

Person Identification using Rotation Invariant SIFT Algorithm on IRIS Images

Shashwati Mishra, Mrutyunjaya Panda

Abstract— Scale Invariant Feature Transform (SIFT) is one of the popular feature extraction algorithm in computer vision which is used to detect and describe local features. One of the important characteristic of this algorithm which makes it popular is that SIFT is invariant to change in image scale, illumination, rotation etc. In this paper we have used SIFT algorithm for the identification of iris image and also proved that SIFT is invariant to rotation. Our experimental observations concludes that SIFT algorithm can be used for identifying a person from the iris images of the eye.

Index Terms— Biometric identifier, feature point, descriptor, SIFT, iris

1 INTRODUCTION

BIOMETRICS is the metrics related to human characteristics and is used in computer science as a form of identification and access control. Biometric identifiers can be classified into two types, one considers physiological characteristics and other behavioural characteristics. Physiological characteristics include fingerprint, palm veins, face recognition, DNA, palm print, hand geometry, iris recognition, retina etc. and basically related to the shape of the body. Behavioural characteristics include typing rhythm, gait and voice etc. which are related to the pattern of behaviour of a person. These identifiers are unique to individuals and more reliable in verifying identity than token and knowledge-based methods (password, personal identification number etc.)[2]. Here we have considered the physiological characteristic for identifying different persons by analyzing the iris images of eye using SIFT (Scale Invariant Feature Transform) algorithm.

As a sense organ that reacts to light and allows vision, the eye is a fused two piece unit. The cornea is the smaller frontal unit, transparent and more curved which is linked to the sclera the larger white unit. The corneal segment is typically about 8 mm in radius and the sclerotic chamber is typically about 12mm in radius. The ring like structure called limbus connects the cornea and sclera. The coloured circular structure concentrically surrounding the center of the eye, the pupil, is the iris.

After passing through the cornea, the pupil and then through the lens controlled by ciliary muscles, light falls on the light-sensitive cells of the retina, which is converted to

electrical signals. These electrical signals are carried to the brain by the optic nerves.

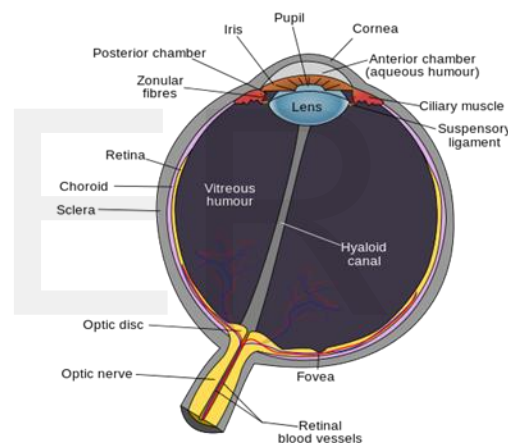


Fig. 1. Schematic diagram of human eye

The iris (irides or irises in plural) is a thin, circular structure in the eye. It controls the amount of light reaching the retina by adjusting the diameter and size of the pupil. Colour of the iris defines the colour of the eye like green, blue, brown, hazel (a combination of light brown, green and gold), grey, violet or even pink.

The process of recognizing a person by analyzing the random pattern of the iris is nothing but the iris recognition. In iris recognition two irides are matched for identification and verification of individual identity. This requires extraction of features and comparison of those features. J. Daugmann [4] explained the iris recognition algorithms along with the results of comparison among 9.1 million eye images. K. M. Gandhi et al. [5] presented an algorithm of iris feature extraction using a modified version of SIFT algorithm. For applying their proposed method they have compared feature points using java imaging API and SWT kit of eclipse. C.M. Patil et al. developed a method of identifying persons from iris images using lifting wavelet transform and similarity matching using

- Shashwati Mishra, Ph. D. research scholar, Department of Computer Science and Application, Utkal University, Vanivihar, Bhubaneswar, Odisha, India. E-mail: shashwati.mishra@gmail.com
- Mrutyunjaya Panda, Reader, Department of Computer Science and Application, Utkal University, Vanivihar, Bhubaneswar, Odisha, India. E-mail: mrutyunjayapanda@yahoo.in

Euclidean distance. SIFT algorithm is also applied on teeth images for feature extraction and matching by D. Rindhe et al. [2]. S. Malathi et al. [8] proposed a method of using SIFT algorithm for partial fingerprint matching. A. Jain et al. [9] used minutiae and texture features for fingerprint matching. In this paper we have used point feature extracted using SIFT algorithm for iris recognition.

