

A Comparative Study of Sift and Surf Approaches

M. Chitra Devi

Assistant Professor, Dept. of Computer Science, M.K.U
Evening College, Dindigul

G. Daniel Kirubakaran

Assistant, M.K.U Evening College,
Dindigul

Abstract- Biometrics is the method for uniquely recognizing humans based upon one or more intrinsic physical or behavioral traits. Biometric image registration, required steps are: Feature detection, Feature matching, derivation of transformation function based on corresponding features in images and reconstruction of images based on derived transformation function. Accuracy of registered image depends upon accurate feature detection and matching. This paper presents two different methods such as Scale Invariant Feature Transform (SIFT) and Speed Up Robust Features (SURF). It also presents a way to extract distinctive invariant features from images that can be used to perform consistent matching between different views of an object.

Keywords: *Biometrics, Feature extraction, SIFT, SURF.*

I. INTRODUCTION

Biometric is the powerful tool for reliable automated person's identification and there exist several established biometrics-based identification techniques including fingerprint, geometry methods, speaker identification, and face recognition and iris identification[1]. These techniques have been broadly applied to access system, security system and other occasion needing identity witnesses.

Face Recognition[2] is one of the few biometric methods that possess the merits of both high accuracy and low intrusiveness. It has the accuracy of a physiological approach without being intrusive. For, this reason, face recognition has drawn the attention of researchers in fields from security, psychology and image processing. To computer vision. While network security and access control are it most widely discussed applications, face recognition has also proven useful in other multimedia information processing areas.

Lowe (2004) presented SIFT for extracting distinctive invariant features from images that can be invariant to image scale and rotation. Then it was widely used in image mosaic, recognition, retrieval and etc [3]. Bay and Tuytelaars (2006) speeded up robust features and used integral images for image convolutions and Fast-Hessian detector. Their experiments turned out that it was faster and it works well [13].

The features of the face are called biometric identifiers. The biometric identifiers are not easily forget, misdirected, or shared. For this reason, Face recognition has drawn the consideration of the researchers in fields security, psychology, and image processing.

II. REVIEW OF LITERATURE

SIFT is firstly developed by D. Lowe in 1999 [3], the main idea of SIFT is to extract features from images to match the reliable features between different parts of the same object. The extracted features are invariant to scale and orientation, and are very distinctive of the image.

In 2008, F. Alhwarin et al. [6], proposed an improvement on the original SIFT algorithm by producing more reliable feature matching for the object recognition purposes. This is done by splitting the features extracted from both the test and the model object image into several sub groups before they are matched.

They also proposed in [7] a new method for fast SIFT feature matching and the experimental results show that the feature matching can be speeded up by 1250 times with respect to exhaustive search without lose of accuracy.

In 2013, Tong Liu et al. [9] proposed a face recognition system based on SIFT feature and its distribution on feature space. The proposed method gave a higher face recognition rate than other methods included matching and total feature based methods in three famous databases.

Scale Invariant Feature Transform (SIFT) proposed by D. Lowe [4] became popular in face recognition. One interesting thing about the SIFT is its capability to capture the main gray level features of an objects view by means of local patterns extracted from a scale-space decomposition of an image [5].

In 2011, Jacob Toft Pedersen, produced a technical report on Feature detection and implementing the Speeded-Up Robust Features(SURF) algorithm.

Herbert Bay, discussed the experimental results on a standard evaluation set, as well as on imagery obtained in the context of a real-life object recognition application using Speeded-Up Robust Features Algorithm.

III. OVERVIEW OF METHODS

A. Scale Invariant Feature Transform

In 2004 David Lowe presented a method to extract distinctive invariant features from images [11]. He named

them Scale Invariant Feature Transform (SIFT). This particular type of features are invariant to image scale and rotation, and are shown to provide robust matching across a substantial range of affine distortion, change in 3D viewpoint, addition of noise, and change in illumination. They are well localized in both the spatial and frequency domains, reducing the probability of disruption by occlusion, clutter, or noise. Large numbers of features can be extracted from typical images with efficient algorithms. A typical image of size 500 x 500 pixels will give rise to about 2000 stable features (although this number depends on both image content and choices of various parameters). In addition, the features are highly distinctive, which allows a single feature to be correctly matched with high probability against a large database of features, providing a basis for object and scene recognition. The cost of extracting these features is minimized by taking a cascade filtering approach, in which the more expensive operations are applied only at locations that pass an initial test.

