Available online at [www.sciencedirect.com](http://www.sciencedirect.com/)



ScienceDirect

Procedia Computer Science 70 (2015) 185 – 197

4th International Conference on Eco-friendly Computing and Communication Systems, ICECCS 2015

Two Novel Detector-Descriptor Based Approaches for Face Recognition using SIFT and SURF

Vinay Aa, Dixit Hebbarb, Vinay S Shekhara , K N Balasubramanya Murthya, S Natarajanb

*aPES University, 100 Feet Ring Road, BSK III Stage, Bangalore -560085*

*bPES Institute of Technology, 100 Feet Ring Road, BSK III Stage, Bangalore -560085*

**Abstract**

A plethora of promising detectors and descriptors are available in Computer Vision for carrying out Face Recognition (FR) and although these techniques that form the backbone of FR have yielded reasonable efficacy, they are yet to advance to those levels where they demonstrate robust performance in unconstrained scenarios. In our deliberations, we employ the popular SIFT and SURF algorithms that are ubiquitously implemented in FR due to their remarkable potency in handling a variety of FR tasks. In this paper, we proffer two novel detector-descriptor variants to augment the proficiency of contemporary FR systems:

(1) SURF detector with SIFT descriptor and (2) SIFT detector with SURF descriptor. We demonstrate the proficiency of the proposed techniques by utilizing pertinent mathematical arguments and performing comprehensive comparisons with the classical SIFT and SURF algorithms by employing a number of standard metrics over the benchmark LFW and Face94 datasets.

© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license ([http://creativecommons.org/licenses/by-nc-nd/4.0/).](http://creativecommons.org/licenses/by-nc-nd/4.0/))

© 2015 The Authors. Published by Elsevier B.V.

Peer-review under responsibility of the Organizing Committee of ICECCS 2015.

Peer-review under responsibility of the Organizing Committee of ICECCS 2015

*Keywords:*Face Recognition; Feature Description; Keypoint Detection; SIFT; SURF.

# Introduction and Related Work

Face Recognition (FR)[13][27][29] is one of the most prominent fields in the domain of Computer Vision that

**Nomenclature**

FLANN Fast Library for Approximate Nearest Neighbour Search SIFT Scale Invariant Feature Transform

SURF Speeded Up Robust Features

\* Corresponding author. Tel.: +91-80-26721983

*E-mail* [*address:*a.vinay@pes.edu](mailto:a.vinay@pes.edu)

1877-0509 © 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license ([http://creativecommons.org/licenses/by-nc-nd/4.0/).](http://creativecommons.org/licenses/by-nc-nd/4.0/))

Peer-review under responsibility of the Organizing Committee of ICECCS 2015 doi:10.1016/j.procs.2015.10.070

strives to simulate through automated machine techniques and inventive algorithms, the inherent capability of human beings in recognizing a face in an effortless and instantaneous manner. In today’s world there is a dire need to ensure the security of the massive influx of sensitive information and curb security breaches that are becoming increasingly prevalent in corporations and institutions leading to loss of earnings and reputation.FR systems can significantly augment security systems in a number of ways: (1) it is unique to every user, unlike PINS and passwords which can be misused or forgotten, (2) they require the physical presence or action by the user for authentication, (3) can be implemented on existing infrastructure through security cameras, which can capture facial data without any explicit action by the user, whereas other biometrics such as iris and fingerprint identification require expensive hardware for acquisition and (4) data acquisition is an easy task, unlike other biometric mechanisms (such as Fingerprint or Iris) which involve massive enrollment expense.

Although FR has tendered a number of effective applications in a wide gamut of cross-domain commercial and law-enforcement settings, its performance tends to decline in unconstrained scenarios [15] i.e. images acquired under scenarios where there is a sharp variation in parameters such as scale, illumination, rotation, pose, expression, affine, translation and so on. Typically, real world images tend to contain such variations, as they are acquired from assorted sources due to which they do not conform to any common specifications. Hence the search for better FR techniques that can demonstrate improved robustness in a wide variety of unconstrained settings is ongoing [27][29].

The key-point detection and feature description stages are significantly crucial in an FR system, as they form the backbone of the entire process and hence their selection dictates the overall performance and recognition accuracy. There are a number of accomplished detector and descriptor mechanisms that are available (see [14] for an in-depth review of contemporary feature detectors and descriptors) for developing an FR system. SIFT (Scale Invariant Feature Transform) [4][11][12]is a potent algorithm developed by David Lowe, that dubs as both a key point detector and feature descriptor and has been ubiquitously employed in a number of computer vision tasks such as object recognition, image stitching, visual mapping etc. and is preferred for the wide gamut of standout characteristics it possesses such as its robustness to variations in scale, zoom, translation, rotation, illumination and so on. Similarly, SURF (Speeded up Robust Features) [3][26][28]is a computationally less expensive alternative to SIFT that was developed by Herbert Bay in 2006,which has demonstrated better performance and accuracy over SIFT [4][19]. SURF has several standout characteristics like repeatability, distinctiveness, robustness, scale and rotation invariance, and can be computed quickly as it relies on integral images for image convolutions to reduce the computation time by employing a fast Hessian matrix based measure. Both SIFT and SURF are landmark techniques in the field of FR [1], whose effectiveness has been extensively demonstrated [14], so we opted to employ them for our deliberations. Furthermore, in order to match the feature descriptors to determine if it matches a face in the database, we employ the FLANN (Fast Library for Approximate Nearest Neighbour Search) [16] algorithm.

