International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:8, No:8, 2014

Enhanced Face Recognition with Daisy Descriptors Using 1BT Based Registration

Sevil Igit, Merve Meric, Sarp Erturk

***Abstract*—**In this paper, it is proposed to improve Daisy Descriptor based face recognition using a novel One-Bit Transform (1BT) based pre-registration approach. The 1BT based pre-registration procedure is fast and has low computational complexity. It is shown that the face recognition accuracy is improved with the proposed approach. The proposed approach can facilitate highly accurate face recognition using DAISY descriptor with simple matching and thereby facilitate a low-complexity approach.

***Keywords*—**Face Recognition, Daisy Descriptor, One-Bit Transform, Image Registration.

I. Introduction

FACE recognition systems are targeted at automatically identifying or verifying a person from a digital image or video. While security is still the most frequent application of face recognition, new areas such as social networking or gaming have also started to make use of this technology.

Reference [1] has presented a classification category for face recognition systems that is well accepted [2]. In the corresponding work methods are divided into four categories that may overlap, so that an algorithm could belong to two or more categories. This classification can be made as follows:

- Knowledge-based methods: Algorithms that are rule-based and encode available knowledge of human faces.
- Feature-based methods: Algorithms that try to find scale, rotation or translation invariant features.
- Template matching methods: Algorithms that compare input images with stored patterns of faces or features.
- Appearance-based methods: Template matching methods whose pattern database is learnt from a set of training images.

A Daisy descriptor based face recognition system has been proposed in [3]. This approach proposed to use a dense computation of descriptors, that differs in this point from previous approaches such as Scale Invariant Feature Transform (SIFT) based approaches [4] that rely on a limited number of key points. The dense description is proposed to account for limited feature points that appear insufficient because a feature present in one pose can disappear or be considered unstable in another one.

An important aspect in key point extraction based approaches is that it is required to match corresponding key points extracted in query and template images, whose position may vary due to pose or expression. In [3] it is proposed to use a recursive grid search for this purpose, nonetheless the effectiveness of the approach is limited.

In this paper it is proposed to initially align (i.e. register) the faces prior to the computation and matching of Daisy features. Because the Daisy features are proposed as a computationally efficient way of computing dense features [5], it is important that this initial registration introduces a low complexity burden. Therefore it is proposed in this paper to use a low complexity One-Bit Transform (1BT) based registration approach before the feature extraction and matching process.

The One-Bit Transform (1BT) has been used previously for low complexity motion estimation in [6]. The 1BT provides a single bit-depth representation that covers the general structure of image objects/content. Therefore it is very suitable for low complexity image matching/registration and outperforms low bit-depth approaches such as bit truncation or truncated gray coding [7].

It is shown that the proposed 1BT based pre-registration approach improves the accuracy of Daisy descriptor based face recognition by allowing a simple matching criterion.

II. The Daisy Descriptor [5]

The Daisy descriptor provides a dense representation of image features. For a given image, the initial step is to compute a number (H) of orientation maps $\boldsymbol{G_i}$ $(1 \leq i \leq H)$. Here, i shows the quantized direction and $\boldsymbol{G_i}(u,v)$ is the image gradient at pixel location (u,v) if the gradient is larger than zero, otherwise it is set to zero to preserve gradient polarity. Each orientation map is then convolved with multiple Gaussian kernels of different standard deviation (Σ) to control the region size. This process can be formulated as

$$\mathrm{G}_i^{\Sigma} = G_{\Sigma} * G_{\mathrm{i}} \tag{1}$$

where G_{Σ} presents the Gaussian kernel. For computational efficiency the convolution can be carried out successively (detail is provided in [5]).

The Daisy descriptor is constructed by sampling the convolved orientation maps G_i^{Σ} at neighbor positions as shown in Fig. 1. In this figure, each circle represents a region with a radius that is proportional to the standard deviation of the Gaussian kernels G_{Σ} and "+" shows the location of the samples that make up the descriptor.

Sevil Igit, Merve Meric, and Sarp Erturk are with Kocaeli University Laboratory of Image and Signal processing (KULIS), Kocaeli, 41380, Turkey. (corresponding author: S. Erturk phone: 90-262-3033002; fax: 90-262-3033003; e-mail: sertur@ kocaeli.edu.tr).

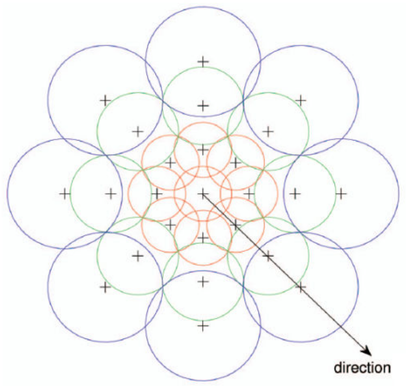

Fig. 1 The Daisy Descriptor [5]

It is possible to construct a vector $h_{\Sigma}(u,v)$ by concatenation of the values in the orientation maps after convolution with the Gaussian kernels so that

$$h_{\Sigma}(u,v) = [G_1^{\Sigma}(u,v) \quad \dots \quad G_H^{\Sigma}(u,v)] \tag{2}$$

where G_i^{Σ} is the convolved orientation map as defined in (1). All these vectors are normalized to unit norm, $\tilde{h}_{\Sigma}(u,v)$, to compensate for variations.