2 FEATURE EXTRACTION USING SIFT

2.1 Feature

Feature represents a distinctive attribute or aspect of something. Features can be global or local. Global features describe the whole image and are generally not vary robust, because a change in part of the image may cause it to fail as it will affect the resulting features. Local features describe a patch within an image. Multiple local features are used to match an image and this is more robust as not all the features need to match for the comparison to be made. This makes them more robust to changes between the matched images [1], [3].

2.2 Feature Extraction

Feature extraction plays an important role in machine learning, pattern recognition and image processing [10]. Feature extraction process starts from an initial set of measured data and builds derived values or features which are informative and non-redundant, helps in subsequent learning and generalization steps, and leads to better human interpretations. In image processing feature extraction algorithms are used to detect and isolate desired features of a digital image or video stream. To extract low-level features several techniques like edge detection, corner detection, blob detection, ridge detection and scale invariant feature transform are used [1], [7].

2.3 Advantages of using SIFT

SIFT (Scale Invariant Feature Transform) detects and describes local features in an image. This algorithm was proposed by David Lowe in 1999. The Harris operator is not invariant to scale and correlation is not invariant to rotation. For better image matching, Lowe's goal was to develop an interest operator that is invariant to scale and rotation. Lowe was also aimed to create a descriptor that was robust to the variations corresponding to typical viewing conditions. In SIFT the image content is transformed into local feature coordinates that are invariant to translation, rotation, scale, and other imaging parameters. This algorithm is widely used in image search, object recognition, video tracking, gesture recognition etc. [1], [3].

3 PROPOSED METHODOLOGY

The proposed method for identifying the percentage of similarity between two iris images involves the following stages. Training image represents the input image that is compared with each image in the test image database to find a match.

3.1 Preprocessing Steps

Preprocessing steps involve any preprocessing activity that is done on the input images (training image and test image) be-

fore performing the actual processing to find the percentage of match between two images. This preprocessing activity involves rgb to gray conversion, segmentation, noise removal, enhancement of the image etc. In this proposed approach we have converted the input image from rgb to grayscale and removed the noise from the image using median filter [1].

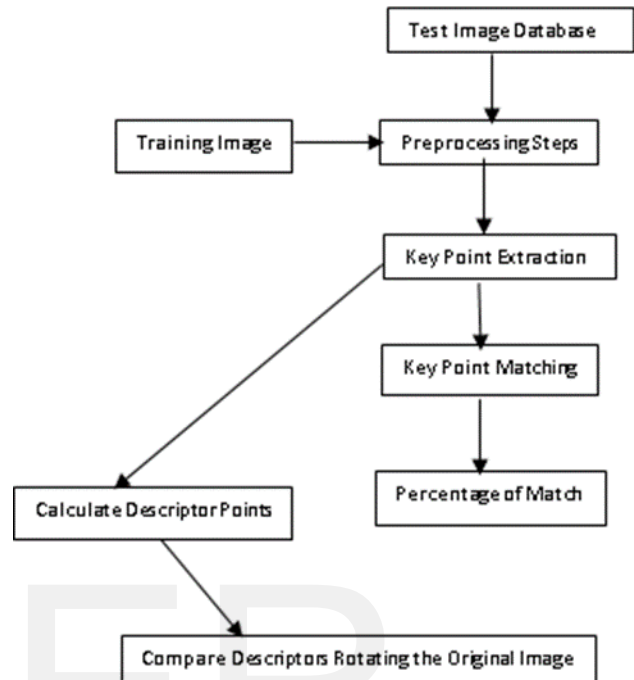


Fig. 2. Block diagram of proposed method

3.2 Keypoint Extraction

For any object interesting point on it can be extracted, which provides a feature description of the object. This descriptions extracted from a training image can then be used to identify the object when attempting to locate the object in the test image containing many other objects. These features may be affected by several factors like changes in image scale, noise, illumination etc. To perform reliable recognition it is required the features extracted from training image be detectable even changes in image scale, noise and illumination. Such points usually lie on high contrast regions of the image, such as object edges. Another important characteristic of these features is that the relative position in the original scene should not change from one image to another. We have used SIFT (Scale Invariant Feature Transform) algorithm for extracting the key points from an image [1].

SIFT is an algorithm in computer vision to detect and describe local features in an image. Each feature vector is invariant to any scaling, rotation or translation of the image. Applications include object recognition, robot localization and mapping, Panorama stitching, 3D scene modelling, recognition and tracking. For extracting these features the SIFT algorithm uses a four stage filtering approach [1].

i) Scale space extrema detection:

To detect the key points initially those locations and scales

are identified that can be repeatedly assigned under different views of the same object. For finding these locations scale space is used which is a continuous function of scale. According to Koenderink (1984) and Lindeberg (1994) Gaussian function is the best possible scale space kernel [1].