In this paper, we introduce two novel techniques for FR. The first employs SIFT as a keypoint detector and SURF as the feature descriptor and the second uses SURF as the key-point detector and SIFT as the feature descriptor. We will demonstrate that these two methods are considerably potent and can markedly augment FR accuracy over the classical SIFT (SIFT as both detector and descriptor) and classical SURF (SURF as both detector and descriptor) methods using pertinent mathematical arguments and by conducting extensive experiments over the benchmark LFW [17] and Face94 [18] databases. The rest of the paper is organized as follows: Section 2 describes the proposed system, Section 3 furnishes the experimental design, Section 4 details the experimental results and finally, Section 5 proffers a discussion of the proposed work and outlines future work.

# Proposed system

The section describes the distinct sequence of steps involved in our approach along with the necessary background information of the applied techniques. The proposed approach is depicted in Fig.1.

**Input**

**Face Image**

**KeyͲpoint Detection**

**Feature**

**Extraction**

**Descriptor**

**Matching with FLANN**

**Announce**

**Match/Mis match**

Fig.1.Framework for the Applied Methodology

In the proposed approach, we consider the query image (input face) and the gallery image from the database (LFW or Face94) and perform key-point detection and feature description (according to the two novel variants and classical SIFT and classical SURF). Finally, we match the descriptors of the query and gallery images using the FLANN [16] algorithm (section 2.3) in order to conclusively determine whether they belong to the same face i.e. match/mismatch status.

In our implementation, we make several prudent modifications to the existing SIFT and SURF techniques to

proffer the two novel techniques: SURF-Detector-SIFT-Descriptor (section 2.3) and SIFT-Detector-SURF-

Descriptor (section 2.4), which are compared with SIFT-Detector-SIFT-Descriptor (classical SIFT, section 2.1) and SURF-Detector-SURF-Descriptor (classical SURF, section 2.2) to demonstrate the proficiency of our proposed method.

* 1. *SIFT Algorithm (SIFT-Detector and SIFT-Descriptor)*

The SIFT algorithm [2][10][11][12]consists of four crucial steps: Scale-space Extrema Detection, Key point Localization, Orientation Assignment and Key-point Description.

* + 1. *Scale-Space Extrema Detection*

The first computational stage [12] involves the scanning of all the scale and image locations, by implementing the Difference-of-Gaussian (DoG) function. DoG distinctly identifies the potential interest points [30] (which are invariant with respect to scale and orientation).

ܩሺݔǡ ݕǡ ߪሻ ൌ భ ିሺ௫మା௬మሻȀʹߪଶ

೐

మഏ഑మ

(1)

DoG is carried out by convolving the image with the Gaussian filters at different scales. Subsequently, the computation of the difference of the successive Gaussian-blurred images is performed, followed by the extraction of key-points as maxima/minima of the difference of Gaussians at multiple scales. A DoG image is represented as follows:

ሻݕ ǡݔሺܫ כ ሻߪ ǡݕ ǡݔሺܩ ൌ ሻߪ ǡݕ ǡݔሺܮ

where ܮሺݔǡ ݕǡ ߪሻ denotes the convolution of the original image:

(2)

ܦሺݔǡ ݕǡ ߪሻ ൌ ൫ܩሺݔǡ ݕǡ ߪ݇ሻ െ ܩሺݔǡ ݕǡ ߪሻ൯כ ܫሺݔǡ ݕሻ ൌ ܮሺݔǡ ݕǡ ߪ݇ሻ െ ܮሺݔǡ ݕǡ ߪሻ (3)

This computation of DoG is essentially the process of scale-space extrema detection with the SIFT algorithm involves the convolution of the image with Gaussian-blurs at different scales and subsequently, the difference-of- Gaussian images are acquired from the adjacent Gaussian-blurred images based on a per octave basis. The Local Extrema is detected by comparing a pixel marked x against its 26 neighbors (represented in green) in a 3\*3\*3

neighborhood that spans adjacent DoG images (discretized scale-space volume). 26 was chosen as the threshold pixel as it showed promising results in [11][12].

2*po*.1*i*.*n*2*t KLoecyalization*

The successive stage in the application of the SIFT algorithm is *key-point localization* [12][30]. In order to compute the location and scale information, a model is fitted for each key-point candidate location. These key-points are chosen on the basis of the measurement of their stability and the interpolation process is carried out by utilizing the quadratic Taylor expansion of the Difference-of-Gaussian scale-space function [30]. The Taylor’s expansion is represented as follows [12]:

ሺ ሻ ఋ஽೅ ଵ ் ఋమ஽

ܦ ݔ ൌ ܦ ൅ ݔ ൅ ݔ

ఋ௫ ଶ

ఋ௫మ ݔ (4)

where x= (x, y) denotes the offset from this point. Once the key point is localized and refined, the low contrast points are rejected as depicted in Equations (5) and (6).