The Daisy descriptor is then constructed to take the different circular layers into account, in the form of

$$D(u_0,v_0) = \begin{matrix} [\tilde{h}_{\Sigma_1}(u_0,v_0), & & \\ \tilde{h}_{\Sigma_1}(I_1(u_0,v_0,R_1)), & \dots, & \tilde{h}_{\Sigma_1}(I_T(u_0,v_0,R_1)), \\ \tilde{h}_{\Sigma_2}(I_1(u_0,v_0,R_2)), & \dots, & \tilde{h}_{\Sigma_2}(I_T(u_0,v_0,R_2)), \\ \dots & \dots & \dots \\ \tilde{h}_{\Sigma_Q}(I_1(u_0,v_0,R_Q)), & \dots, & \tilde{h}_{\Sigma_Q}(I_T(u_0,v_0,R_Q))] \end{matrix} \tag{3}$$

where Q represents the number of different circular layers and T is the number of orientations. In this case, $I_j(u_0,v_0,R)$ is the location with distance R from (u_0,v_0) in the direction represented by j.

A detailed computational complexity analysis of the Daisy descriptor is provided in [5]. It is shown that the Daisy descriptor can be computed about 65 times faster than the SIFT descriptor if it is efficiently implemented.

III. The One Bit Transform (1BT) [6]

The 1BT transforms images into 1 bit/pixel representations by providing a local pixel-wise threshold that is computed using a multi band-pass filter kernel. The multiplication free 1BT presented in [6] provides a low-complexity approach for obtaining the 1BT of images. Therefore, this approach is also adopted in this paper to provide lower computational complexity. The multi band-pass filter K proposed in [6] is used for the 1BT computation.

The filtered image is computed by convolving the input image with the multi band-pass filter kernel in the form of

$$I_f(u,v) = I(u,v) * K \tag{4}$$

Because of the definition of the kernel this computation can be carried out using addition and shift operations only, without the need of multiplication.

The 1BT is then computed as

$$B(u,v) = \begin{cases} 1, & \text{if } I(u,v) \geq I_f(u,v) \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

by a simple comparison process.

Fig. 2 shows an example face image from the Olivetti Research Lab (ORL) database together with the filtered version and the 1BT result. It is seen that the 1BT result captures and extracts facial key characteristics successfully with a low complexity approach.

(a) (b) (c)

Fig. 2 (a) Original face image (b) Multi band-pass filtered image (c) 1BT result

IV. Proposed Approach

The presented approach proposes to pre-register the facial images using a low-complexity 1BT based approach and then carry out the Daisy descriptor based matching for facial recognition. The pre-matching will compensate for variations, such as in pose, to some amount, improving the accuracy of the face recognition process.

Fig. 3 shows two separate facial images of the same person form the ORL database before and after registration. Lines are superposed to the images at the same position to enable easier visual interpretation. While there is a significant amount of variation between the two facial images, it is easily observed that the presented pre-registration reduces the positional variations compared to the original version. In practice it is observed that the provided reduction is sufficient and an exact match is not required, as the matching process itself also compensated for some degree of variation.

(a) (b) (c)

Fig. 3 (a) Original first face image (b) Original second face image (c) Registered second face image

The proposed approach is depicted in Fig. 4. Initially the images are registered using 1BT based matching. For this purpose the 1BT of both images is constructed and a simple full-search strategy as presented in [6] is utilized to find the best matching position. After 1BT based registration of the facial images, the Daisy descriptor is computed for both images and the Daisy descriptors are matched to determine if the facial images are of the same person or not.

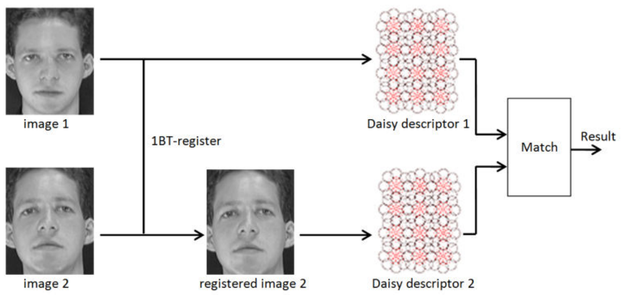

Fig. 4 Proposed face recognition system

While [3] uses a Support Vector Machine (SVM) based classification method and a recursive grid search matching for face recognition, a more simple approach is utilized in this paper. In the utilized approach, simply a $N \times N$ window is constructed around the pixel position in the test image and the minimum Euclidean distance between the Daisy descriptor of a pixel in the test image and all pixels in the window of the reference image is computed. The average Euclidean distance error for the entire face image is utilized to make the recognition decision.

