8th IEEE International Conference on Advanced Video and Signal-Based Surveillance, 2011

Frontal–to–side face re–identification based on hair, skin and clothes patches

Antitza Dantcheva
Multimedia Communications Department
EURECOM, Sophia Antipolis, France
Antitza.Dantcheva@eurecom.fr

Jean-Luc Dugelay
Multimedia Communications Department
EURECOM, Sophia Antipolis, France
Jean-Luc.Dugelay@eurecom.fr

Abstract

Despite recent advances, face-recognition algorithms are still challenged when applied in the setting of video surveillance systems which inherently introduce variations in the pose of subjects. The present work addresses this problem, and seeks to provide a recognition algorithm that is specifically suited for a frontal-to-side re-identification setting. Deviating from classical biometric approaches, the proposed method considers color- and texture- based soft biometric traits, specifically those taken from patches of hair, skin and clothes. The proposed method and the suitability of these patch-based traits are then validated both analytically and empirically.

1. Introduction

Typically biometric face-recognition algorithms are developed, trained, tested and improved under the simplifying assumption of frontal-to-frontal person recognition. Such algorithms though are challenged when facing scenarios that deviate from the training setting, such as for example in the presence of non-constant viewpoints, including the frontal-to-side scenario. Most person recognition algorithms, whether holistic or based on facial features, only manage to optimally handle pose differences that are less than about 15 degrees. As a result, a variation in the pose is often a more dominant factor than a variation of subjects. This aspect of pose variation comes to the fore in video surveillance, where a suspect may be pictured firstly frontal, whereas the corresponding test images could be captured from the side, thus introducing a *frontal-to-side recognition problem*.

Towards handling this problem, we draw from the way humans perform frontal-to-side recognition, that is by using simple and evident traits like hair, skin and clothes color. One of our tasks here is to get some insight into the significance of these traits, specifically the significance of using hair, skin and clothes patches for frontal-to-side re-identification. We mention that we work on the color Feret dataset [1] with frontal gallery images for training, and side (profile) probe images for testing. Towards achieving re-identification, the proposed algorithm first analyzes the color and texture of the three patches, as well as their intensity correlations. This analysis is then followed by the construction of a single, stronger classifier that combines the above measures, to re-identify the person from his or her profile.

1.1. Related work

Pose invariant face recognition has been addressed in different approaches which, as described in [11], can be classified in following three categories:

- mapping methods: construction of a 3D model based on more than one 2D images (cf. [6])
- geometric methods: construction of a 3D model based on a single 2D image (cf. [14])
- statistical methods: statistical learning methods that relate frontal to non-frontal poses (cf. [11].

An overview of these frontal-to-side face recognition methods was given in [18] which also addressed some of the methods' limitations in handling different pose variations. Such methods can be originally found in [16] and [17], which recorded a true detection rate of 50-60% over an authentication group of 100 subjects. Better results on pose-variant face recognition were recorded in [11] which employed statistical methods to achieve reliability of 92% over an authentication group with the same size.

This present work takes a rather different and more direct approach in the sense that the proposed method does not require mapping or a-priori learning. In applying this approach, we provide a preliminary study on the role of a specific set of simple features in frontal-to-side face recognition. These selected features are based on *soft biometrics*, as such soft biometric based features are highly applicable in video surveillance scenarios. To clarify, soft biometrics are weak biometric characteristics, which are not as permanent and distinctive as classical biometrics, but which instead include related benefits such as computational efficiency and the ability to be acquired without consent and cooperation by the user. Prevalently soft biometrics have been used for search pruning (cf. [7], [4]), but they have been furthermore

978-1-4577-0845-9/11/$26.00

employed for continuous authentication [10], [12] and re-identification [2], [8], [3]. *Re-identification* in this context is the correctly relating of data to an already explicitly identified subject (cf. [9]). Our used traits of hair, skin and clothes color and texture also belong in this category of soft biometric traits, and are thus of special interest for video surveillance applications.

2. System model, simulations and analysis

In what follows we empirically study the error probability of a system for extraction and classification of properties related to the above described color soft biometric patches. We first define the operational setting, and also clarify what we consider to be system reliability, addressing different reliability related factors, whose impact we examine. The clarifying experiments are then followed by a preliminary analytical exposition of the error probability of a combined classifier consisting of the above described traits.