ෝݔ ൌ െ ఋమ஽షభ ఋ஽

(5)

ఋ௫మ ఋ௫

ܦሺݔොሻ ൌ ܦ ൅ ଵ ఋ஽೅ ݔො (6)

ଶ ఋ௫

*a*2*t*.1*io*.3*n OAsrsieignnt ment*

The orientation(s) are assigned to every key-point location on the basis of the local image gradient directions and subsequent operations are carried out on the image data [12][30] (after it has undergone transformation, relative to the assigned orientation scale and location for every such feature), thereby rendering the transformations invariant with respect to the aforementioned parameters. Initially, the Gaussian-smoothed image ܮ ሺݔǡ ݕǡ ߪሻ at key-point's

scale ı is considered in order to ensure that all the computations are conducted in a scale-invariant fashion. Consider

that for a given sample image ܮ ሺݔǡ ݕሻ, at a scale ı, the gradient magnitude is represented by ݉ ሺݔǡ ݕሻ, and the orientation is denoted by ߠ ሺݔǡ ݕሻ (all of which are pre-computed by utilizing the pixel differences). Then ߠ ሺݔǡ ݕሻ and ݉ ሺݔǡ ݕሻcan be represented in the following manner [12][30]:

݉ሺݔǡ ݕሻ ൌ ඥሺܮሺݔ ൅ ͳǡ ݕሻ െ ܮሺݔ െ ͳǡ ݕሻሻଶ ൅ ሺܮሺݔǡ ݕ ൅ ͳሻ െ ܮሺݔǡ ݕ െ ͳሻሻଶ (7)

ߠሺݔǡ ݕሻ ൌ ିଵሺሺܮሺݔǡ ݕ ൅ ͳሻ െ ܮሺݔǡ ݕ െ ͳሻሻΤሺܮሺݔǡ ݕ ൅ ͳሻ െ ܮሺݔǡ ݕ െ ͳሻሻሻ (8)

*p*2*o*.1*i*.*n*4*t KDeeyscription*

This phase consists of computing the key-point descriptors, which is carried out by measuring the local image gradients at the selected scale in the proximity of the region around each key-point [12][30]. These local image gradients are transformed into a representation that permits significant levels of local shape distortion and illumination changes [30]. This process is depicted in Fig.2.

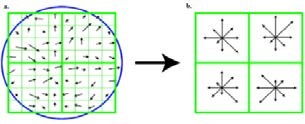


Fig.2.Sift Feature Descriptor

SIFT also performs the computation of vectors that can sufficiently characterize the local image appearance around the location of particular features. A typical instance of a descriptor can be a window of either colour or grayscale values around the detected point. The SIFT descriptor utilizes image gradients, instead of intensity values because the image derivatives do not vary if an addition of a constant value is performed on each pixel’s intensity. Essentially, SIFT considers the direction of the gradient, instead of their raw magnitude because the gradient directions are highly invariant to changes with respect to brightness and contrast. SIFT operates by considering the local image gradient directions and computes their histogram by creating 4 × 4 histogram grid around a selected feature point [30] (each such histogram consists of eight bins for gradient directions, yielding a 4 × 4 × 8 = 128dimensional descriptor. Hence, a feature f contains a 2D location (fx,fy) along with a descriptor vector fd[12][30].

* 1. *SURF Descriptor (SURF-Detector and SURF-Descriptor)*

SURF (Speed Up Robust Features) algorithm [3][6][26][28] is a distinctly proficient algorithm based on the multi-scale space theory that is robust to variations in scale, illumination, rotation and so on. The SURF algorithm consists of the following four critical stages [3][26][28]: (1) Integral image generation, (2) Fast-Hessian detector,

(3) Descriptor orientation assignment and(4) Descriptor generation.

* + 1. *Integral Image Generation*

The Integral Image (as illustrated in the Fig 3, where S=A-B-C+D) is an intermediate representation for the image, that plays a pivotal role in improving speed of the SURF algorithm. It contains the sum of the gray scale pixel values of the image (S).

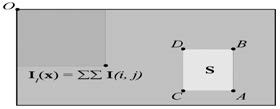


Fig. 3: Evaluation of Integral Image

The integral image, as depicted in Equation (9) is employed by the succeeding stages of algorithm to improve the computation speed [3][26][28].

Iஊሺݔǡ ݕሻ ൌ σ௫

௜ୀ଴

௬

௝ୀ଴

σ

Iሺ݅ǡ ݆ሻ

(9)

When the integral image is used, it is necessary to read only four pixel values so as to compute the surface integral of any size from the given image (the choice of four pixels is demonstrated to be optimal [3][26][28]).

*2.2.2 Keypoint Detection (Fast Hessian Detector)*

SURF employs the determinants of the Hessian matrices to locate the interest points in the given image. The determinants of the Hessian matrix for a 2D function are depicted in Equation (10) [3][26].

࣢ሺݔǡ ݕሻ

డమ௙ డ௫మ

ൌ  ൮ డమ௙

൲

డೣడ೤

డమ௙ డೣడ೤ డమ௙ డ௬మ

(10)

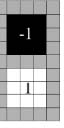
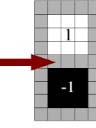
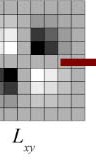
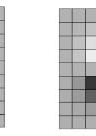
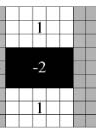
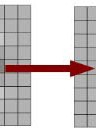
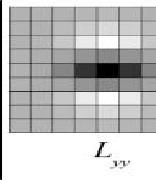


Fig. 4. LoG Approximations.

The Fast-Hessian detector revises Equation (10) in two prominent ways [3][26][28].First, the second order partial derivatives are replaced by the convolutions of the image with approximated Gaussian kernel second order derivatives [26][28]. The Laplacian of Gaussian (LoG) approximations are performed by using the box filters with the following coefficients [26]: 1, í1, 2, í22(as depicted in Fig.4). The value of the coefficient is set to 0.9 in Equation (11) to compensate for the aforementioned approximation [3][26][28]. For the second revision, the Gaussian kernels are parameterized according to their position and size in the image. Then to tender scale invariance the crucial points on multiple image scale levels are found [3][26]. Since, scaling the entire image bears high

computation expense, SURF only

scales the Gaussian kernels (only the box filters that

approximate them)

[3[26][28]. The scaling component is represented by the parameter ‘s’ in Equation (11). ‘s’ forms the Scale Space (3D space of determinant results)[3][26]. The Scale values vary and are quantized with respect to the octaves and intervals [26] and these aforementioned revisions render Equation (11).

