distributed uniformly.



Fig. 3. Typical aging face sequence in FG-NET Aging Database

Table 1. Age range distribution of the images in FG-NET databas

Age Range	FG-NET (%)
0-9	37.03
10-19	33.83
20-29	14.37
30-39	7.88
40-49	4.59
50-59	1.50
60-69	0.80

Table 2. Age age distribution of the images in MORPH database

Age Range	MORPH (%)
18-29	951
30-39	445
40-49	126
50+	32

A typical aging sequence from the FG-NET database is shown in Figure 2. Besides the aging variation, most aging sequences display variations in pose, illumination, facial expression, occlusion, etc.

Although these variations may increase computational complexity.

5.1.2 MORPH-DB

MORPH [56] is a publicly available aging database created by the Face Aging Group at the University of North Carolina.

This dataset is composed of two sets. Age distribution in this dataset range from 18 to 50+ years (See Table 1). There are 1430 images for males and 294 images for females with age gap ranging from 46 days to 29 years. Set 2 contains 55,134 images of 13,000 individuals collected over 4 years. Both albums contain metadata for race, gender, date of birth, and date of acquisition. The eye coordinates of the dataset can be requested. A commercial version of album 2 contains a larger set of images collected over a longer time span and includes information like the height and weight of individual. Database images samples are shown in figure 3.

5.1.3 Iranian Face Database (IFDB)

The Iranian Face Database (IFDB) [38], the first image database in middle-east, contains color facial imagery of a large number of Iranian subjects. IFDB is a large database that can support studies of the age classification systems. It contains over 3,600 color images. IFDB can be used for age classification, facial feature extraction, aging, facial ratio extraction, percent of facial similarity, facial surgery, race detection and other similar researches.

5.1.4 Specs on Faces (SoF) Dataset

The SoF dataset [57] is a collection of 42,592 ($2,662 \times 16$) images for 112 persons (66 males and 46 females) who wear glasses under different illumination conditions. The dataset is FREE for reasonable academic fair use. The dataset presents a new challenge regarding face detection and recognition. It is devoted to two problems that affect face detection, recognition, and classification, which are harsh illumination environments and face occlusions.



Fig. 4. EURECOM database samples



Fig. 5. Superface-Kinect-Face-dataset samples



Fig. 6. OU-ISIR database samples

The glasses are the common natural occlusion in all images of the dataset. However, the glasses are not the sole facial occlusion in the dataset; there are two synthetic occlusions (nose and mouth) added to each image. Moreover, three image filters, that may evade face detectors and facial recognition systems, were applied to each image. All generated images are categorized into three levels of difficulty (easy, medium, and hard).

The glasses are the common natural occlusion in all images of the dataset. However, the glasses are not the sole facial occlusion in the dataset; there are two synthetic occlusions (nose and mouth) added to each image. Moreover, three image filters, that may evade face detectors and facial recognition systems, were applied to each image. All generated images are categorized into three levels of difficulty (easy, medium, and hard).

That enlarges the number of images to be 42,592 images (26,112 male images and 16,480 female images). Furthermore, the dataset comes with a metadata that describes each subject from different aspects. The original images (without filters or synthetic occlusions) were captured in different countries over a long period.

Usage: 1 - Gender classification; 2 - Face detection; 3 - Facial landmark estimation; 4 - Emotion Recognition; 5 - Eyeglasses detection; 6 - Age classification.

5.2 RGB-D Face Databases

Kinect is a new hardware device that is recently used in computer vision applications. Indeed, there exist a few RGB-D face databases with subject age annotation which are publicly available. Two existing databases filmed with Kinect such as.

5.2.1 EURECOM Database

The EurecomKinect face database [58] contains both RGB and depth facial images of 52 subjects acquired using Kinect sensor. The people in the database belong to 2 different age-groups (young and adult). The data is captured in two sessions separated by two weeks. In each session, the facial images of each person are captured under 9 different facial variations (neutral, smile, open mouth, strong light, eyes occlusion, mouth occlusion, paper occlusion, left profile and right profile). Face image samples from this database are shown in Figure 4 a).