2.1. Operational scenario

The setting of interest corresponds to the re-identification of a randomly chosen subject (*target subject*) out of a random authentication group, where subjects in gallery and probe images have approximately 90 degrees pose difference.

Patches retrieval: We retrieve automatically patches corresponding to the regions of hair, skin and clothes, cf. Figure 1, based on the coordinates of face and eyes. The size of each patch is determined empirically by one half and one third distance of center-to-center eye distance. The frontal placement of the patches is following, horizontally / vertically respectively:

- hair patch: from nose towards left side / top of the head
- skin patch: centered around left eye / centered between mouth and eyes
- clothes patch: from left side of the head towards left / mouth-(mouth-top of head distance)

The horizontal side face placements of the patches are: nose-(nose-chin distance), eye towards left, head length-chin. Those coordinates were provided along with the images from the Feret database, but are also easily obtainable by state of the art face detectors, like the OpenCV implementation of the Viola and Jones algorithm [15].

Following the extraction of the patches we retrieve a set of feature vectors including the color and texture information. In this operational setting of interest, the *reliability* of our system captures the probability of false identification of a randomly chosen person out of a random set of N subjects. This reliability relates to the following parameters:

- Factor 1. The number of categories that the system can identify. This number is denoted by ρ.

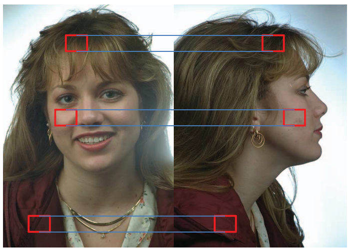

Figure 1. Frontal / gallery and profile / probe image of a subject. Corresponding ROIs for hair, skin and clothes color.

- Factor 2. The degree with which these features / categories represent the chosen set of subjects over which identification will take place.
- Factor 3. The robustness with which these categories can be detected.

Finally reliability is related (empirically and analytically) to N, where a higher N corresponds to identifying a person among an increasingly large set of possibly similar-looking people.

We will proceed with the discussion and illustration of the above three named pertinent factors.

2.2. Factor 1. Number of categories that the system can identify

Our system employs soft biometric traits, where each trait is subdivided into *trait-instances*, *e.g.* the trait hair color can take the values (trait-instances) *blond*, *brown*, *red*, *black*. The number of overall categories ρ is determined by the traits and instances a system is endowed with. We here note that in the scientific work [12] the terminology trait and semantic term was used instead. For this specific scenario, our system is endowed with the traits hair, skin and shirt color as well as texture and intensity correlation. It is to be noted that the more categories a system can classify subjects into, the more reliable the system is in terms of distinctiveness.

We proceed with the description of population of categories.

2.3. Factor 2. Category representation of the chosen set of subjects

The distribution of subjects over the set of ρ categories naturally has an impact on the specific importance of the different traits and thus also affects the behavior of the system for a given population. Table 1 and Table 2 give example distributions, with respect to skin and hair color traits, based on 265 subjects from the color Feret database. For clarifying experiments we annotated this subset into three subcategories of skin color and eight subcategories of hair color. It is evident that the trait hair color has a higher distinctiveness and thus importance than skin color, since it has more instances in which the subjects are subdivided into. This observation is illustrated in Figure 2 and we follow with the explanation.

Categories	1	2	3
Skin Color	62.64%	28.66%	8.7%

Table 1. Distribution of Feret subjects in the three skin color categories.

Categories	1	2	3	4	5	6	7	8
Hair Color %	4.2	26.8	46.4	3	3	3.4	1.5	11.7

Table 2. Distribution of Feret subjects in eight hair color categories.

In the case of (re-)identification the only way to unambiguously recognize a person is an exclusive category membership, with other words a subject can only be (re-)identified if he or she is the only one assigned to a category. Hereby it is of interest, that the target subject does not collide with any other subject inside the authentication group in terms of category. *Collision* in this context is the event of two or more subjects belonging to the same category. An analysis relating to the birthday paradox can shed light on the related collision probability. For the uniformity assumption the collision probability $q(N)$ carries following closed form expression:

$$q(N) = 1 - (\frac{\rho - 1}{\rho})^N. \tag{1}$$

In the example, of subjects-categories distribution probabilities illustrated in Table 1 and Table 2 the probabilities for collision for skin and hair color as functions of N are portrayed in Figure 2. For this experiment we randomly pick N subjects out of the color Feret subset, pick randomly one of the N subjects as target subject and examine if the target subject collides with somebody else from the given authentication group. We repeatedly perform this experiment and compute finally an averaged collision error probability as a function of N.

