On the Quantitative Estimation of Short-Term Aging in Human Faces

Marcos Ortega, Linda Brodo, Manuele Bicego, and Massimo Tistarelli

[1] University of A Coruña, A Coruña, Spain
[2] Computer Vision Laboratory
University of Sassari, Italy

Abstract. Facial aging has been only partially studied in the past and mostly in a qualitative way. This paper presents a novel approach to the estimation of facial aging aimed to the quantitative evaluation of the changes in facial appearance over time. In particular, the changes both in face shape and texture, due to *short-time aging*, are considered. The developed framework exploits the concept of "distinctiveness" of facial features and the temporal evolution of such measure. The analysis is performed both at a global and local level to define the features which are more stable over time.

Several experiments are performed on publicly available databases with image sequences densely sampled over a time span of several years. The reported results clearly show the potential of the methodology to a number of applications in biometric identification from human faces.

1   Introduction

The demanding needs of the security industry as well as the introduction of electronic passports, increased the interest in automatic face identification and verification. In both operational modalities, digital patterns extracted from the face image are compared with similar patterns stored on the passport or in a database to validate the user's identity. A very challenging issue is the definition of the optimal face representation, either holistic (based on global features) or component-based (based on local features) [1,2,3]. Regardless of the adopted representation, biometric patterns are always affected by the changes in the face appearance due to aging. The impact on the intra-class variations can often spoil the distinctiveness of a face template even after a short time since the user's enrollment. It has been reported in several evaluation studies that, even a short time elapse of one or two weeks between acquisition sessions, can induce variations in the face appearance which degrade the performance of the recognition system. In order to overcome this problem a quantitative evaluation of the changes in the facial appearance over time as well as a regular update of the stored facial template are mandatory.

Only recently aging effects on face recognition have been rigorously studied, both for adult and young people faces[1] [4,5,6,7,8,9]. In general terms the analysis

---

[1] Face images of people from childhood into adulthood (18-20 years hold).

of the changes in the face appearance over time can be analyzed over two different time scales: *long-time* as opposed to *short-time* effects. As for the former, it involves gradual and deep modifications in the face structure and texture. These can be accounted for reduction in the fat tissue, progressive degradation of the skin elasticity, reduction of the tone of the facial muscles, development of the skull bones and non uniform pigmentation of the skin [10]. The short time effects can be either due to the same phenomena as well as to other, even transient, effects such as the change in facial tan, hair and beard growth, or small textural changes in the skin. As these are related to finer and sometimes temporary changes in the face appearance, they require a different methodology than the analysis of long time aging. First and outmost, long time changes can be analyzed from coarsely sampled face images over a considerably long time span (ten years or more), while short-time aging requires to analyze face images sampled very closely within a short time (one year or less).

In this paper a quantitative approach to facial aging is presented, where the evolution of changes in the face appearance are *measured* over a short time period. Most of the work reported in the literature is related to the analysis of long time variations. In [4,9] algorithms to simulate the age progression in adult face images are presented. In [8] a computer-based automatic aging simulation technique in 3D is described, with some qualitative results. The work in [6] describes the textural changes produced by three types of geometrical manipulations on face images, which are comparable to aging effects on the face appearance. In [11] an approach to automatically estimate age based on facial patterns is presented.

The most similar study to the one presented in this paper, is reported in [7]. In this case a quantitative approach is adopted to establish the identity between two age separated face images of an individual. The computational framework is based on Bayesian functions to estimate the intra-personal subspace (differences between two age separated faces of the same individual) and the extra-personal subspace (differences between different individuals). The estimated subspaces are used to correctly classify the face image of a given subject.

The estimation framework presented in this paper does not rely on a probabilistic measurement of the changes in the facial appearance but rather on the direct measurement of the time evolution of the face shape. In order to capture the global changes in the face an holistic approach is adopted. Given the short time variability and the transient nature of the analyzed changes, the physical phenomena underpinning the aging effects are not systematically analyzed in this paper. Rather, a global face map is computed, which allows to define the face areas showing the same *aging factor*. This estimation allows to determine the variability of facial features over time.