ଶ

࣢ሺݔҧሻ ൌ ܦ௫௫ሺݔҧሻܦ௬௬ሺݔҧሻ െ ቀͲǤͻ ܦ௫௬ሺݔҧሻቁ (11)

whereݔҧ ൌ ሺݔǡ ݕǡ ݏሻ

Svab et al. [26] illustrated that, as the size of the box filter increases, the image-sampling rate nearly doubles between octaves and accordingly, the determinants in all the predefined scales and intervals are calculated and filtered with respect to the positive threshold value. Subsequently, the non-local maxima suppression is carried out for the 26 closest neighbors in the determinant scale space and the determinant local maxima represents the position of the interest point [3][26] and then this position is refined via interpolation to the sub-pixel precision [3][26].

The second order Taylor polynomial approximation of the Hessian within the scale space centered around the examined local maxima is denoted by Equation (12), where x = (x, y, s) [26].

࣢ሺݔሻ ൌ ࣢ ൅ డு೅

ଵ ் డమ࣢ ݔ

(12)

ݔ ൅ ݔ

డ௫ ଶ

డ௫మ

The position correction is then attained through the computation of the zero value of the directive of the polynomial. The position correction is represented by Equation (13), whereݔො ൌ ሺݔොǡ ݕොǡ ݏƸሻ [3][26][28]. The local maxima is then selected as an interest point if none of the absolute values of the position correction vector ݔො elements are larger than 0.5 [26].

ݔො ൌ െ డమ࣢షభ డ࣢

(13)

డ௫మ డ௫

* + 1. *Descriptor Orientation Assignment*

The assignment of a dominant direction to each of the interest points renders the Rotational Invariance characteristic [3][26]. If rotation invariance is not required, this step can be skipped and the orientation may be assigned as (0,0) [26]. Consequently, we consider a circle with a radius of 6s around the interest point [3][26][28] and then within the circle, the Haar wavelet responses along the x and y directions are calculated from the 4s-sized filters by using a sampling step of `s` [3][26]. The responses that are thus procured are weighted by the Gaussian

function using a value of ı = 2.5s, which is cent red on the interest point and these are summed by employing a sliding sector window of ʌ/3 with an approximate step of ʌ/18 [3][26[28]. The longest vector that is obtained as this window is selected to be the descriptor dominant direction [26].

* + 1. *Descriptor Generation*

As a final step, the SURF descriptor is calculated from the neighborhood of the square interest points with the edge size of 20s [3][26][28]. The selected neighborhood is then rotated to match the dominant direction of the given interest point [3][26]. Then, the square area is further divided into 16 equal sub-squares with edge sizes of 5s each

[26] and inside each of the sub-square areas, 25 Haar wavelet filter response pairs are calculated with the following: a filter size of 2s, sampling step s, one direction along the interest point dominant direction െ݀ௗ and one perpendicular in the positive sense െ݀௣ [3][26]. These responses are then weighted by the Gaussian function with ı

= 3.3s and then summed within the sub-square into a 4D vector ݒ ൌ ሾσ ݀ௗǡ σ ݀௣ ǡ σȁ݀ௗȁ ǡ σห݀௣หሿ [26]. These individual vectors are chained together by utilizing one vector for each of the 16 sub-squares to form a 64 dimensional SURF descriptor [3][26][28].

Subsequently, the descriptor elements are scaled again to ensure that the resultant descriptor is a unit vector [26]. The output of the algorithm for a given image produces a set of interest point locations in the image scale space along with a set of descriptors that describe the regions around the interest points [3][26].

* 1. *SURF-detector-SIFT-descriptor*

The first technique is formulated by conducting the key-point detection process using the SURF algorithm i.e. the stages of *Integral Image Generation* (Section 2.2.1), *Interest Point Detection using Fast-Hessian Detector* (Section 2.2.2) and *Descriptor Orientation Assignment* (Section 2.2.3) are conducted using SURF. Then, we perform *Feature Description* (Section 2.1.4) using the SIFT algorithm. This sequence of steps is elucidated in Section 4.1.

* 1. *SIFT-detector-SURF-descriptor*

The second technique is formulated by conducting the key-point detection process using the SIFT algorithm i.e. the stages of *Scale-space Extrema Detection* (Section 2.1.1), *Key point Localization* (Section 2.1.2) and *Orientation Assignment* (Section 2.1.3) are conducted using SIFT. Then, we perform *Feature Description* (Section 2.2.4) using the SURF algorithm. This sequence of steps is elucidated in Section 4.2. Subsequently, once the keypoint detection and feature description are carried out using one of the algorithms (Sections 2.1-2.4), the descriptors of the query and database images are matched by employing the FLANN [16] algorithm.