V. Experimental Results

The proposed approach is evaluated using the ORL facial image database. The database includes images with a significant amount of variation (pose, gesture, etc.). The database has 40 subjects with 10 pictures each and the image resolution is 112×92 pixels.

In the experimental evaluation, the matching window size is selected as $N = 5$, so that a total of 25 computations are made for each descriptor. This is only about a quarter of the 99 feature vector elaborations utilized in [3]. A total of eight orientations are used (i.e. $T = 8$) with three circular layers (i.e. $Q = 3$).

In the experimental setup, in turn, one image is taken as reference and the distance between the test image and the reference image is computed as explained in Section IV. If the distance is below the matching threshold the two facial images are marked as matching, they are marked as non-matching otherwise. In order to avoid bias of threshold selection, experimental results are presented in terms of Receiver Operation Characteristic (ROC) curves. For this purpose the matching threshold is changed and the true positive and false positive rates are computed successively.

Fig. 5 shows the ROC curves for the ORL facial database for the proposed approach and also the original approach that does not use pre-registration. It is shown that the proposed approach significantly improves the recognition accuracy and thereby the performance of the face recognition system. Because the introduced 1BT based pre-registration is computational efficient, the low-computational complexity of the overall face recognition system is retained.

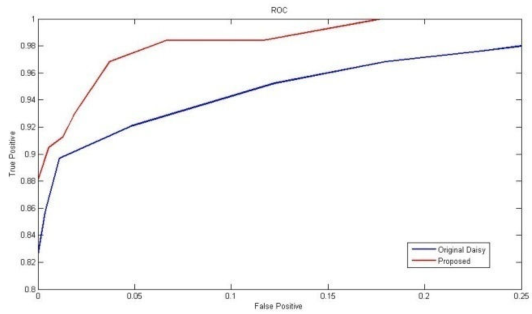

Fig. 5 ROC curve for the ORL database

VI. Conclusion

This paper proposes a novel 1BT based pre-registration before Daisy descriptor based face recognition. The proposed registration approach has low computational complexity and is shown to significantly improve the facial recognition rate of the system.

References

[1] M.-H. Yang, D. Kriegman, and N. Ahuja. Detecting faces in images: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(1): 34–58, January 2002.

[2] Ion Marques, Face Recognition Algorithms, Universidad del Pais Vasco, Thesis, June 2010.

[3] C. Velardo, J. Dugelay, Face recognition with DAISY descriptors. In Proceedings of the 12th ACM Workshop on Multimedia and Security, Roma, Italy, Sep 2010.

[4] J. Luo, Y. Ma, E. Takikawa, S. Lao, M. Kawade, and B. Lu. Person-specific SIFT features for face recognition. In IEEE International Conference on Acoustics, Speech and Signal Processing, 2007. ICASSP 2007, volume 2, 2007

[5] Engin Tola, Vincent Lepetit, Pascal Fua: DAISY: An Efficient Dense Descriptor Applied to Wide-Baseline Stereo. IEEE Trans. Pattern Anal. Mach. Intell. 32(5): 815-830, 2010.

[6] S. Ertürk, "Multiplication-free One-bit Transform for Low-complexity Block-based Motion Estimation", IEEE Signal Processing Letters, Sayı 14, No. 2, pp.109-112, Feb. 2007

[7] A. Yeni, S. Ertürk, "Fast Digital Image Stabilization using One Bit Transform Based Sub-Image Motion Estimation", IEEE Transactions on Consumer Electronics, Sayı 51, No. 3, pp.917-921, Aug. 2005.

S. Igit is a B.Sc. student in the Electronics and Telecommunications Engineering department and double major Computer Engineering department at Kocaeli University, expected to graduate in 2015.

M. Meric is a B.Sc. student in the Electronics and Telecommunications Engineering department at Kocaeli University, expected to graduate in 2015.

S. Erturk (M’99) received the B.S. degree in electrical and electronics engineering from Middle East Technical University, Ankara, Turkey, in 1995, and the M.Sc. degree in telecommunication and information systems and the Ph.D. degree in electronics systems engineering from the University of Essex, Colchester, U.K., in 1996, and 1999, respectively.

From 1999 to 2001, he carried out his compulsory service at the Army Academy, Ankara. He is currently appointed as a Full Professor with Kocaeli

University, Kocaeli, Turkey, where he worked as an Assistant Professor between 2001 and 2003, and an Associate Professor between 2002 and 2007. He is also the head of the Kocaeli University Laboratory of Image and Signal Processing (KULIS). His research interests are in the areas of digital signal and image processing, video coding, remote sensing, and digital communications.

Dr. Ertürk has authored two national books, one on digital communications and one on digital signal processing, and is the author and co-author of more than 150 scientific publications.