The scale space is defined by the function

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (1)$$

Where, * is the convolution operator in x and y

$G(x, y, \sigma)$ is a variable-scale Gaussian

$$\text{and } G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \quad (2)$$

$I(x, y)$ is the input image.

Then difference of Gaussians technique is used to detect stable key-point locations in the scale space and locating the extrema. For finding the extrema $D(x, y, \sigma)$ in the scale space the difference between two images, one with scale k times the other is computed [1].

So, $D(x, y, \sigma)$ is given by:

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (3)$$

For finding the local maxima and minima of $D(x, y, \sigma)$ each point is compared with its 8 neighbours at the same scale, and its 9 neighbours up and down one scale. If this value is the minimum or maximum of all these points then this point is an extrema.

ii) Key point localisation:

In the next step the low contrast and poorly localized points along an edge are extracted from all the detected keypoints. Then such points are rejected to reduce the size of the keypoints extracted.

The initial implementation made by Lowe, locates keypoints at the location and scale of the central sample point. A recent method developed by Brown and Lowe for fitting a 3D quadratic function to the local sample points to determine the interpolated location of the maximum using Taylor expansion of the scale space function [1].

The location of the extremum z is:

$$z = -\frac{\partial^2 D^{-1} \partial D}{\partial x^2 \partial x} \quad (4)$$

The value of z is compared with a threshold value and if the value of z is below a threshold value then the point is excluded. This is how extrema with low contrast is removed. In case of extrema based on poor localization, across the edge there is a large principal curvature and in the perpendicular direction in the difference of Gaussian function there is a small curvature. The keypoint is rejected if this difference is below the ratio of largest to smallest eigenvector, from the 2x2 Hessian matrix at the location and scale of the keypoint.

iii) Orientation assignment:

Based on the local image properties this step assigns a consistent orientation to the extracted keypoints. The key point descriptor, described below, can then be represented relative to this orientation and invariance to rotation can be achieved [1].

The approach taken to find an orientation is:

- Select the Gaussian smoothed image L using the keypoints scale, from above
- Compute gradient magnitude, $m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2}$
- Compute orientation, θ
$$\theta(x, y) = \tan^{-1} \left(\frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)} \right)$$
- From the gradient orientation of sample points create an orientation histogram having 36 bins which covers 360 degree of orientation.
- Then the highest peak in the histogram is located.

iv) Key point descriptor:

To create keypoint descriptors local gradient data is also used. First to line up with the orientation of the keypoint, the gradient information is rotated. Then weighted by a Gaussian with variance of $1.5 * \text{keypoint scale}$. A set of histograms over a window centered on the keypoint is created from this data.

Typically keypoint descriptors uses a set of 16 histograms which are aligned in a 4×4 grid and each having 8 orientation bins. One orientation for each of the main compass directions and one orientation for each of the mid points of the main compass directions. Finally it creates a feature vector having 128 dimensions which are nothing but the SIFT (Scale Invariant Feature Transform) keys [1].

3.3 Keypoint Matching

For finding the matching keypoints the extracted key points of both images are compared. In this case we have compared each point of the training image with every other point of the test image and found number of matching points. Then percentage of match is calculated.

3.4 SIFT as Rotation Invariant

We have also calculated the descriptor vectors and proved that SIFT is rotation invariant. For this we have computed the dot product of descriptor vectors and taken inverse cosine of it. The calculated value is compared with a threshold value to find matching descriptors.

4 EXPERIMENTAL OBSERVATIONS

We have considered a specific iris image (iris1) as training image and other images iris2 to iris10) as test images. Then the feature points extracted from iris1 are compared with

- Feature points of the same image (iris1)
- Feature points of test images (iris2 to iris10)
- Feature points of rotated iris1 image
- Feature points of scaled iris1 image