5.2.2 Superface-Kinect-Face-dataset

The Superface-Kinect-Face-dataset [59] (Figure 5) contained simultaneously a various sequence of different positions of the 2D and 3D facial image captured by the Kinect camera. It contained 920 images pertaining to 20 subjects for the young and the adult age.

5.3 Gait Database

5.3.1 OU-ISIR database

OU-ISIR is a large population dataset [60] which contains gait sequences of 4007 persons (2135 men and 1872 women). As we are concern about the age factor, this database have ages ranging from 1 to 94 years old. Figure 6. b shows the distributions of the subjects' gender and age information. There are several advantages for using this dataset.

First, it is almost twenty times the size of the largest publicly available gait database [7], providing enough subjects for almost any tests or experiments. Besides, it has a good gender balance with the ratio of males to females close to one.

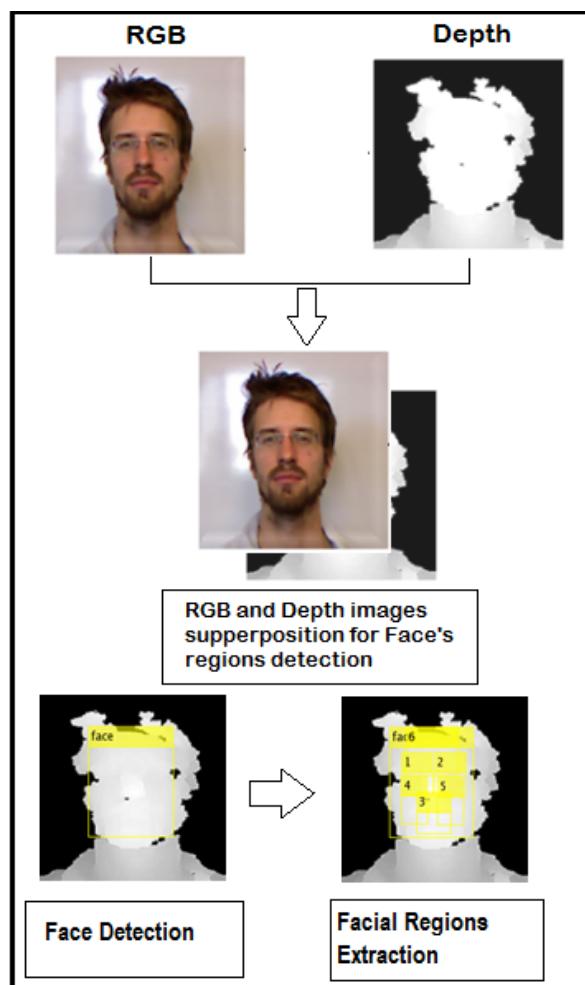


Fig. 7. Proposed process for face detection and IR extraction from RGB-D images

In addition, the wide range of ages from 1 to 94 years old provides an almost ideal dataset for testing.

Furthermore, the dataset is composed of silhouette sequences pretreated with size normalized to 88 by 128 pixels. Peoples are filmed in front-parallel point of view with a great distance to the camera when a face is not available. With these characteristics, the dataset servers as an ideal performance evaluation in our case to verify the reliability of the proposed gait-based age estimation method.

6 Discussion and Proposition

Face-based age estimation problems have been mainly extensively studied using conventional RGB cameras at visible light. 2D face images acquired using the conventional sensors (traditional RGB cameras) may have inherent restrictions that hinder the inference of some specific information in the face. Thus, the classical 2D HOG descriptor describes properly the face's appearance and detect the first primary appearance of aging effect but it can not deal with their accentuation and digging evolution.

Also, it appears that most of the few attempts on using Kinect in face analysis are mainly devoted to the face recognition problem, gender recognition and ethnicity [17], hence overlooking and ignoring other face analysis tasks such as age estimation. Moreover, most of the proposed works focused on the fusion of Kinect depth information and RGB images but did not explicitly explore how much information Kinect facial depth data alone can reveal about the faces' aging [17].

To overcome these limitations, using RGB-D images to describe the local distribution of aging effects, direction and depth evolution from the depth map is the suitable solution. Indeed, the new approach of the 3D descriptors applies the classical 2D features extraction process on the depth maps. These depth maps are extracted from face's Int erst Regions (IR), most affected by aging.