This collision error can be decreased by considering more distinctive or multiple traits. We note here that the collision error probability is the minimum error bound a system can achieve given the specific traits in a certain population.

In what follows we will include automatic category classification and analyze the influence of estimation errors caused by algorithm limitations.

2.4. Factor 3. Robustness of categories estimation

In this section we take into consideration the over all error probability, containing inevitably of the above discussed collision error- and furthermore of the algorithmic categorization error- probabilities. With other words we examine,

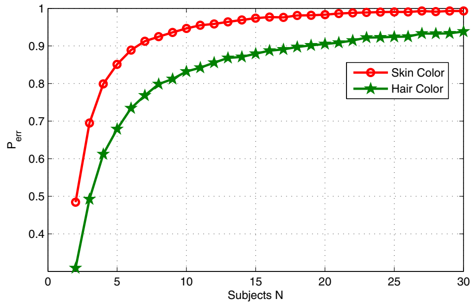

Figure 2. Probability of collision in the Feret database for skin and hair color.

how often the target subject is wrongly re-identified, regardless of the underlying source, which can be both estimation- or collision error character. Thereby we again randomly pick an authentication group of N subjects and randomly declare one of the N subjects as the target subject for authentication. Then we proceed to train our algorithms with the feature vectors extracted from the patches of the gallery images. Following the training, we re-identify the target subject by matching his or her feature vectors of the probe patches (retrieved from the profile image) with the trained ones (frontal). We repeat the procedure and average the error probability over all iterations for all values of N. As feature vectors we firstly consider the color information provided by the patches. For this purpose we work in the HSV (hue, saturation, value) color space and use the hue and saturation information as feature vectors. The according classification is performed by the AdaBoost algorithms [5], which method clearly outperformed classification done in the range of this work by Gaussian Mixture Models. We note here that we do not consider anymore the manual annotation of categories used for the experiment in Section 2.3, instead we use the discrete HS distribution of colors. By doing so we do not have anymore human compliant categories, but a more refined and efficient classification.

2.4.1 Boosting of patch color

Boosting is referred to *additive logistic regression* and is the combination of weak classifiers, which classifiers alone should perform slightly better than random into one higher accuracy classifier. This combined classifier is a sum of the weighed weak classifiers, where the weight of each classifier is a function of the accuracy of the corresponding classifier. This concept is compliant with the concept of soft biometrics, where we have a multitude of weak biometric information that we want to combine to a stronger classification hypothesis. We selected a discrete AdaBoost.MH algorithm, cf. [13] for multi class classification, for its good performance and robustness against overfitting.

We train hue and saturation (HS) per gallery patch using AdaBoost and evaluate posterior probabilities of the HS

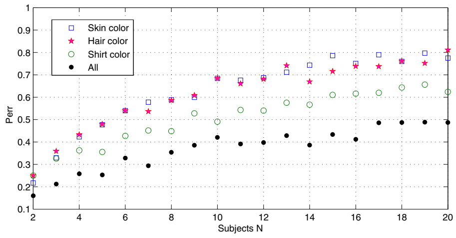

Figure 3. Boosting of soft biometric patches.

vectors of the probe patch related to the target authentication subject and each patch of the N trained subjects. An error occurs, if the HS vectors of gallery and probe patch of the target authentication subject do not match. The results on error probability as a function of the authentication group size N for each patch color and the combined color patches are illustrated in 3. As expected, and similar to the collision experiment in Section 2.3, the re-identification error probability is increasing with an increasing authentication group. We further observe that clothes color has the strongest distinctiveness. The explanation therefore is that clothes color is distributed in a high range of colors and is thus easier to distinguish. Furthermore we surprisingly notice a much higher performance of skin or hair color than in the pure collision analysis, see Figure 2. For example, in an authentication group of 5 subjects, AdaBoost re-identification provides an error probability for skin color of about 0.48, where in the collision analysis we achieve only 0.85. This interesting error decrease is due to the limited human capability for distinguishing and classification of skin color in only three categories, which classification was only used in the collision analysis. In the current AdaBoost estimation setting the classification is hidden, non human compliant but of a higher efficiency.

