

A Study of SURF features for "Face recognition"

Mr.Vivek D. Kamdi

Dr.Atul D. Raut

Ms.Vaishnavi Pardakhe

Mr.Vishal A. Jagtap

Abstract — Face recognition is a fascinating research topic in recent years. The Scale Invariant Feature Transform (SIFT) proposed by David G. Lowe has been used in face recognition and proved to perform well. Recently, a new detector and descriptor, named Speed-Up Robust Features (SURF) suggested by Herbert Bay, attracts people's attentions. SURF is a scale and in-plane rotation invariant detector and descriptor with comparable or even better performance with SIFT. Because each of SURF feature has only 64 dimensions in general and an indexing scheme is built by using the sign of the Laplacian, SURF is much faster than the 128-dimensional SIFT at the matching step. Thus based on the above advantages of SURF, we propose to exploit SURF features in face recognition in this paper. Face recognition is a big research area which takes more attention of many researchers in computer technology. The human face recognition from video sequences is a challenging task, because there are variations present in the Background of the images, facial expression and illumination

Keywords: SURF, SIFT, face recognition

I. INTRODUCTION

Face recognition has been an active area of research over the last two decades. It involves computer recognition of personal identity based on geometric or statistical features derived from face images. There are some traditional algorithms for face recognition such as EigenFace [9], Fisher Face [10], 2D-PCA [11] and Elastic Graph Matching [12]. The Scale Invariant Feature Transform (SIFT) proposed by David G. Lowe [7][8] has been widely used in object detection and recognition. There are also some works on the use of SIFT features in face recognition, such as SIFT_GRID proposed by M. Bicego[6] and SIFT_CLUSTER proposed by Jun Luo[5]. Different from traditional algorithms, the SIFT can extract local personal specific features and the above works show that SIFT can perform well in face recognition. Due to the high computation cost of SIFT in matching, some methods are proposed to speed it up. For example, kd-tree is used in the stage of searching k-nearest neighborhood, and PCA is proposed to reduce the dimensions of the SIFT features. However, these methods still can not's make SIFT satisfy the speed requirement of on-line applications. Recently, a new detector and descriptor, named Speed-Up Robust Features (SURF) suggested by Herbert Bay[13], attracts people's attentions. SURF is a scale and in-plane rotation invariant detector and descriptor with comparable or even better performance with SIFT. Its feature is also personal specific. Just like SIFT, in SURF, detectors are first employed to find the interest points in an image, and then the descriptors are used to extract the feature vectors at each

interest point. However, instead of difference of Gaussians (DoG) filter used in SIFT, SURF uses Hessian-matrix approximation operating on the integral image to locate the interest points, which reduces the computation time drastically. As for the descriptor, the first-order Haar wavelet responses in x and y directions are used in SURF to describe the intensity distribution within the neighborhood of an interest point, whereas the gradient is used by SIFT. In addition, only 64 dimensions are usually used in SURF to reduce the time cost for both feature computation and matching. Because each of SURF feature has only 64 dimensions in general and an indexing scheme is built by using the sign of the Laplacian, SURF is much faster than the 128-dimensional SIFT at the matching step. Furthermore, the method used to speed up the SIFT matching can also be applied to SURF. Thus based on the above advantages of SURF, we propose to exploit SURF features in face recognition in this paper. The performance of the proposed method will be evaluated by experimental comparisons with SIFT features. To our knowledge, the application of SURF to face recognition has not been systematically investigated yet.

The rest of this paper is organized as follows: In section 2, Speed-up robust features are introduced. Our proposed method is described in section 3. Section 4 gives the experimental results on the FERET database. Conclusions are drawn in section 5.

II. SPEED-UP ROBUST FEATURES

A speed-up robust feature (SURF) is a scale and in-plane rotation invariant feature. It contains interest point detector and descriptor. The detector locates the interest points in the image, and the descriptor describes the features of the interest points and constructs the feature vectors of the interest points.