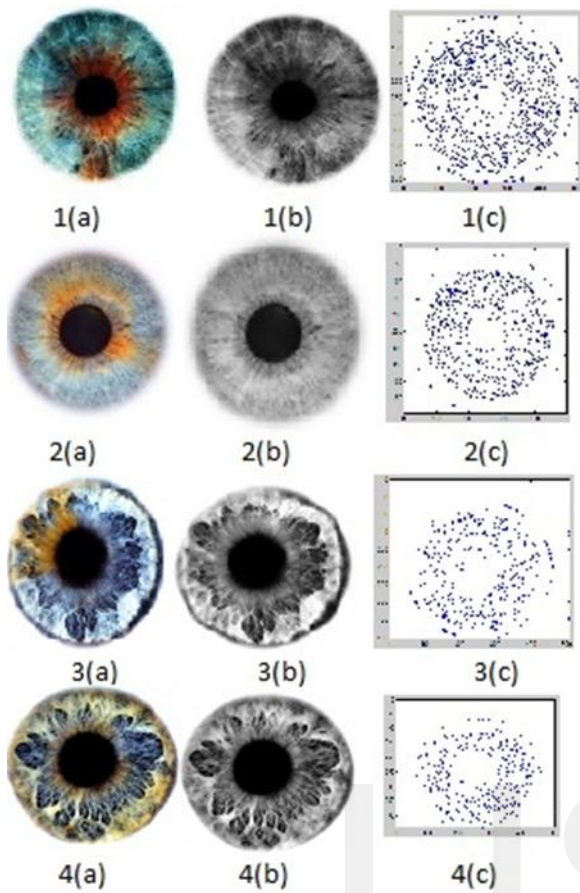


Fig. 3. Iris image (1 (a), 2(a), 3(a), 4(a)), gray scale conversion of iris image (1 (b), 2(b), 3(b), 4(b)), extracted key points (1 (c), 2(c), 3(c), 4(c))

The above figure shows four iris images in rgb format (1(a), 2(a), 3(a), 4(a)), its gray scale conversion (1(b), 2(b), 3(b), 4(b)) and the feature points extracted from the iris images (1(c), 2(c), 3(c), 4(c)). The number of feature points extracted are 996, 636, 426, 448 from first (shown in 1(c)), second (shown in 2(c)), third (shown in 3(c)) and fourth (shown in 4(c)) iris image respectively.

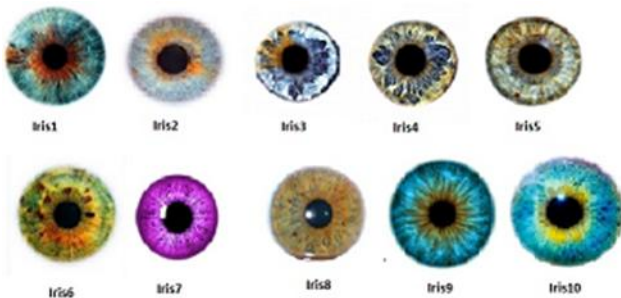


Fig. 4. Iris imageset

TABLE 1
KEYPOINTS MATCH PERCENTAGE OF THE IRIS IMAGES

| Training Image | Test Image | Number of Keypoints in Training Image(Iris1) | Number of Keypoints in Test Image | Number of Matched Keypoints | Percentage of Match |
|----------------|------------|--|-----------------------------------|-----------------------------|---------------------|
| Iris1 | Iris1 | 996 | 996 | 996 | 100 |
| Iris1 | Iris2 | 996 | 636 | 4 | 0.4016 |
| Iris1 | Iris3 | 996 | 426 | 0 | 0 |
| Iris1 | Iris4 | 996 | 448 | 4 | 0.4016 |
| Iris1 | Iris5 | 996 | 429 | 5 | 0.5020 |
| Iris1 | Iris6 | 996 | 836 | 2 | 0.2008 |
| Iris1 | Iris7 | 996 | 211 | 0 | 0 |
| Iris1 | Iris8 | 996 | 562 | 2 | 0.2008 |
| Iris1 | Iris9 | 996 | 598 | 7 | 0.7028 |
| Iris1 | Iris10 | 996 | 1016 | 2 | 0.2008 |

The above table shows the Number of extracted feature points from each iris image (iris1 to iris 10) using SIFT algorithm. The calculated percentage of match obtained from these ten (10) iris images using SIFT algorithm proves that, the percentage of match of iris1 with iris1 is 100% but with other iris images percentage of match of iris 1 is very less. So, this can be used as a good biometric identifier.

TABLE 2
TIMING ANALYSIS OF INDIVIDUAL STEPS OF SIFT

| Image | Time for Gaussian Scale Space Construction | Time for Differential Scale Space Construction | Time for Finding Key points | Number of Key points Extracted | Time for Calculating Descriptor |
|--------|--|--|-----------------------------|--------------------------------|---------------------------------|
| Iris1 | 4.063 s | 0.014 s | 0.066 s | 996 | 0.606 s |
| Iris2 | 4.032 s | 0.015 s | 0.053 s | 636 | 0.415 s |
| Iris3 | 3.999 s | 0.020 s | 0.050 s | 426 | 0.262 s |
| Iris4 | 4.186 s | 0.018 s | 0.050 s | 448 | 0.292 s |
| Iris5 | 4.035 s | 0.013 s | 0.047 s | 429 | 0.277 s |
| Iris6 | 4.167 s | 0.022 s | 0.069 s | 836 | 0.515 s |
| Iris7 | 4.140 s | 0.018 s | 0.036 s | 211 | 0.126 s |
| Iris8 | 3.939 s | 0.026 s | 0.065 s | 562 | 0.348 s |
| Iris9 | 3.954 s | 0.018 s | 0.060 s | 598 | 0.355 s |
| Iris10 | 4.617 s | 0.014 s | 0.062 s | 1016 | 0.664 s |