The main input will be a RGB and RGB-D images superposition for face detection and IR extraction as illustrated by image (7).

7 Conclusion and Future Work

This paper presents a survey of various techniques and approaches used for age estimation. These approaches are classified into RGB-images based approaches and RGB-D images based approaches. Also face aging and gait aging features are presented. Hence we can underline based on this study, although the important results achieved based on face analysis, both gait and RGB-D image are very promising research field in age recognition.

References

1. Sung, E.Ch., Youn, J.L., Sung, J.L., & Jaihie, K. (2014). Hierarchical age estimation from unconstrained facial images. *Pattern Recognition*, pp. 1262–1281.
2. Song, Z., Bingbing, Ni., Dong, G., Terence, S., & Shuicheng, Y. (2010). Learning universal multi-view age estimator by video context. *Proceeding International Conference on Computer Vision*, pp. 1–8.
3. Yanchao, S., Haizhou, A., & Shihong, L. (2008). Real-time face alignment with tracking in video. *Proceeding International Conference on Image Processing*, pp. 1632–1635.
4. Hajizadeh, M.A. & Ebrahimnezhad, H. (2011). Classification of age groups from facial image using histograms of oriented gradients. *Proceeding Machine Vision and Image Processing*, pp. 1–5.
5. Jing-Ming, G., Yu-Min, L., & Hoang-Son, N. (2011). Human face age estimation with adaptive hybrid features. *Proceeding International conference on system science and engineering*, pp. 55–58.
6. Selvi, V.T. & Vani, K. (2011). An efficient age estimation system based on multi-linear principal component analysis. *Journal of Computer Science*, pp. 1497–1504.
7. Mansouri, N., Aouled-Issa, M., & Ben-Jemaa, Y. (2018). Gait-based human age classification using a silhouette model. *IET Biometrics*, Vol. 7, No. 2, pp. 116–124.
8. Mansouri, N., Aouled-Issa, M., & Ben-Jemaa, Y. (2018). Gait features fusion for efficient automatic age classification. *IET Computer Vision*, Vol. 12, No. 1, pp. 69–75.
9. Li, B., Mian, A., Liu, W., & Krishna, A. (2013). Using Kinect for face recognition under varying poses, expressions, illumination and disguise. *IEEE Workshop on Applications of Computer Vision*, pp. 186–192.
10. Goswami, G., Bharadwaj, S., Vatsa, M., & Singh, R. (2013). On RGB-D face recognition using Kinect. *International Conference on Biometrics: Theory, Applications and Systems*, pp. 1–6.
11. Min, R., Choi, J., Medioni, G., & Dugelay, J. (2012). Real-time 3D face identification from a depth camera. *International Conference on Pattern Recognition*, pp. 1739–1742.
12. Pamplona-Segundo, M., Sarkar, S., Goldgof, D., Silva, L., & Bellon, O. (2013). Continuous 3D face authentication using rgb-d cameras. *IEEE Conf. on Computer Vision and Pattern Recognition Workshops*, pp. 64–69.
13. Li, B., Mian, A., Liu, W., & Krishna, A. (2015). Face recognition based on Kinect. *Pattern Analysis and Applications*, pp. 1–11.
14. Fanelli, G., Weise, T., Gall, J., & Gool L. (2011). Real time head pose estimation from consumer depth cameras. *Pattern Recognition*, Vol. 68, No. 35, pp. 101–110.
15. Huang, Y., Wang, Y., & Tan, T. (2006). Combining statistics of geometrical and correlative features for 3D face recognition. *British Conference on Machine Vision*, pp. 879–888.
16. Xiaoguang, L. & Jain, A.K. (2004). Ethnicity Identification from Face Images. *Biometric Technology for Human Identification*.
17. Boutellaa, E., Bengherabi, M., Ait-Aoudia, S., & Hadid A. (2014). How much information Kinect facial depth data can reveal about identity, gender and ethnicity? *European Conference on Computer Vision*. Zurich, Switzerland, pp. 725–736.
18. Han, J., Shao, L., Xu, D., & Shotton, J. (2013). Enhanced computer vision with Microsoft Kinect sensor: A review. *Cybernetics, IEEE Transactions on*, Vol 43, No. 5, pp. 1318–1334.
19. Andersen, M., Jensen, T., Lisouski, P., Hansen, A., Gregersen, T., & Ahrendt, P. (2012). *Kinect depth sensor evaluation for computer vision applications*. Technical report, Department of Engineering, Aarhus University.
20. Fu, Y., Guo, G., & Huang, T. (2010). Age synthesis and estimation via faces: a survey. *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 32, pp. 1955–1976.
21. Fu, Y. & Huang, T.S. (2008). Human age estimation with regression on discriminative aging manifold. *IEEE Trans. Multimedia*, Vol. 10, pp. 578–584.
22. Geng, X., Zhou, Z., & Smith-Miles, K. (2007). Automatic age estimation based on facial aging patterns. *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 29, pp. 2234–2240.
23. Ramanathan, N. & Chellappa, R. (2006). Face verification across age progression. *IEEE Trans. Image Process.*, Vol. 15, pp. 3349–3361.
24. Zebrowitz, L.A. (1997). *Reading faces: Window to the soul*. Westview Press.