Interest point detection

Different from SIFT using DoG to detect interest points, SURF [13] use the determinant of the approximate Hessian matrix as the base of the detector. To locate the interest point, we detect blob-like structures at locations where the determinant is at maximum. Integral images are used in Hessian matrix approximation, which reduce computation time drastically. Given a point $\mathbf{x} = (x, y)$ in an image I , the Hessian matrix $H(\mathbf{x}; \sigma)$ in \mathbf{x} at scale σ is defined as follows:

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

where $L_{xx}(x, \sigma)$, $L_{xy}(x, \sigma)$ and $L_{yy}(x, \sigma)$ are the convolutions of the Gaussian second order partial derivatives with the image I in point x respectively.

To reduce the computation time, a set of 9×9 box filters (Fig.1) is used as the approximations of a Gaussian with $\sigma = 1.2$ and represents the lowest scale (i.e. highest spatial resolution) for computing the blob response maps. We will denote them by $D_{xx}(x, \sigma)$, $D_{yy}(x, \sigma)$, and $D_{xy}(x, \sigma)$. The weights applied to the rectangular regions are kept simple for computational efficiency. This yields:

$$\det(H_{approx}) = D_{xx}D_{yy} - (\omega D_{xy})^2 \quad (2)$$

where ω is a weight for the energy conservation between the Gaussian kernels and the approximated Gaussian kernels, and

$$\omega = \frac{|L_{xy}(1.2)|_F |D_{yy}(9)|_F}{|L_{yy}(1.2)|_F |D_{xy}(9)|_F} = 0.912 \approx 0.9 \quad (3)$$

$|X|_F$ is the Frobenius norm.

For scale invariant, the SURF constructs a pyramid scale space, like the SIFT. Different from the SIFT to repeatedly smooth the image with a Gaussian and then sub-sample the image, the SURF directly changes the scale of box filters to implement the scale space due to the use of the box filter and integral image.

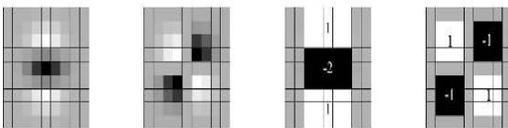


Fig. 1. The box filters of approximations of Gaussian second order partial derivative. The figure shows $L_{yy}(x, \sigma)$, $L_{xy}(x, \sigma)$, $D_{yy}(x, \sigma)$ and $D_{xy}(x, \sigma)$ from left to right.

Interest point description

In [13], the SURF used the sum of the Haar wavelet responses to describe the feature of an interest point. Fig.2 shows the Haar wavelet filters used to compute the responses at x and y directions. For the extraction of the descriptor, the first step consists of constructing a square region centered at the interest point and oriented along the orientation decided by the orientation selection method introduced in [13]. The region is split up equally into smaller

4×4 square sub-regions (as shown in Fig. 3). This preserves important spatial information. For each sub-region, we compute Haar wavelet responses at 5×5 equally spaced sample points. For simplicity, we call d_x the Haar wavelet response in horizontal direction and d_y the Haar wavelet response in vertical direction. To increase the robustness towards geometric deformations and localization errors, the responses d_x and d_y are first weighted with a Gaussian centered at the interest point.



Fig. 2. The Haar wavelet filters used to describe the interest points.

Then, the wavelet responses d_x and d_y are summed up over each sub-region and form a first set of entries in the feature vector. In order to bring in information about the polarity of the intensity changes, we also extract the sum of the absolute values of the responses, $|d_x|$ and $|d_y|$. Hence, each sub-region has a four-dimensional descriptor vector \mathbf{v} for its underlying intensity structure $\mathbf{v} = (\sum dx, \sum dy, \sum |dx|, \sum |dy|)$. Concatenating this for all 4×4 sub-regions, this results a descriptor vector of length 64. The wavelet responses are invariant to a bias in illumination (offset). Invariance to contrast (a scale factor) is achieved by turning the descriptor into a unit vector.

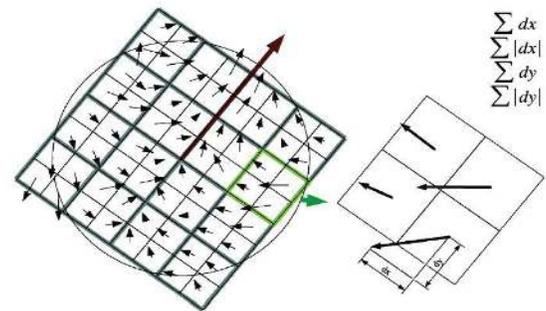


Fig. 3. The demonstration of descriptor building.