*L*2.*A*5*NFN(Fast Library for Approximate Nearest Neighbour Search)*

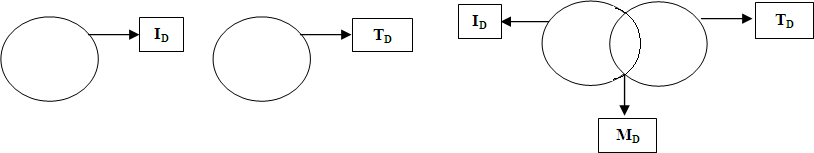


Fig.5 Descriptor Matching using FLANN

The FLANN [16] algorithm (as illustrated in Fig.5) performs descriptor matching by initially assuming that I and T represent the Query and Gallery image respectively. Further, Ik = {Ik1, Ik2, Ik3.....Ikn} and Tk = {Tk1, Tk2, Tk3. Tkn},

where Ik and Tk denote a set of key-points of Test Image (I) and Template image (T) respectively. Let us Consider ID= {ID1, ID2, ID3....IDmn}, where ID is a set of descriptors (feature vectors) of image I and TD= {TD1, TD2, TD3. TDmn},

where TD is a set of descriptors of image T. Then, MD = {MD11, MD12, MD13. MDmn} denotes the set of matched

descriptors.

# Experimental Design

* 1. *Face Databases*
     1. *LFW (Labeled Faces in the Wild) Database*



Fig.6 Sample Faces from the LFW Database

The LFPW (Labeled Faces in the Wild) [17] dataset, as illustrated in Fig.6, is a collection of face photographs that was specifically devised for the task of accurately understanding and investigating unconstrained face recognition. LFPW contains an excess of 13,000 face images that were acquisitioned from various sources from the World Wide Web. The associated person’s name has been utilized to label every individual face in the dataset. For about 1680 individuals that appear in the set, the dataset holds two or more distinct photos. There is however one perceivable constraint that all the faces in the dataset were detected by employing the Viola-Jones face detector [20]. In our experimentations, we have considered 200 images from the LFW dataset and while identifying True Negative (TN) and False Negative (FN), we have included 40 mismatched images (2 subjects) from the Face-94 dataset, which should not match on the LFW Dataset. The results yielded on this dataset are furnished in Table A.

*a*3*c*.1*e*.*9*2*4FDatabase*



Fig.7 Sample Faces from the LFW Database

The Face94 [18] dataset, as illustrated in Fig.7, consists of 3060 images from 153 subjects with a sequence of 20 images per person acquired using a fixed camera. The images are normalized to 48 x 48 pixels from 180 x 200 pixels and are in portrait format. There are two directories in the dataset, the female directory comprises of 20 subjects, the male directory has 113 subjects and the male staff directory has 20 subjects. The images are acquired against a plain green background, with minor variations in head turn, tilt, slant and the position of the face in the image. Furthermore, there is considerable expression variation but contains no head scale or image lighting variation. In our experimentations, we have considered 400 images from the Face-94 dataset comprising of 40 subjects/persons and while identifying TN and FN, we have included 40 mismatched images (2 subjects) from LFW Data set, which should not match on the Face-94 dataset. The results yielded on this dataset are furnished in Table B.

* 1. *Evaluation Metrics*

Precision (Positive Predictive Value) represents the fraction of the retrieved instances that are relevant and Recall (Sensitivity) denotes the fraction of relevant instances that were retrieved. Essentially, high precision value indicates that the algorithm returned remarkably more relevant results than irrelevant and high recall suggests that the algorithm returned majority of the relevant results and therefore both are decisively significant for an ideal algorithm [24].

Further, False Acceptance Rate (FAR) and False Rejection Rate (FRR) are two crucial metrics that have been widely employed for comparison of biometrics verification performances [23] because they are single index measures and easy to interpret. FAR is the measure of the likelihood that a given biometric system will incorrectly accept a given image and is generally defined as the ratio of the number of false acceptances to the number of identification attempts. FRR is the measure of the likelihood that a given biometric system will incorrectly reject a given image and is defined as the ratio of the number of false rejections to the number of identification attempts [22]. Hence for an efficient system, FRR and FRR should exhibit a decreasing trend.

Finally, Sensitivity (True Positive Rate) measures the proportion of actual positives that were correctly identified and is typically examined with Specificity (True Negative Rate), which measures the proportion of negatives that were correctly identified. Thus, an ideal predictor should be highly sensitive and highly specific. The procedural information concerning the computation of these metrics have been comprehensively detailed in [21].

# Experimental Results

The experimentations were conducted on the benchmark LFW andFace94databases to establish the proficiency of our proposed methods: *SURF-Detector-SIFT-Descriptor*and *SIFT-Detector-SURF-Descriptor* over the state-of-the-art *SIFT- detector-SIFT-descriptor* (classical SIFT) and *SURF-detector-SURF-descriptor* (classical SURF) algorithms.

* 1. *SURF-Detector-SIFT-Descriptor*

We formulate our first technique by employing the SURF algorithm to perform key-point detection, and utilize SIFT to carry out the feature description process. We find an image from the Face94 Data Set with the minHessian value on 10 (threshold value for selecting key-points as described in [6]). Given a Query image (Fig. 8.1), we initially select the face from it i.e. face detection (as illustrated in Fig 8.2) and subsequently we locate the key- points using SURF and then we perform feature description using SIFT (Fig 8.3). The same process is repeated for the database images. Finally, we perform descriptor matching of the query image against the database images using FLANN (Fig. 8.4) to determine the match/mismatch status. This combination yields superior accuracy and precision rate than the other variants (as depicted in Table A).