As expected the combined classifier has a stronger classification accuracy than the single classifiers. The performance of the combined classifier is though enticed by several factors. Firstly the traits used are partly correlated, so is for example the Pearson's correlation coefficient defined for two variables λ_1 and λ_2 as:

$$r_{\lambda_1,\lambda_2} = \frac{cov(\lambda_1,\lambda_2)}{\sigma_{\lambda_1}\sigma_{\lambda_2}} = \frac{E[(\lambda_1-\mu_{\lambda_1})(\lambda_2-\mu_{\lambda_2})]}{\sigma_{\lambda_1}\sigma_{\lambda_2}} \quad (2)$$

for hair and skin color given by $r_{HairColor,SkinColor} = -0.14$. Furthermore the estimation errors of traits are correlated as well, $r_{Estimationerror(HairColor,SkinColor)} = 0.22$, which shows a tendency of jointly occurrence of classification errors, for both hair and skin color classification. On a different note we point that each further trait and its classification error contributes negatively to the over all categorization error and thus the over all error probability is increasing with an increasing number of traits and categories ρ. On the other hand with each further trait the collision probability decreases.

We proceed with the description and inclusion of two further properties of the employed patches, namely texture and intensity difference.

2.4.2 Patch texture

We formalize a descriptor for texture $\overrightarrow{x}$ including following four characteristics and compute them on the graylevel images for each patch.

Contrast: measure of the intensity contrast between a pixel and its neighbor over the whole image. The contrast in an image is related to its variance and inertia and is:

$$x_1 = \sum_{i,j} |i-j|^2 p(i,j), \quad (3)$$

where i and j are the gray scale intensities of two pixels, p is the gray level co-occurrence matrix, which describes the co-occurrence of gray scale intensities between two image areas. Each element (i, j) in gray level co-occurrence matrix specifies the number of times that the pixel with value i occurred horizontally adjacent to a pixel with value j.

*Correlation*is a measure for how correlated pixels are with their neighbors and is denoted as:

$$x_2 = \sum_{i,j} \frac{(i-\mu_i)(j-\mu_j)\vec{p}(i,j)}{\sigma_i\sigma_j}, \quad (4)$$

where μ_i and μ_j are the mean values of the two areas around i and j, σ_i and σ_j are the related standard deviations.

*Energy*is a sum of squared elements or angular second moment. Energy equal to one corresponds to a uniform color image.

$$x_3 = \sum_{i,j} p(i,j)^2 \quad (5)$$

Homogeneity is a measure of the closeness of distribution of elements.

$$x_4 = \sum_{i,j} \frac{p(i,j)}{1+|i-j|} \quad (6)$$

2.4.3 Patch histogram distance

Along with the color information we integrate into our classifier a simple relation measure for the divergence between the intensity probability density functions (pdf) of patches concerning one subject. With other words we express the three relationships between intensities within a subject: hair–skin, skin–clothes and hair–clothes. Speaking in an example we expect to have a higher distance measure for a person with brown hair and light skin than for a person with blond hair and light skin. For the computation we convert the patches to gray level intensities and assess the L1–distance three times per person for all relations between the patches. For two distributions r and s of discrete

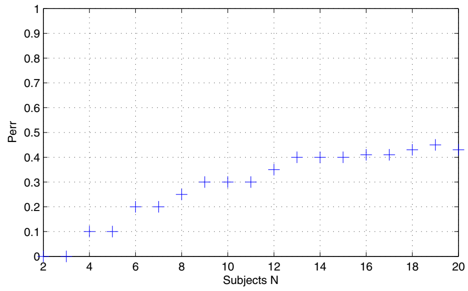

Figure 4. Over-all-classifier obtained by boosting color, texture and intensity differences.

random character the measure is given as:

$$D = \|r - s\|_1 = \sum_{k=1}^{255} |r(k) - s(k)|, \quad (7)$$

where k represents a bin of the 255 intensity bins in a gray scale image.

2.5. Combined over-all-classifier

The combined over–all–classifier, which boosts all described traits, color, texture and intensity differences performs with a decreased error probability and thus outperforms expectedly the color classifier shown in Figure 3. Still the achieved error probability of 0.1 in an authentication group of 4 subjects is not sufficient enough for a robust re-identification system. This limited enhanced performance is due to the correlations between traits, e.g. hair color–skin color or skin color–skin texture, cf. [2]. To improve this performance the amount of sub-classifiers can be further extended. The system in its current constellation can be used as a pruning system for more robust systems or as an additional system for multi-trait biometric systems.