Then we calculate the descriptor vectors for each key point. The above table shows the time taken by different steps of SIFT algorithm. To prove that the feature descriptors are invariant to rotation we have compared the descriptors of iris1 image with the descriptors of rotated iris1 image. For this we have calculated the angle between the two vectors by taking dot product and inverse cosine transform and sort the calculated values. For finding the matched points these angular values are compared with the help of a threshold value. In this case we have taken this value as 1.1.

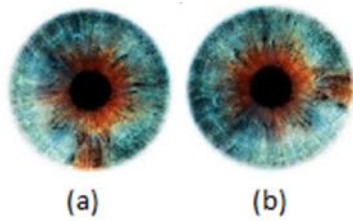


Fig. 5. Iris1 image(shown as (a)) and rotated iris1 image(shown as (b))

TABLE 3
MATCHED KEYPOINTS OF IRIS1 WITH IRIS1 USING SIFT DESCRIPTORS

| Image | Number of Keypoints Extracted | Number of Matched Descriptor Vectors |
|-------|-------------------------------|--------------------------------------|
| Iris1 | 996 | 980 |
| Iris1 | 996 | |

TABLE 4
MATCHED KEYPOINTS OF IRIS1 WITH ROTATED IRIS1 USING SIFT DESCRIPTORS

| Image | Number of Keypoints Extracted | Number of Matched Descriptor Vectors |
|---------------|-------------------------------|--------------------------------------|
| Iris1 | 996 | 980 |
| Rotated Iris1 | 1096 | |

The analysis of above two tables (table III and IV) shows that the number of matched key points is same for original iris image and rotated iris image. This proves that SIFT feature vectors are invariant to rotation.

5 CONCLUSION

In this paper we have applied SIFT algorithm on different iris images of eye. The comparison of key points extracted from different iris images proves that it can be used as a good biometric identifier. Our experimental observations also concludes that the SIFT algorithm is invariant to rotation. This is another advantage of using this algorithm. The time taken by this algorithm increases with the number of key points extracted. This approach may further be improved to reduce the time taken for extracting the keypoints.

REFERENCES

[1] D.G. Lowe, "Distinctive Image Features from Scale Invariant Keypoints," International Journal of Computer Vision, 2004.
[2] D. Rindhe and G. Sable, "Teeth Feature Extraction and Matching for Human Identification using Scale Invariant Feature Transform Algo-

rithm," European Journal of Advances in Engineering and Technology, vol. 2, no. 1, pp. 55-64, 2015.
[3] G. A. Montazer and D. Giveki, "Content Based Image Retrieval System using Clustered Scale Invariant Feature Transforms," ELSEVIER, pp. 1695-1699, 2015.
[4] J. Daugman, "How Iris Recognition Works," IEEE Transactions on Circuits and Systems for Video Technology, vol. 14, no. 1, pp. 21-30, Jan 2014.
[5] K.M. Gandhi and R.H. Kulkarni, "SIFT Algorithm for Iris Feature Extraction," Global Journal of Computer Science and Technology: F Graphics and Vision," vol. 14, no. 3, pp. 30-36, 2014.
[6] C. M. Patil and S. Patilkulkarni, "Iris Feature Extraction for Personal Identification using Lifting Wavelet Transform," International Journal of Computer Applications, vol. 1, no.14, pp. 68-72, 2010.
[7] A. Aichert, "Feature extraction techniques," CAMP Medical Seminar WS0708, pp. 1-8, 2008.
[8] S. Malathi and C. Meena, "Partial Fingerprint Matching Based on SIFT Features," International Journal on Computer Science and Engineering, vol. 2, no. 4, pp. 1411-1414, 2010.
[9] A. Jain and S. Prabhakar, "Fingerprint Matching using Minutiae and Texture Features", Proceedings of International Conference on Image Processing, pp. 282-285. 2001.
[10] I. Khan, A. Khan and R. A. Shaikh, "Object Analysis in Image Mining," 2nd International Conference on Computing for Sustainable Global Development, pp. 1985-1988, 2015.