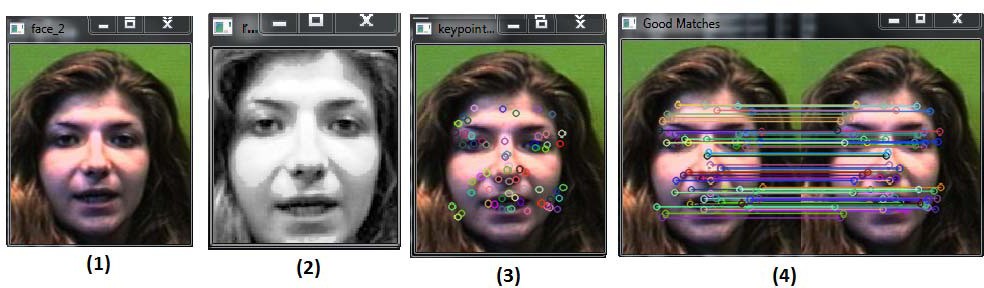


Fig. 8: (1) Original Image (2) Face selection (3) Key point detecting (4) Matching

* 1. *SIFT-Detector-SURF-Descriptor*

We formulate our second technique by employing the SIFT algorithm to perform key-point detection, and utilize SURF to carry out the feature description process. We find an image from the Face94 Data Set with the minHessian value on 20. Given a Query image (Fig. 9.1), we initially select the face from it i.e. face detection (as illustrated in Fig 9.2) and subsequently we locate the key-points using SIFT and then we perform feature description using SURF (Fig 9.3). The same process is repeated for the database images. Finally, we perform descriptor matching of the query image against the database images using FLANN (Fig. 9.4) to determine the match/mismatch status.

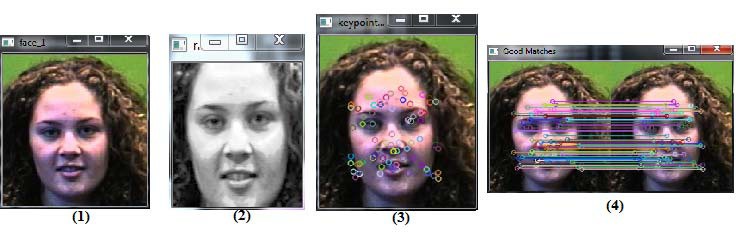


Fig. 9: (1) Original Image (2) Face Selection (3) Keypoint Detection (4) Matching

This technique was found to be relatively slow in finding the keypoints and hence in order to address this, we opted to increase the value of minHessian from 10 to 20. Furthermore, since LFW consists of images acquired under unconstrained scenarios, carrying out FR on this dataset is considerably arduous and hence the overall accuracy suffers over this database (still the proposed methods fared better than the classical algorithms).

* 1. *SIFT-Detector-SIFT-Descriptor (Classical SIFT)*

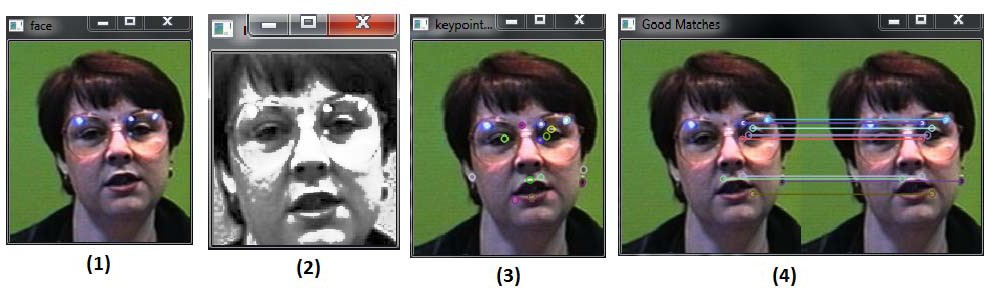


Fig. 10. (1) Original Image (2) Face Selection (3) Keypoint Detection (4) Matching

In this approach, we run the classical SIFT algorithm (SIFT for both key-point detection and feature description) in order to find an image in the Face-94 Dataset with the minHessian value on 10. Given a Query image (Fig. 10.1), we initially select the face from it i.e. face detection (as illustrated in Fig 10.2) and by using that we perform both key-point detection and feature description using SURF (Fig 10.3). The same process is repeated for the database images. Finally, we perform descriptor matching of the query image against the database images using FLANN (Fig. 10.4) to determine the match/mismatch status.

* 1. *SURF-Detector-SURF-Descriptor (Classical SURF)*

In this approach, we run the classical SURF algorithm (SURF for both key-point detection and feature description) in order to find an image in the LFW Dataset with the minHessian value on 10.



Fig. 11: (1) Original Image (2) Face Selection (3) Keypoint Detection (4) Matching

Given a Query image (Fig. 11.1), we initially select the face from it i.e. face detection (as illustrated in Fig 11.2) and by using that we perform both key-point detection and feature description using SIFT (Fig 11.3). The same process is repeated for the database images. Finally, we perform descriptor matching of the query image against the database images using FLANN (Fig. 11.4) to determine the match/mismatch status.

4*or*.5*mPanercfe Evaluation*

The following results were obtained on a machine with the following configuration: Intel Core i-7, 2.3 GHz processor and 8 GB RAM with windows 8 as the operating system.