Following are the major stages of computation used to generate the set of image features:

STEP1. Choose N candidates" based on the Euclidean distance. We define $X = \{x_{11}; \dots; x_{ij}; \dots\}$ as a test sample, $Y = \{y_{11}; \dots; y_{ij}; \dots\}$ as a training sample, where x_{ij} and y_{ij} are the pixels at position (i,j) of the test and training samples respectively. Let $B = \{Y_1; Y_2; \dots\}$ denote the training set. Let $D(X; Y_k)$ be the Euclidean distance between a test sample X and a training sample Y_k .

$$D(X, Y_k) = \sqrt{\sum_{i=1, j=1}^{I=H, j=W} (x_{ij} - y_{ij})^2}$$

Where H and W are the height and width of a test sample respectively. The training samples should have the same size as the test sample. We sort the whole training samples in the ascending order of Euclidean distance, and then choose N training samples with smaller distance as "candidates" set C.

$$C = \{(Y_1:::Y_N) | D(X; Y_1) \leq D(X; Y_2) \leq \dots \leq D(X; Y_N) \leq D(X; Y_{N+1})\}$$

Where $N \leq n$, n is the number of training samples in training set B.

STEP2. Choose P new candidates" based on SIFT features. In this step, we choose P new "candidates" from C based on the number of well matched pairs of SIFT features. First of all, we define the criterion of well matched pair of SIFT features. We build a KD-tree [42] using the descriptors of SIFT features in a training sample. And then, for each descriptor a in the test sample, we employ Best-Bin-First search algorithm to find out k nearest nodes $b_1; b_2; \dots; b_k$ in the KD-tree (usually $k=2$), which are sorted in descending order. Let d_1, d_2 respectively be the distances between a and b_1 , a and b_2 . We then calculate the ratio of $d_1; d_2$:

$$ratio = d_1/d_2$$

If $ratio < Threshold$ (defined manually), we define a and b_1 are a well matched pair of SIFT features. Fig.3 shows the effect of Threshold on the recognition accuracy. When Threshold is below a certain value, the recognition accuracy increases rapidly and reaches the highest while Threshold is 0.5. Thus, we fix it as 0.5 in our method.

In this step, we count the number of well matched pairs of SIFT features between the test sample and each "candidate" in C, and sort the "candidate" samples in descending order. Then we choose P samples from C with the greater number of well-matched pairs as the new "candidates" set $C' = \{C'1; C'2; \dots; C'P\}$.

STEP3. Calculate the average similarity between classes. Let M_i be a class with the same class label as new "candidates" $c_i \in C'$, e_{ij} the angle of SIFT feature descriptors between a test sample and the j-th training sample in M_i , and \bar{e}_i be the average angle between the test sample and each class M_i .

$$e_{ij} = \cos^{-1} \left(\frac{f_{test} \cdot f_{ij}}{\|f_{test}\| \cdot \|f_{ij}\|} \right)$$

$$\bar{e}_i = \frac{\sum_{j=1}^L e_{ij}}{L}$$

Where f_{test} and f_{ij} are the descriptors of the test sample and the j-th training sample in M_i respectively, and L is the number of training samples in M_i . In our method, we describe the similarity of a test sample and the class M_i by using e_i and choose the class with the minimum angle as the recognition result. To get high recognition accuracy, we adjust some parameters involved in our method by using a small validation set. The parameters mainly include the numbers N,P of "candidates" chosen in the first two steps and the Threshold used to count the number of well matched pair of SIFT.

Matlab is used to implement the SIFT matching algorithms. In an attempt to assess the significant number of SIFT features required for reliable matching of face and skull images, several experiments were performed using only a subset of the extracted SIFT features in the matching process.

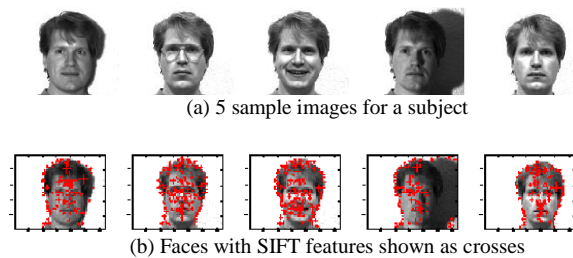


Fig. 1: Yale Face Database

Clearly, the accuracy increases rapidly with increasing the number of SIFT features used and then starts to saturate. This can considerably decrease the run time for SIFT matching process, as the number of matching operations is $O(n^2)$ where n is the number of features to be matched.