25. **Alberta, A.M., Ricanek, K., & Patterson, E. (2007).** A review of the literature on the aging adult skull and face: implications for forensic science research and applications. *Forensic Sci. Int.*, Vol. 172, pp. 1–9.
26. **Mark, L.S., Pittenger, J.B., Hines, H., Carello, C., Shaw, R.E., & Todd, J.T. (1980).** Wrinkling and head shape as coordinated sources for age-level information. *Percept Psychophys*, Vol. 27, pp. 117–124.
27. **Park, U., Tong, Y., & Jain, A.K. (2010).** Age invariant face recognition. *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 32, pp. 947–954.
28. **Ramanathan, N. & Chellappa, R. (2005).** Face verification across age progression. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 462–469.
29. **Dehshibi, M.M. & Bastanfard, A. (2010).** A new algorithm for age recognition from facial images. *Signal Process*, Vol. 90, pp. 2431–2444.
30. **Zimblar, M.S., Kokosa, M.S., & Thomas, J.R. (2001).** Anatomy and pathophysiology of facial aging. *Facial Plast. Surg. Clin. N. Am.*, Vol. 9, pp. 179–187.
31. **El-Dib, M. & El-Saban, M. (2010).** Human age estimation using enhanced bio-inspired features (EBIF). *Proc. 17th IEEE Int. Conf. on Image Processing (ICIP)*, pp. 1589–1592.
32. **Cootes, T.F., Taylor, C.J., Cooper, D.H., & Graham, J. (1995).** Active shape models-their training and application. *Comput. Vision Image Understand*, Vol. 61, No. 1, pp. 38–59.
33. **Van-Ginneken, B., Frangi, A.F., Staal, J.J., Romeny, B.M., & Viergever, M. (2002).** Active shape model segmentation with optimal features. *IEEE Trans. Med. Imaging*, Vol. 21, No. 8, pp. 924–933.
34. **Cootes, T.F., Taylor, C.J., Cooper, D.H., & Graham, J. (1995).** Active shape models — their training and application. *Comp. Vision Image Underst.*, Vol. 61, pp. 38–59.
35. **Fu, Y. & Huang, T.S. (2008).** Human age estimation with regression on discriminative aging manifold. *IEEE Transactions on Multimedia*, Vol. 10, No. 4, pp. 578–584.
36. **Wu, T., Turaga, P., & Chellappa, R. (2012).** Age estimation and face verification across aging using landmarks. *IEEE Transactions on Information Forensics and Security*, Vol. 7, No. 6, pp. 1780–1788.
37. **Mussel-Cirne, M.V. & Pedrin, H. (2018).** Combination of texture and geometric features for age estimation in face images. *13th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP)*, Vol. 4, pp. 395–401.
38. **Angulu, R., Tapamo, J.R., & Adewumi, O. (2018).** Age estimation via face images: a survey. *EURASIP Journal on Image and Video*, Vol. 42.
39. **Dyche (2001).** *The CRM handbook: A business guide to customer relationship management.* Addison-Wesley,
40. **Sarhan, S., Hamad, S., & Elmougy, S. (2016).** Human injected by botox age estimation based on active shape models, speed up robust features, and support vector machine. *Pattern Recognition and Image Analysis*, Vol. 26, pp. 617–629.
41. **Lanitis, A., Taylor, J., & Cootes, T.F. (2002).** Toward automatic simulation of aging effects on face images. *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 24, pp. 442–455.
42. **Guo, G., Fu, Y., Dyer, C., & Huang, T. (2008).** Image-based human age estimation by manifold learning and locally adjusted robust regression. *IEEE Trans. Image Process.*, Vol. 17, pp. 1178–1188.
43. **Guo, G. & Mu, G. (2013).** Joint estimation of age, gender and ethnicity: CCA vs PLS. *Proceedings of IEEE Conference on Face and Gesture Recognition*, pp. 1–6.
44. **Guo, G., Fu, Y., Huang, T.S., & Dyer, C. (2008).** Locally adjusted robust regression for human age estimation. *Proceedings of IEEE Workshop on Applications of Computer Vision*.
45. **Yan, S., Wang, H., Tang, X., & Huang, T.S. (2007).** Learning auto-structured regressor from uncertain non-negative labels. *Proceedings of IEEE Conference on Computer Vision*.
46. **Nguyen, D.T., Cho, S.R., & Park, K.R. (2015).** Age estimation-based soft biometrics considering optical blurring based on symmetrical sub-blocks for MLBP. *Symmetry*, pp. 1882–1913.
47. **Lu, J. & Tan, Y. (2013).** Ordinary preserving manifold analysis for human age and head pose estimation. *IEEE Trans. Hum. Mach. Syst.*, Vol. 43, pp. 249–258.