2   Estimation of Time-Related Face Variability

The primary objective of the proposed approach is to quantitatively measure the face changes, either local or global, within a given time interval.

The estimation is based on the algorithm presented in [12] to compute the most distinctive facial regions of an individual by comparing its face with a

number of other faces. The concept of face differences is applied to define the distinctiveness of faces [13], where a feature is selected as distinctive if it is significantly different from any other feature in a given set (like a facial scar). In the context of short time aging, the same algorithm is then applied to a series of face images of the same person but sampled at different times. As a result, the facial regions which are more stable over time are located. These can be also considered as the most time-invariant regions in the face.

2.1   Computing Face Differences

The face difference process is defined in three steps: extraction of face patches, vectorization of each patch, and weighting of the patches. The weight represents a measure of the difference of the feature in one face from any other feature in the other face.

Since face analysis may involve processing information contained at a variety of spatial resolutions, the extraction of each patch is performed at multiple scales. Toward this end each patch is re-sampled using a log-polar mapping [14]. This mapping has been motivated by its resemblance to the distribution of the receptive fields in the human retina, where the sampling resolution is higher at the central fovea and decreases toward the periphery. The resulting sampling process ensures that each patch contains both low scale (fine resolution) and contextual (low resolution) information.

In principle, all the patches of the face may be equally useful for the analysis of the aging effects. To ensure translation-independence and to side-step coarse registration issues, the patches are extracted at the loci centered on the edges extracted from the image. In this way, translation-independence is enforced, while reducing the number of points.

It is worth noting that the use of edge points does not imply that the analysis occurs only at edges. The extracted points represent the center (the *fovea*) of the patches, which extends the sampled area also to the neighborhood. In particular, each face patch is re-sampled following a log-polar scheme representing a local space-variant remapping of the original patch. Analytically, the log-polar scheme describes the mapping postulated to occur between the retina (retinal plane (r, q)) and the visual cortex log-polar or cortical plane (x, h). As illustrated in Fig. 1(a) the size of the "receptive fields" follows a linear increment moving from the central region (fovea) outwards into the periphery.

The set of log-polar image patches, sampled at each point, are vectorized, e.g. a 20×20 pixels patch is transformed into a vector of 400 raw gray-level values which represents the face in the feature space.

Once vectorized, the patches of each face are weighted. Namely, the similarity between each patch in a face and all patches in the second face is computed. When projected to an higher dimensional space, the patches from one face image will tend to form their own cluster in the feature space, while those of the other face image ought to form a different cluster.

The degree of difference of each patch with respect to the other face can be related to its locus in the feature space. All patches of the first face, lying near
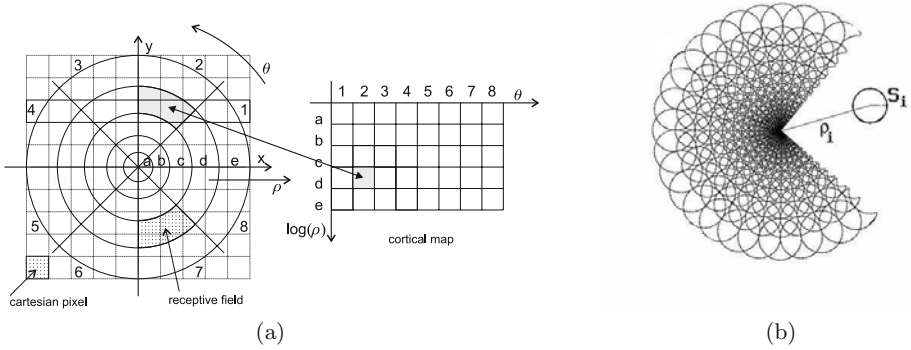
Fig. 1. (a) Log-polar sampling strategy and (b) the adopted log-polar model

the loci of a second face are not very different, and they may easily be confused with the patches of the second face. Conversely, a patch lying on the limb of its own cluster, that is most distant from any other cluster, turns out to be very different.