Table A. Experiment Results - LFW Dataset (of 200 images + 40 Unknown Images)

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Detector | Descriptor | Accuracy | F1 Score | Sensitivity | Specificity | Precision | FAR | FRR |
|  |  | (%) | (%) | (%) | (%) |  |  |  |
| SURF | SIFT | 78.86 | 86.71 | 98.41 | 33.33 | 77.5 | 0.667 | 0.016 |
| SIFT | SURF | 76.67 | 85.1 | 98.35 | 31.07 | 75 | 0.690 | 0.016 |
| SURF | SURF | 75.48 | 84.93 | 93.93 | 25 | 77.5 | 0.750 | 0.061 |
| SIFT | SIFT | 71.1 | 81.64 | 93.55 | 21.43 | 72.5 | 0.786 | 0.065 |
| Table B. Experiment Results – Face 94 Dataset (of 400 images + 40 Unknown Images) | | | | | | | | |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Detector | Descriptor | Accuracy | F1 Score | Sensitivity | Specificity | Precision | FAR | FRR |
|  |  | (%) | (%) | (%) | (%) |  |  |  |
| SURF | SIFT | 96.67 | 98.18 | 98.74 | 81.81 | 97.5 | 0.182 | 0.013 |
| SIFT | SURF | 93.34 | 96.16 | 98.65 | 64.28 | 93.75 | 0.357 | 0.013 |
| SURF | SURF | 92.22 | 95 | 95 | 70 | 96.203 | 0.400 | 0.050 |
| SIFT | SIFT | 90 | 94.27 | 96.16 | 53.84 | 92.5 | 0.462 | 0.039 |

An ideal predictor [21] should demonstrate an increasing trend in sensitivity and specificity, and a decreasing trend in FAR and FAR; and as illustrated by Tables A and B; our algorithm adheres to these stipulations.

4.6 *Accuracy Comparison*

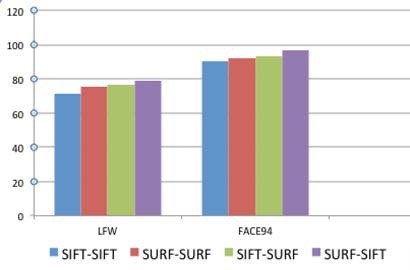


Fig.12 Accuracy Comparison over the Benchmark Databases

Over the LFW database, the proposed SURF-detector-SIFT-descriptor method outperformed SIFT-detector-SIFT-descriptor (classical SIFT) by 7.75% and SURF-detector-SURF-descriptor (classical SURF) by

3.38%. Our second proposed technique SIFT-detector-SURF-descriptor outperformed SIFT-detector-SIFT- descriptor (classical SIFT) by 5.57% and SURF-detector-SURF-descriptor (classical SURF) by 1.19%.

Over the Face94 database, the proposed SURF-detector-SIFT-descriptor method outperformed SIFT-detector-SIFT-descriptor (classical SIFT) by 6.67% and SURF-detector-SURF-descriptor (classical SURF) by 4.45%. Our second proposed technique SIFT-detector-SURF-descriptor outperformed SIFT-detector-SIFT- descriptor (classical SIFT) by 3.34% and SURF-detector-SURF-descriptor (classical SURF) by 1.12%. Cumulatively, over the two databases, the SURF-detector-SIFT-descriptor method yielded an accuracy of 87.765% and SIFT-detector-SURF-descriptor yielded accuracy of 85.005% and between the two proposed techniques, the SURF-detector-SIFT-descriptor technique demonstrated an accuracy improvement of 2.76% over the SIFT-detector- SURF-descriptor.

# Discussion and Future Work

Two novel performance oriented detection-description techniques, which are considerably proficient in carrying out FR tasks has been discussed. Our experimentations on two benchmark datasets: LFW and Face94 convey that the SURF–detector-SIFT-descriptor method outperformed the others methods as it was able to detect more number of features and was robust even under unconstrained scenarios (as demonstrated on the LFW dataset).The SIFT- detector-SURF-descriptor method fared better than SIFT-detector-SIFT-descriptor (classical SIFT) and SURF- detector-SURF-descriptor (classical SURF), but yielded slightly less accuracy then the SURF–detector-SIFT- descriptor method and suffers slightly in terms of speed when compared to classical SIFT and SURF. Hence for general FR tasks, the SURF–detector-SIFT-descriptor technique is a more viable choice.

Future work is currently being focused towards further enhancing the accuracy of the algorithm by implementing other detector/descriptor combinations and validating them on various benchmark databases to ascertain the various parameters over which the combinations are effective.