B. .Speeded-Up Robust Feature

Scale-Invariant Feature Transform, SIFT is a successful approach to feature detection introduced by Lowe [3]. The SURF-algorithm [13] is based on the same principles and steps, but it utilizes a different scheme and it should provide better results, faster.

SURF uses a hessian based blob detector to find interest points. The determinant of a hessian matrix expresses the extent of the response and is an expression of the local change around the area.

$$H(x, \sigma) = \begin{bmatrix} L_{xx}(x, \sigma) & L_{xy}(x, \sigma) \\ L_{xy}(x, \sigma) & L_{yy}(x, \sigma) \end{bmatrix} \quad (1)$$

Where
$$L_{xx}(x, \sigma) = I(x) * \frac{\partial^2}{\partial x^2} g(\sigma) \quad (2)$$

$$L_{xy}(x, \sigma) = I(x) * \frac{\partial^2}{\partial x \partial y} g(\sigma) \quad (3)$$

$L_{xx}(x, \sigma)$ in equation 2 is the convolution of the image with the second derivative of the Gaussian. The heart of the SURF detection is non-maximal-suppression of the determinants of the hessian matrices. The convolutions is very costly to calculate and it is approximated and speeded-up with the use of integral images and approximated kernels.

An Integral image² $I(x)$ is an image where each point $x = (x, y)^T$ stores the sum of all pixels in a rectangular area between origo and x (See equation 4).

$$I(x) = \sum_{i=0}^{i \leq x} \sum_{j=0}^{j \leq y} I(x,y) \quad (4)$$

The use of integral images enables calculating the response in a rectangular area with arbitrary size using 4 look-ups. The second order Gaussian kernels $\frac{\partial^2}{\partial y^2} g(\sigma)$ used for the hessian matrix must be discretized and cropped before we can apply them, a 9×9 kernel[10] is illustrated in Figure 2.

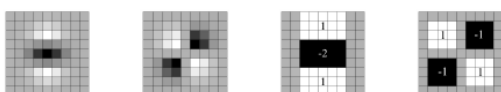


Fig. 2. Left to right: the (discretised and cropped) Gaussian second order partial derivative in y- (L_{yy}) and xy-direction (L_{xy}), respectively; our approximation for the second order Gaussian partial derivative y- (D_{yy}) and xy-direction (D_{xy}). The grey regions are equal to zero.

The SURF algorithm approximates these kernels with rectangular boxes, box filters. In the illustration grey areas corresponds to 0 in the kernel where as white are positive and black are negative. This way it is possible to calculate the approximated convolution effectively for arbitrarily sized kernel utilizing the integral image.

$$\text{Det}(H_{\text{approx}}) = D_{xx}D_{yy} - (wD_{xy})^2 \quad (5)$$

The approximated and discrete kernels are referred to as D_{yy} for $L_{yy}(x, \sigma)$ and D_{xy} for $L_{xy}(x, \sigma)$. The illustrated kernels corresponds to a σ of 1.2 and are the lowest scale that the SURF algorithm can handle. When using the approximated kernels to calculate the determinant of the Hessian matrix - we have to weight it with w in equation 5, this is to assure the energy conservation for the Gaussians. The w term is theoretically sensitive to scale but it can be kept constant at 0.9 [13].

To detect features across scale we have to examine several octaves and levels, where SIFT scales the image down for each octave and use progressively larger Gaussian kernels, the integral images allows the SURF algorithm to calculate the responses with arbitrary large kernels Figure 3.

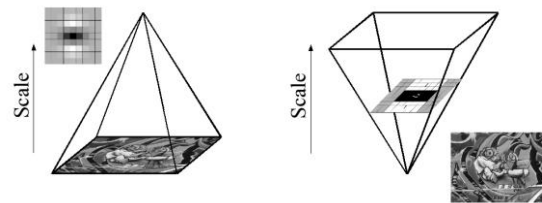


Fig. 3. Increasing the size of kernels, and keeping the lobes correctly scaled

This does pose two challenges, how to scale the approximated kernels and how this influences the possible values for σ . The kernels has to have an uneven size to have a central pixel and the rectangular areas, the lobes, has to have the same size Figure 4.