Fast index for matching

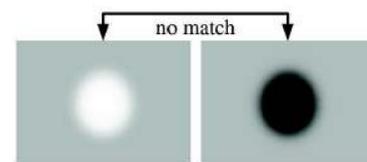


Fig. 4. The fast index for matching.

To speed up matching step, the sign of the Laplacian (i.e, the trace of the Hessian matrix) for the interest point is used. Only the point-pair with the same sign will be matched with the features. Fig.4 shows the example blobs of the sign.

III. SURF FOR FACE RECOGNITION

SURF feature extraction

Like using SIFT feature in face recognition, SURF features should be extracted from images through SURF detectors and descriptors[2]. Interest points are first extracted from each face image after pre- processing, such as normalization and histogram equalization (shown in Fig.5)[2]. This turns out to obtain about 30-100 interest points per image. The SURF feature vectors of the set of interest points are then computed to describe the image and these feature vectors are normalized to 1. These features are person-specific, since the number and the positions of points selected by SURF detector as well as the features around these points computed by SURF descriptor are different in each person's image[2].

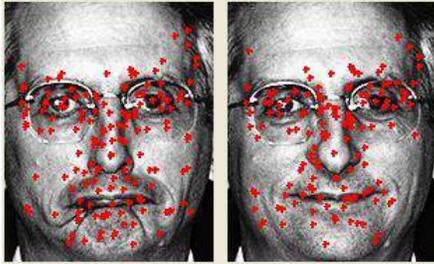


Fig. 5. Interest points in face image.

SURF feature matching

Point matching is commonly employed in face recognition with SIFT. In order to effectively match two face images, M. Bicego [6] and Jun Luo [5] suggested different sub-region matching strategies. But both methods have to compute the sub-region similarities and global similarities, with increase of computation cost in matching. In [6][5] they all used the point matching method mentioned in [7][8] as a part of their evaluation of matching. In this paper, based on point matching method suggest in [7][8], we introduce geometric constraints into point-matching based on SURF features to increase the matching speed and robustness. Because in face recognition, face images are usually upright and normalized, the matching points in two images must have the similar locations on the two faces. Thus for an interest point (x,y) in the probe image the search area for its mate is limited within a rectangular window centered at (x,y) of the gallery image. The point-pair with the minimum distance between descriptors will be considered as a candidate matching pair. To verify the validity of the candidate point-pair, the next minimal distance of point-pair, which contains the same point of the probe image, is then searched over the whole

area of the gallery image. If the ratio of these two distances is smaller than a pre-defined threshold, the point-pair with the minimum distance is confirmed as a matched pair. Since location information is introduced in search of the minimum-distance point-pair, and the ratio of the minimum distance and next to minimum distance measures the matching reliability of two interest points in some degree, the above method can avoid mismatching effectively. Finally, based on the result of the point-matching, we define a similarity measure in Eq.(4), which contains the number of matched points, the average value of the Euclidean distance, and the average distance ratio of all matched points, for face recognition.

$$Sim = \begin{cases} (Dis.Avg + Ratio.Avg)/2 & \text{if } N \geq 10 \\ (Dis.Avg + Ratio.Avg)/2 + 1 & \text{if } N < 10 \end{cases} \quad (4)$$

$$Dis.Avg = \frac{1}{N} \sum_n MinDis \quad (5)$$

$$Ratio.Avg = \frac{1}{N} \sum_n DisRatio \quad (6)$$

where N is the number of matched points of two face images, $MinDis$ is the Euclidean distance between two matched points, and $DisRatio$ is the distance ratio of the matched points. Here, $n = 1, \dots, N$.

When the number of the matched points of two images is smaller than the predefined threshold (10 in our experiment), it is thought that the matching result is not reliable even if the similarity measure Sim is small. Therefore, in Eq. (4), 1 is added to minish the similarity in this case. Fig.6 shows an example of the point matching result. The red lines indicate the corresponding matched interest points.