# References

1. Nabeel Younus Khan, Brendan McCane, and Geoff Wyvill, “SIFT and SURF Performance Evaluation against Various Image Deformations on Benchmark Dataset”, International Conference on Digital Image Computing: Techniques and Applications, pp.501-506, 2011.
2. Geng, Cong, and Xudong Jiang. "Face recognition using sift features." In *Image Processing (ICIP), 2009 16th IEEE International Conference on*, pp. 3313-3316. IEEE, 2009.
3. Bay, Herbert, Tinne Tuytelaars, and Luc Van Gool. "Surf: Speeded up robust features." In *Computer vision–ECCV 2006*, pp. 404-417. Springer Berlin Heidelberg, 2006.
4. Jafri, Rabia, and Hamid R. Arabnia. "A Survey of Face Recognition Techniques." *JIPS* 5, no. 2 (2009): 41-68.
5. Miksik, Ondrej, and Krystian Mikolajczyk. "Evaluation of local detectors and descriptors for fast feature matching." In *Pattern Recognition (ICPR), 2012 21st International Conference on*, pp. 2681-2684. IEEE, 2012.
6. Luo Juan, and Oubong Gwun, “A Comparison of SIFT, PCA-SIFT and SURF”, International Journal of Image Processing (IJIP), Vol. 3, Issue 4, pp. 143-152.
7. B. Heisele, P. Ho, J. Wu, and T. Poggio, “Face recognition: Component-based versus global approaches,” Comput. Vis. Image Understand, vol. 91, no. 1, pp. 6–12, 2003.
8. Philippe Dreuw,Pascal Steingrube,Harald Hanselmann and Hermann Ney.”SURF-Face: Face Recognition under Viewpoint Consistency Constraints”T. Lindeberg, “Scale-space for discrete signals,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 12, pp. 234–254, 1990.
9. Mikolajczyk, K., Schmid, C.: A performance evaluation of local descriptors. PAMI 27 (2005) 1615–1630.
10. Jurie, F., Schmid, C.: Scale-invariant shape features for recognition of object categories. In: CVPR. Volume II. (2004) 90 – 96
11. Lowe, David G. "Object recognition from local scale-invariant features." In *Computer vision, 1999. The proceedings of the seventh IEEE international conference on*, vol. 2, pp. 1150-1157. Ieee, 1999.
12. D.G. Lowe. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 60(2):91–110, 2004.
13. Ramanathan, Narayanan, Rama Chellappa, and A. K. Roy-Chowdhury. "Facial similarity across age, disguise, illumination and pose." In Image Processing, 2004. ICIP'04. 2004 International Conference on, vol. 3, pp. 1999-2002. IEEE, 2004.
14. Bhumika G. Bhatt, Zankhana H. Shah “Face Feature Extraction Techniques: A Survey”, National Conference on Recent Trends in Engineering & Technology, 13-14 May 2011
15. Best-Rowden, Lacey, et al. "Unconstrained face recognition: Establishing baseline human performance via crowdsourcing." *Biometrics (IJCB), 2014 IEEE International Joint Conference on*. IEEE, 2014.
16. M. Muja and D. G. Lowe. Fast approximate nearest neighbors with automatic algorithmic configuration. In Proc. VISAPP, 2009
17. LFW Home “Labeled Faces in the Wild Face Database”[:http://vis-www.cs.u](http://vis-www.cs.umass.edu/lfw/(accessed)m[ass.edu/lfw/(accessed](http://vis-www.cs.umass.edu/lfw/(accessed) May 17, 2015)
18. Libor Spacek's Facial Image Database, “Face94 Databas[e”:http://cswww.essex.ac.uk/](http://cswww.essex.ac.uk/mv/allfaces/faces94.html)mv[/allfaces/faces94.ht](http://cswww.essex.ac.uk/mv/allfaces/faces94.html)m[l](http://cswww.essex.ac.uk/mv/allfaces/faces94.html) (accessed May 17, 2015)
19. Bay, Herbert, Andreas Ess, Tinne Tuytelaars, and Luc Van Gool. "Speeded-up robust features (SURF)." *Computer vision and image understanding* 110, no. 3 (2008): 346-359.
20. Jones, Michael, and Paul Viola. "Fast multi-view face detection." *Mitsubishi Electric Research Lab TR-20003-96* 3 (2003): 14.
21. Wikipedia article, “Confusion Matrix”: <http://en.wikipedia.org/wiki/Confusion_matrix>(accessed May 18, 2015)
22. Biometric Solution Provider article “False Acceptance Rate (FAR) & False Recognition Rate (FRR)”: <http://www.bayometric.com/blog/false-acceptance-rate-far-false-recognition-rate-frr>(accessed May 18, 2015)
23. K. Toh, L. Kim, S. Lee Biometric scores fusion based on total error rate minimization, Pattern Recognition, 41 (5) (2008), pp. 1066–1082
24. Wikipedia article, “Precision and Recall”: <http://en.wikipedia.org/wiki/Precision_and_recall>(accessed May 18, 2015)
25. N. Poh, S. Bengio, "Database protocol and tools for evaluating score-level fusion algorithms in biometric authentication", Pattern Recognition, 39 (2) (2006), pp. 223–233
26. Svab, Jan, Tomas Krajnik, Jan Faigl, and Libor Preucil. "FPGA based speeded up robust features." In *Technologies for Practical Robot Applications, 2009. TePRA 2009. IEEE International Conference on*, pp. 35-41. IEEE, 2009.
27. B. Heisele, P. Ho, J. Wu, and T. Poggio, "Face recognition: component-based versus Global Approaches ", Computer Vision and Image Understanding, vol 91, nos 1-2, pp 6- 21, 2003.
28. Bay, Herbert, Andreas Ess, Tinne Tuytelaars, and Luc Van Gool. "Speeded-up robust features (SURF)." *Computer vision and image understanding* 110, no. 3 (2008): 346-359.
29. Luo, Jing, Shuze Geng, Zhaoxia Xiao, and Chunbo Xiu. "A review of recent advances in 3D face recognition." In *Sixth International Conference on Graphic and Image Processing (ICGIP 2014)*, pp. 944303-944303. International Society for Optics and Photonics, 2015.
30. S.Jain, M. Chouhan, “Content Based Image Retrieval By Clustering using (SIFT) Algorithm”, In the Proceedings of International Journal of Computer Science and Information Security (IJCSIS) Volume (1): Issue (1), June 2013.
31. Morel, Jean-Michel, and Guoshen Yu. "ASIFT: A new framework for fully affine invariant image comparison." *SIAM Journal on Imaging Sciences* 2, no. 2 (2009): 438-469.