More formally, let P_1, P_2 be the set of patches of face 1 and 2, respectively. The weight of distinctiveness ω of a patch $p_{P_1}(x, y)$, centered at the position (x, y) in the face 1 is computed as:

$$\omega(p_{P_1}(x, y)) = d(p_{P_1}(x, y), P_2) \tag{1}$$

where

$$d(p_{P_1}(x, y), P_2) = \min_{(x', y')} d_M(p_{P_1}(x, y), p_{P_2}(x', y')) \tag{2}$$

where d_M is some distance metric between features vectors. Here, for clarity, we adopt an Euclidean metric. It might be worthwhile investigating other metrics, such as those due to transforming feature space via say a Principal Component Analysis or Linear Discriminant Analysis.

Another possibility for the extraction of weights is to train a classifier, which is able to separate the two sets P_1 and P_2 to compute the difference on the basis of the distance from the separating surface. This has been investigated in [12], using a Support Vector Machine with mixed results. However, SVM proved to be computationally demanding and much affected by the choice of the learning parameters. Still relaying on the same principle, in this paper a simpler approach is adopted to compute the facial differences which is easier to control and apply within the aging estimation context.

2.2   Face Evolution

The algorithm defined to computed the face differences can be applied to obtain a quantitative measure of the differences of the face over time. To apply the algorithm in this context, we first define the time evolution for two samples. The time evolution between two samples, I_1 and I_2, of the same face separated by a

given time t is defined as the total difference between the two face appearance. Consequently, the difference between I_1 and I_2 is computed from the patch weights of both samples. The weights are averaged to obtain a single value of the difference.

The difference between I_1 and I_2 at a point (x, y) is computed considering only those patches extracted from I_1 and I_2 containing the point (x, y). Therefore, the difference $d_{I_1,I_2}(x, y)$, between two face samples I_1 and I_2 of a sequence of images from the same face, is defined by:

$$d_{I_1,I_2}(x, y) = \frac{1}{2} \left(\frac{1}{|P_{I_1}|} \sum_{p \in P_{I_1}} \omega(p) + \frac{1}{|P_{I_2}|} \sum_{q \in P_{I_2}} \omega(q) \right) \tag{3}$$

where P_{I_1} and P_{I_2} are, respectively, the sets of patches from I_1 and I_2 containing the point (x, y), $|P_{I_i}|$ represents the cardinality of the set P_{I_i} and w is the weight determining the distinctiveness of a patch.

The computed difference value can be used as a measure of the time evolution between two samples. In a sequence of faces $S = I_1, I_2, ..., I_t, ...$, the evolution for a point (x, y), within the time span t, can be computed using $d_{I_1,I_{t+1}}$, i.e. taking the initial sample as the reference point in time and computing the difference with the sample separated by the time t. A problem arising is that the results highly depend on the actual quality of both face samples being analyzed. Changes in illumination and image noise can introduce artifacts in the estimated time evolution. On the other hand, this approximation does not take advantage of the high temporal sampling of the sequences.

To overcome these issues, instead of a single difference value between two samples, the evolution within a time span is computed by considering a window of samples of size T and comparing each one of them with its respective sample separated by t. The average of the T differences returns a more robust evolution value, as noise and illumination changes are averaged out over time.

Following this concept, the evolution $E(x, y, t)$ in a sequence, for a point (x, y) and within the time span t, is defined by:

$$E(x, y, t) = \frac{1}{T} \sum_{i=1}^{T} d_{I_i,I_{i+t}}(x, y) \tag{4}$$

2.3 Evolution of Facial Regions

In order to assess the coherence of the aging-related differences in different subjects, the quantitative results for standard facial features are compared in sequences from different individuals. The value of the time evolution for a *standard* feature (either nose, eyes or mouth) of the same individual is expected to stay within a limited and small range. At the same time, when comparing samples from different individuals for the same feature, a substantial, consistent difference in the range of values is expected over time.