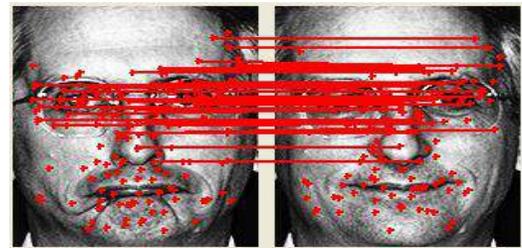


Fig. 6. Example of point matching result.

IV. EXPERIMENTAL RESULTS

Experiments are performed to evaluate the performance of the proposed algorithm, and the results are compared with that of using SIFT. The data sets we used in the experiments are FERET standard testing subsets of Facial Expression.

There are 1195 images in probe set and 1196 images in gallery set, one image for one subject. In our experiments, features of SURF-64, SURF-128, SIFT-128, SURFdbl-128 and SIFTdbl-128 are used respectively, where 64 and 128 indicate the dimensions of the feature vectors, and dbl means the size of the image doubled before feature extraction.

In Table 1, the recognition rates on all used feature types are given. It is obvious that SURF-64 has a similar recognition rate to that of SIFT-128 features. Feature vectors of 128 dimensional SURF (SURF-128) gives a slightly better result than SURF-64 and SIFT-128. For the feature sets of dbl, the experiments show that the recognition rate is higher than non-doubled feature sets except for 64 dimension features. SURF dbl-64 is slightly worse than SURF-64, which implies that it can not offer enough description of the image characteristics with 64 dimensions. Because a doubled image will generate more interest points than no doubled one, the 128 dimensions will supply more discrimination information than 64 dimensions in match. We also compare our results with that of the methods proposed in [5] and [6] in Table 1. It can be seen that our proposed method gives better or comparable performance.

Table 2 shows the matching time and performance of different SIFT or SURF features in face recognition. As we seen, the average matching speed of SURF-64 is the fastest one, and decrease almost one half compared to that of SIFT-128, and SURF-128 could save more than 30% computation time compared to that of SIFT-128. In [5] and [6], they not only used the point matching method proposed in [7] but also spent time to compute the similarities of sub-regions. In our method the point matching is faster than [7] because of limitation of the searching area. And no sub-region matching is needed. Obviously the speed of our method can be expected faster than that of [5] and [6] either using SURF or SIFT features.

Table 3 shows the effects of the ratio threshold on matching performance in cases of SURF and SIFT. The experimental results show that the SURF features are more robust to the change of the ratio threshold than SIFT features do.

Table 1. Recognition rate comparison among SURF and SIFT features.

Recognition rate of different features	
Feature used:	Recognition rate(%):
SURF-64	95.6
SURF dbl-64	95.2
SURF-128	96.0
SURF dbl-128	96.6
SIFT-128	95.9
SIFT dbl-128	96.6
SIFT_GRID[4]	94.0
SIFT_CLUSTER[3]	97.0

Table 2. Computation time cost with different feature types.

Average Matching Time of different features			
Feature Used:	Ratio Threshold:	Recognition Rate (%):	Average Matching Time (ms/image):
SURF-64	0.45	95.6	0.378
SURF-128	0.45	96.0	0.597
SIFT-128	0.45	95.9	0.911

Table 3. Recognition rate with different ratio thresholds.

Different ratio threshold		
Feature Used:	Ratio Threshold:	Recognition Rate(%):
SURF-64	0.5	95.6
SURF-128	0.5	96.1
SIFT-128	0.5	95.5
SURF-64	0.45	95.6
SURF-128	0.45	96.0
SIFT-128	0.45	95.9

CONCLUSION

This paper deals with using SURF features in face recognition and gives the detailed comparisons with SIFT features. Experimental results show that the SURF features perform only slightly better in recognition rate than SIFT, but there is an obvious improvement on matching speed. Therefore, SURF features are proven to be suitable for face recognition.

ACKNOWLEDGMENT

Working on this project on "Face Detection and Recognition for Automatic Attendance System" was a source of immense knowledge to us. We would like to express our sincere gratitude to the department of Computer Technology for the guidance and valuable support thought out the course of this project work. We acknowledge with a deep sense of gratitude, the encouragement and inspiration received from our faculty members and colleagues.