3. Conclusions

Motivated by realistic surveillance scenarios, the work addressed the problem of frontal-to-side facial recognition, providing re–identification algorithms/classifiers that are specifically suited for this setting. Emphasis was placed on classifiers that belong in the class of soft biometric traits, specifically color–, texture– and intensity– based traits taken from patches of hair, skin and clothes. Towards providing insight, the work presented different identification experiments that adhere to the frontal–to–side setting, as well as presented a preliminary analytical study that seeks to impart intuition on the role of the above traits in improving algorithmic reliability. Our analysis described the overall error probability, both as a function of collisions and of erroneous categorizations for given sizes of authentication groups. In the presence of a moderate reliability of the patches-based method, the analysis suggests promising applications of this method in settings such as pruning of searches.

This work was supported in part by the French national project BIORAFALE.

References

[1] face.nist.gov/colorferet/. 1

[2] A. Dantcheva, J.-L. Dugelay, and P. Elia. Person recognition using a bag of facial soft biometrics (BoFSB). In *Proceedings of IEEE MMSP*, 2010. 2, 5

[3] A. Dantcheva, J.-L. Dugelay, and P. Elia. Soft biometrics systems: Reliability and scaling laws. In *Proceedings of BTAS*, 2010. 2

[4] A. Dantcheva, A. Singh, P. Elia, and J.-L. Dugelay. Search pruning with soft biometric systems: efficiency-reliability tradeoff. *Research Report*, RR-11-255. 1

[5] J. Friedman, T. Hastie, and R. Tibshirani. Additive logistic regression: a statistical view of boosting. *Technical Report*, 1998. 3

[6] R. Ishiyama, M. Hamanaka, and S. Sakamoto. An appearance model constructed on 3-D surface for robust face recognition against pose and illumination variations. *IEEE Transactions System, Man, Cybernetics*, 3(35):326–334, 2005. 1

[7] A. Jain, S. Dass, and K. Nandakumar. Soft biometric traits for personal recognition systems. In *Proceedings of ICBA*, pages 1–40, 2004. 1

[8] A. Jain and U. Park. Facial marks: soft biometric for face recognition. In *Proceedings of IEEE ICIP*, 2009. 2

[9] B. Malin, L. Sweeney, and E. Newton. Trail re-identification: learning who you are from where you have been. *Data privacy laboratory technical report*, LIDAP-WP12, 2003. 2

[10] K. Niinuma, U. Park, and A. Jain. Soft biometric traits for continuous user authentication. *IEEE Transactions on information forensics and security*, 5(4), December 2010. 2

[11] S. Prince, J. Elder, J. Warrell, and F. Felisberti. Tied factor analysis for face recognition across large pose differences. *IEEE Transactions on pattern analysis and machine intelligence*, 30(6):970–984, 2008. 1

[12] S. Samangooei, B. Guo, and M. S. Nixon. The use of semantic human description as a soft biometric. In *Proceedings of BTAS*, 2008. 2

[13] R. Schapire and Y. Singer. Improved boosting algorithms using confidence-rated predictions. In *Proceedings of CCLT*, 1998. 3

[14] R. Singh, M. Vatsa, A. Ross, and A. Noore. A mosaicing scheme for pose-invariant face recognition. *IEEE Transactions System, Man, Cybernetics*, 5(37):1212–1225, 2007. 1

[15] P. Viola and M. Jones. Robust real–time face detection. In *Proceedings of IEEE ICCV*, 2001. 2

[16] F. Wallhof, S. Mueller, and G. Rigoll. Recognition of face profiles from the mugshot database using a hybrid connectionist / hmm approach. In *"Proceedings of IEEE ICASSP"*, 2001. 1

[17] F. Wallhoff, D. Arsic, B. Schuller, G. Rigoll, J. Stadermann, and A. Stoermer. Hybrid profile recognition on the mugshot database. 2005. Proceedings of IEEE EUROCON. 1

[18] X. Zhang and Y. Gao. Face recognition across pose: a review. *Pattern Recognition*, 42(1):2876–2896, 2009. 1