If the ranges of the time evolutions for different individuals can be separated, it would be possible to classify each individual on the basis of the face evolution as a signature. In order to verify these assumptions, six standard facial features are analyzed: the nose, eyes, mouth, chin, cheeks and subnasal area. In this case, a feature in the face is defined as a window in the image. For every feature (or landmark) L the evolution over time t is defined by:

$$E(L,t) = \frac{1}{w_L h_L} \sum_{(x,y) \in L} E(x,y,t) \tag{5}$$

where w_L, h_L are, respectively, the width and the height of the region around the feature L and $E(x,y,t)$ is the evolution function for an image point as defined in equation (4).

3 Experimental Evaluation

In order to analyze the effects of short-time aging and its implications, two experiments have been carried out using sequences of faces over a time frame from 1 to 4 years. The data used in the reported experiments were extracted from a series of publicly available videos which can be downloaded from few Internet web sites[2]. This data set was chosen because the high temporal sampling in the video streams allows to analyze the time evolution of facial features over a short time span. On the other hand, other publicly available databases, such as the MORPH [3] and the FG-NET databases, contain images of adults with a much larger time difference and therefore are not well suited to perform the analysis performed in this paper.

Three video sequences (S_1, S_2 and S_3) from three different individuals have been used. The resolution of the images in all the sequences is 352x240 pixels. In Fig. 2 some samples of the sequences are shown. S_1, S_2 and S_3 cover, respectively, a time span of 4, 3 and 1 year(s). As it can be noted, the faces have been acquired approximately centered in the image and with the head always in the same position and orientation. For this reason the position of facial features does not vary over time, thus simplifying the localization process. The data are highly sampled in time at a rate of one frame per day. This allows to process close in time images thus reducing the effects of noise and temporal aliasing. Only a very limited preprocessing was performed on the face images. In order to avoid face differences due to illumination changes, the samples are preprocessed with a standard histogram equalization. To minimize the influence of outliers, the images with extreme variations such as eyeglasses (weared just once) or long beard were removed from the sequence.

---

[2] The video streams are published as *Living my life faster*. Every subject acquired a picture of his/her face every day for a time period of one to eight years.

[3] http://faceaginggroup.com
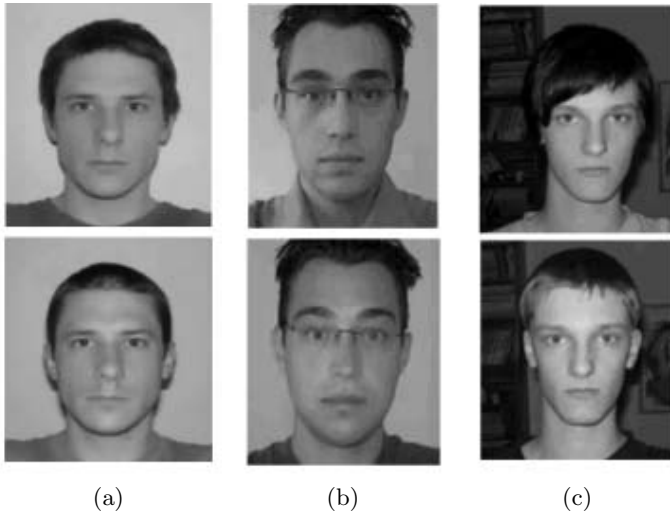
Fig. 2. Sample images extracted from the three image sequences used in the experiments. Images in (a), (b) and (c) correspond to images from sequence S_1, S_2 and S_3, respectively.

A first series of experiments was devoted to the analysis of the global evolution of the face appearance. Toward this aim, the methodology described in section 2.2 has been applied to the sequence S_1. In Fig. 3 the evolution for every face point over a time span of 4 years is shown. The brighter the value in the picture, the faster the temporal evolution of the point. These results are in line with the qualitative results reported in [10].

A second series of experiments was aimed to the analysis of the time evolution of specific, manually selected, facial regions. In particular, six regions were analyzed: the nose, eyes, mouth, chin, cheeks and subnasal area. The evolution value $E(L, t)$ has been computed for sequences S_2 and S_3. Another temporal evolution has been synthesized by crossing samples of both sequences into a single image stream. This procedure allows to compare the range of variability of the differences related to samples of the same individual, and of different individuals as well.