The work has been supported by National Natural Science Foundation of China (60772114) and the Key Project of The Ministry of Education of P. R. China (108012)

Portions of the research in this paper used the FERET database of facial images collected under the FERET program. SURF features provide obvious advantage of invariance of shifting, rotation and scaling, and partial invariance to illumination and affine transformation. Almost same advantages are also provided by SIFT features. In

future we can also think of including image preprocessing so for better illumination and scaling invariance.

REFERENCES

1. Geng Du*, Fei Su, Anni Cai, " Face recognition using SURF features," *Pattern Recognition and Computer Vision*, edited by Mingyue Ding, Bir Bhanu, Friedrich M. Wahl, Jonathan Roberts, *Proc. of SPIE Vol. 7496*, 749628, MIPPR 2009.
2. Bhaskar Anand, Prashant K Shah , " Face Recognition using SURF Features and SVM Classifier" *International Journal of Electronics Engineering Research. ISSN 0975-6450 Volume 8, Number 1 (2016) pp. 1-8*,
3. P. J. Phillips, H. Moon, S. A. Rizvi, P. J. Rauss, "The FERET evaluation methodology for face-recognition algorithms, " *IEEE transactions on pattern analysis and machine intelligence*, 22(10), (2000).
4. P. J. Phillips, H. Wechsler, J. Huang, P. Rauss, "The FERET database and evaluation procedure for face recognition algorithms, " *Image and Vision Computing J*, 16(5), 295-306 (1998).
5. J. Luo, Y. Ma, E. Takikawa, S. H. Lao, M. Kawade, B. L. Lu, "Person-Specific SIFT features for face recognition," *International Conference on Acoustic, Speech and Signal Processing (ICASSP2007)*, Hawaii, 2, 593-596 (2007).
6. M. Bicego, A. Lagorio, E. Grosso, M. Tistarelli, "On the use of SIFT features for face authentication," *Proc. of IEEE Int Workshop on Biometrics*, in association with CVPR, NY, (2006).
7. D. G. Lowe, "Object recognition from local scale-invariant features," *International Conference on Computer Vision*, Corfu, Greece, 1150-1157 (1999).
8. D. Lowe, "Distince image features from scale-invariant keypoints," *Int. Journal of Computer Vision*, 60(2), 91-110 (2004).
9. M. A. Turk, A. P. Pentland, "Face recognition using Eigenfaces," *In Proc. IEEE Conf. Computer Vision and Pattern Recognition*, (1991)
10. P. N. Belhumeur, J. P. Hespanha, D. J. Kriegman, "Eigenfaces vs. Fisherfaces: recognition using class specific linear projection," *IEEE transactions on pattern analysis and intelligence*, 19(7), (1997).
11. Jian Yang, D. Zhang, A. F. Frangi, Jingyu Yang, "Two-Dimensional PCA: a new approach to appearance-based face representation and recognition," *IEEE transactions on pattern analysis and machine intelligence*, 26(1), (2004).
12. M. Lades, J. C. Vorbruggen, J. Buhmann, J. Lange, C. von der Malsburg, R. P. Wurts, W. Konen, "Distortion invariant object recognition

in the dynamic link architecture," *IEEE transactions on computers*, 42(3), (1993).

13. H. Bay, A. Ess, T. Tuytelaars, L. Van Gool, "Speeded-up robust features(SURF)," *Comput. Vis. Image Underst.*, 110(3), 346-359 (2008).

AUTHOR'S PROFILE



Mr. Vivek D. Kamdi

Final Year, Information Technology.
 Jawaharlal Darda Institute of Engineering & Technology,
 Yavatmal
 .Email ID:vicky.kamdi20@gmail.com



Dr. Atul D. Raut

Professor & HOD Information Technology.
 Jawaharlal Darda Institute of Engineering & Technology,
 Yavatmal
 .Email ID:atuldraut@gmail.com



Ms. Vaishnavi S. Pardakhe

Therd Year, Information Technology.
 Jawaharlal Darda Institute of Engineering & Technology,
 Yavatmal
 Email ID:Pardakhevaish24@gmail.com



Mr. Vishal A. Jagtap

Second Year, Information Technology.
 Jawaharlal Darda Institute of Engineering & Technology,
 Yavatmal
 Email ID:Vishaljagtap@gmail.com