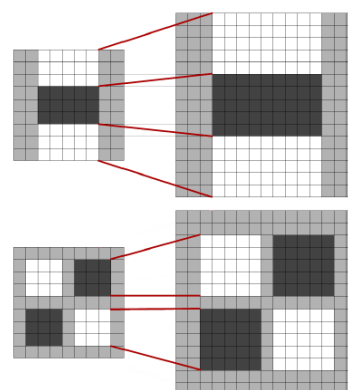
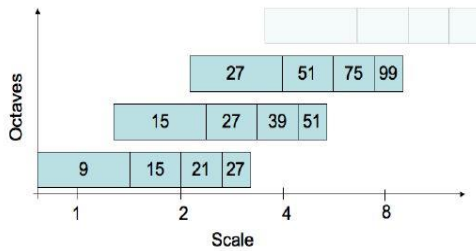


Figure 4. Filters D_{yy} (top) and D_{xy} (bottom) for two successive scale levels (9x9 and 15x15). The length of the dark lobe can only be increased by an even number of pixels in order to guarantee the presence of a central pixel(top).

The SURF paper [2] goes into detail with these considerations. The result is that division of scale space into levels and octaves becomes fixed as illustrated in Figure 7.



The filter size will be large and if the convolution were to be done with a regular Gaussian kernel this would be prohibitively expensive. The use of integral images not only makes this feasible - it also does it fast, and without the need to downscale the image. It should be noted that this approach with large box filters can preserve and be sensitive to high frequency noise. When finding a extrema at one of the higher octaves the area covered by the filter is rather large and this introduces a significant error for the position of the interest point. To remedy this the exact location of the interest point are interpolated by fitting a 3D quadratic in scale space [14]. An interest point is located in scale space by (x, y, s) where x, y are relative coordinates $(x, y \in [0; 1])$ and s is the scale.

IV. CONCLUSION

In this paper, an overview of Scale Invariant Feature Transform (SIFT) and Speed Up Robust Features (SURF) methods are discussed. Based on the referenced papers experimental results, it is found that the SIFT has detected more number of features compared to SURF but it is suffered with speed. The SURF is fast and has good performance as the same as SIFT. In future, the new method will be proposed to overcome these problems by combining SIFT and SURF methods.

REFERENCES

[1] M.Chitra Devi, M.Pushpa Rani, "A Survey of Face Recognition in Current Techniques" International Journal of Advanced Research in Computer Science and Software Engineering, Volume 4, Issue 10, October 2014.

[2] Shang-Hung Lin, "An Introduction to Face Recognition Technology", Informing Science Special Issue on Multimedia Informing Technologies-Part 2, Vol.3, No.1, 2000.

[3] D. Lowe. Object recognition from local scale-invariant features. In Int. Conf. on Computer Vision, pages 1150–1157, 1999.

[4] D. Lowe. Distinctive image features from scale-invariant keypoints. Int. Journal of computer Vision, 60(2):91–110, 2004.

[5] A. Mohamed, "Face Recognition using SIFT Features," CNS/Bi/EE report 186, 2006.

[6] F. Alhwarin, C.Wang, D. Ristic-Durrant and A. Gräser, "Improved SIFT-Features Matching for Object Recognition," In BCS International Academic Conference, 2008, pp. 178-190.

[7] F. Alhwarin, C. Wang, D. Ristic-Durrant and A. Gräser, "VF-SIFT:very fast SIFT feature matching," Pattern Recognition. Springer Berlin Heidelberg, 2010, pp. 222-231."

[8] Isra'a Abdul-Ameer Abdul-Jabbar, Jieqing Tan, and Zhengfeng Hou, "Adaptive PCA-SIFT Matching Approach for Face

Recognition Application" Proceedings of the International MultiConference of Engineers and Computer Scientists 2014 Vol I, IMECS 2014, March 12 - 14, 2014.

[9] T. Liu , S. H. Kim, S. K. Lim, "Selection of Distinctive SIFT Feature based on its Distribution on Feature Space and Local Classifier for Face Recognition," International Arab Journal of Information Technology, vol. 10, no. 1,2013 , pp.95-101.

[10] Jacob Toft Pedersen, "Study group SURF: Feature detection & description", SURF: FEATURE DETECTION & DESCRIPTION, Q4 2011.

[11] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool, "SURF-Speeded Up Robust Features", 2006.

[12] Scott Smith, "Speeded Up Robust Features",Advanced Image Processing, March 15, 2011.

[13] S. U. R. F. (SURF), "Herbet bay, andreas ess, tinne tuytelaars, luc van gool," Elsevier preprint, 2008.

[14] M. B. David Lowe, "Invariant features from interest point groups," BMVC, 2002.