In Fig. 4 the results obtained analyzing different features over different time spans are shown. It is worth noting that the time evolution of every single feature, for the same subject, remains within a limited range. At the same time, the mean value of the temporal evolution of the same feature, but in different subjects, is always considerably different. The estimated difference in the range of values for different individuals is consistently in the order of 50% to 80%. This difference can be easily exploited to devise a robust template embedding the time evolution of each subject's feature for identification or verification purposes.
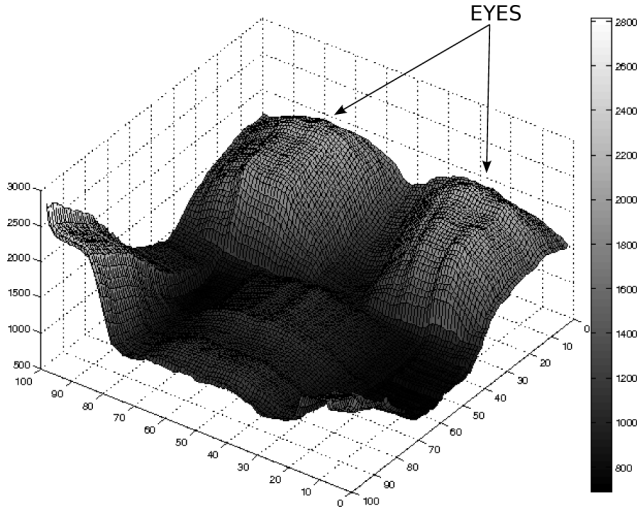
Fig. 3. Evolution for the subject in sequence 1 over 4 years. The area depicted corresponds to a 100x100 image window centered on the face. The elevation of each pixel represents the temporal evolution of the corresponding face point. As expected, the most variable regions (the most elevated in the face) appear to be those containing the eyes.

3.1   Discussion

Even the best learning algorithm can not cope for every possible variation in the face appearance. This is especially true for natural developments and changes due to aging. For this reason most identification algorithms fail to recognize individuals whenever the acquired face image is distant in time from the enrollment session. Sometimes even acquiring images of the same subject at different times in the day can result in noticeable changes in the face image. For these reasons, the analysis of long-time aging effects must be coupled with a short-time quantitative estimation of the time variability of facial features.

The quantitative results obtained in the experimental analysis can be further exploited in a number of practical applications of biometric identification from face images. As an example we may consider the following:

- The time evolution can allow to define the stability of facial features over time to devise more robust facial templates, which incorporate this measure in the face model.
- The analytical formulation of the temporal evolution of a subject's face may facilitate the update of a facial template. Both the global and local estimation of the temporal evolution can be applied as an indicator of the optimal time to update the template or part of it.
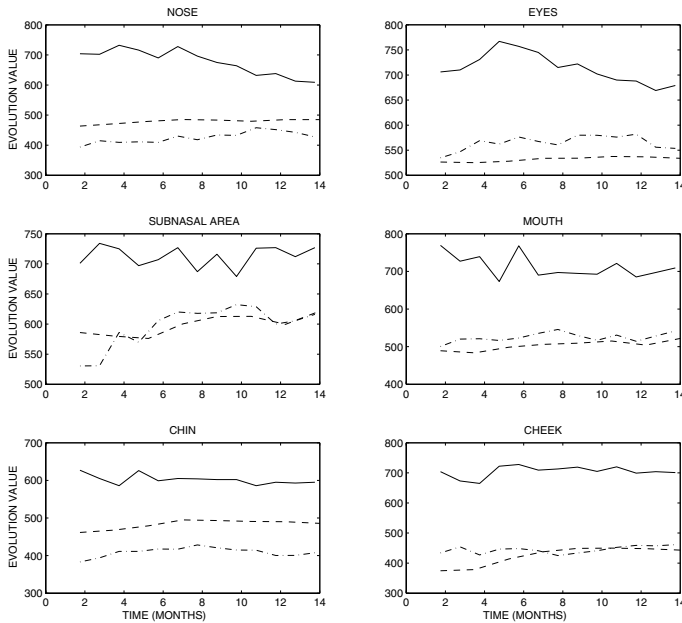
Fig. 4. Time evolution of facial differences obtained for six selected facial features: eyes, nose, mouth, chin, cheek and subnasal area. The solid lines report, for each facial feature, the differences between samples from sequences of different subjects. The dashed lines represent the computed differences between samples of sequence S_2 and the dash-dotted lines correspond to the computed differences between samples of sequence S_3.