48. **Onifade, O.F.W. & Akinyemi, D.J. (2015).** A groupwise age ranking framework for human age estimation. *Int. J. Image Graphics Signal Process*, Vol. 5, pp. 1–12.
49. **Lanitis, A., Draganova, C., & Christodoulou, C. (2004).** Comparing different classifiers for automatic age estimation. *IEEE Trans. Man Syst. Cybern.*, Vol. 34, pp. 621–628.
50. **Ueki, K., Hayashida, T., & Kobayashi, T. (2006).** Subspace-based age group classification using facial images under various lighting conditions. *Proceedings of IEEE Conference on Automatic Face and Gesture Recognition*, pp. 43–48.
51. **Huerta, I., Fernandez, C., Segura, C., Hernando, J., & Prati, A. (2015).** A deep analysis on age estimation. *Pattern Recognit. Lett.*, Vol. 68, pp. 239–249.
52. **Bay, H., Tuytelaars, T., & Gool, L.V. (2006).** Surf: Speeded up robust features. *Comput. Vis.-ECCV*, Vol. 3951, pp. 404–417.
53. **Triggs, B. & Dalal, N. (2005).** Histograms of oriented gradients for human detection. *Proceedings of IEEE on Computer Vision and Pattern Recognition*, pp. 886–893.
54. **Hu, Z., Wen, Y., Wang, J., Wang, M., Hong, R., & Yan, S. (2016).** Facial age estimation with age difference. *IEEE Trans. Image Process.*, pp. 1–13.
55. **Fgnet (2016).** <http://sting.cyclcollege.ac.cy/alanitis/fgnetaging/>
56. **Ricanek, K. & Tesafaye, T. (2006).** Morph: A longitudinal image database of normal adult age-progression. *Proceedings of IEEE 7th International Conference on Automatic Face and Gesture Recognition*, pp. 341–345.
57. **FaceR (2015).** <http://www.face-rec.org/databases>
58. **Min, R., Kose, N., & Dugelay, J. (2014).** KinectFaceDB: A Kinect database for face recognition. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, Vol. 44, No. 11, pp. 1534–1548.
59. **Stefano, B., Alberto, D.B., & Pietro, P. (2012).** Superfaces: A super-resolution model for 3D faces. *Proc. European Conference on Computer Vision*, pp. 73–82.
60. **Iwama, H., Okumura, M., Makihara, Y., & Yagi, Y. (2012).** The OU-ISIR gait database comprising the large population dataset and performance evaluation of gait recognition. *Transactions on Information Forensics and Security*, Vol. 7, No. 5, pp. 1511–1521.

Article received on 23/10/2019; accepted on 19/12/2019X.
Corresponding author is Nabila Mansouri.