– The quantitative evaluation of the temporal evolution for a subject may allow to forecast critical times when the template update or even re-enrollment would be required.

4   Conclusions

In this paper a novel approach to the quantitative estimation of facial aging has been presented. Differently from the majority of previous approaches, the proposed algorithm is aimed to the quantitative estimation of the temporal evolution of face patterns, within a short time span. The analysis has been performed both at a local and global scale allowing to define the time variability of individual facial features. Experiments have been performed on several sequences from different subjects highly sampled in time. The obtained results, not only demonstrate the possibility to quantitatively evaluate the temporal evolution of the face appearance over a short time scale, but also the high potential for

applications to robust personal identity verification. As a further development of this work the application of the time evolution of facial features as a signature for personal identification will be exploited.

References

1. Zhao, W., Chellappa, R., Phillips, P.J., Rosenfeld, A.: Face recognition: A literature survey. ACM Computing Surveys 35(4), 399–458 (2003)
2. Zhao, W., Chellappa, R.: Face Processing: Advanced Modeling and Methods. Academic Press Inc., USA (2006)
3. Abate, A.F., Nappi, M., Riccio, D., Sabatino, G.: Face Recognition: A Survey on 2D and 3D Techniques Pattern Recognition Letters 28(14), 1885–1906 (2007)
4. Lanitis, A., Taylor, C.J., Cootes, T.F.: Toward Automatic Simulation of Aging Effects on Face Images. IEEE Transactions on Pattern Analysis and Machine Intelligence 24(4), 442–455 (2002)
5. Lanitis, A., Draganova, C., Christodoulou, C.: Comparing different classifiers for automatic age estimation. IEEE Transactions on Systems, Man and Cybernetics - Part B 34(1), 621–628 (2004)
6. Miyoshi, T., Hyodo, M.: Aging Effects on face images by varying vertical feature placement and transforming face shape. Proceedings of IEEE International Conference on Systems, Man and Cybernetics 2, 1548–1553 (2006)
7. Ramanathan, N., Chellappa, R.: Face verification across age progression. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 2, pp. 462–469 (2005)
8. Park, U., Tong, Y., Jain, A.K.: Face Recognition with Temporal Invariance: A 3D Aging Model. In: Proceedings of IEEE Conference on Automatic Face and Gesture Recognition (to appear)
9. Wang, J., Shang, Y., Su, G., Lin, X.: Simulation of Aging Effects in Face Images. LNCIS, vol. 345, pp. 517–527. Springer, Heidelberg (2006)
10. Patterson, E., Sethuram, A., Albert, M., Ricanek, K., King, M.: Aspects of age variation in facial morphology affecting biometrics. In: Proceedings of IEEE Conference on Biometrics: Theory, Applications and Systems, pp. 1–6 (2007)
11. Geng, X., Zhou, Z.H., Smith-Miles, K.: Automatic age estimation based on facial aging patterns. IEEE Pattern Analysis and Machine Intelligence 29(12), 2234–2240 (2007)
12. Bicego, M., Grosso, E., Tistarelli, M.: On finding differences between faces. In: Kanade, T., Jain, A., Ratha, N.K. (eds.) AVBPA 2005. LNCS, vol. 3546, pp. 329–338. Springer, Heidelberg (2005)
13. Bicego, M., Brelstaff, G., Brodo, L., Grosso, E., Lagorio, A., Tistarelli, M.: Distinctiveness of faces: a computational approach. ACM Trans. on Applied Perception 5(2), 11:1–11:18 (2008)
14. Grosso, E., Tistarelli, M.: Log-polar stereo for anthropomorphic robots. In: Vernon, D. (ed.) ECCV 2000. LNCS, vol. 1842, pp. 299–313. Springer, Heidelberg (2000)