© 2018 IJSRSET | Volume 4 | Issue 8 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
Themed Section : Engineering and Technology

# Feature Extraction using SIFT for Privacy Preserving Over Encrypted Image

**Pranjali M Marne[1], Prof. P. M. Kamde[2]**

[1]Department of Computer Engineering, Sinhgad College of Engineering, Savitribai Phule, Pune University, Pune, Maharashtra, India

[2]Department of Computer Engineering, Sinhgad College of Engineering, Pune, Savitribai Phule, Pune University, Pune, Maharashtra, India

## ABSTRACT

In past a huge number of the data owners stores their individual multimedia information so for the communication to be carried out securely, the private information shouldn't be reveal by its originated application of the owners as it does outsourced multimedia. Thus it is important to maintain the privacy by preserving the computation in outsource multimedia data. Thus the privacy-preservation will be implemented using SIFT (Scale Invariant Feature Transform) algorithm over the image. The data set owner is responsible to have the image data set thus which is responsible to send the image to server thus the feature extraction is done by using SIFT. The Batched Secure Multiplication Protocol for the communication between the servers and Batched Secure Comparison protocol for the comparison. It uses somewhat holomorphic encryption technique. TheCaltech256 standard dataset is used and the application users will have high level of security on image data.

Keywords : SIFT, Encryption, Feature Selection, Secure Data Outsourcing.

## I. INTRODUCTION

Data Owner mostly prefer remote server for storing large amount of data/information which is multimedia files. Due to high utilization of data services the process is little costly for storing the data. If we use the unlimited storage at the a time as well as estimation of resources for the prospect of minimizing cost and flexibility. Large amount of security with privacy is provided by Storing data on remotely located server . Basically providers like Flickr; Facebook Especially uses outsourced private image information. For increasing experience of user the data is used for conducting behavioral advertising and preference analytics. While dealing with large dataset, system retrieve features of image as an input to characterized model of data mining.

If we use unlimited storage at the a time and estimation of resources for the prospect of minimizing cost and flexibility. Large amount of security with privacy is provided by Storing data on remotely located server . Basically providers like Flickr, Facebook Especially uses outsourced private image information.

At the owner's side workload in extraction of image feature as well as use on the Big Data for building functional user-defined applications is over whelming. the purpose of main picture data to the semi-trusted service provider may essentially uncover the information owner individual data for example monetary profiles individual identity and areas. Before outsourcing it is important to encrypt delicate multimedia information to ensure the security of data

While giving solid plans security, encryption additionally turns into prevention to information computation or usage. Currently the main problem is the manner by which to provide privacy preserving picture feature extractions over huge picture information while as if mitigating the database proprietor of its high calculation load and depending on the server for giving quick and powerful image feature extraction service. In the current literature, the issue of privacy-preserving outsourcing is focus on modular exponentiation, linear programming, and keyword-based search.

On various data types, like numerical data, text data, spatial data etc., has been hugely observed lately. Not with standing, the issue of privacy-preserving computation outsourcing over multimedia information has gotten restricted research consideration up until now. As of late, privacy-preserving information search in the domain of cipher text has been developed to content based multimedia retrieval recovery and face recognition. In view of the significance of image feature extraction in processing of multimedia data and its overwhelming operations on generous information, particularly for satellite informational set for its tremendous size as well as feature points, the extraction or disclosure of picture features from the cipher text domain has started to attract more and more research interest. To the best of our insight, Hsu et al. was the first to discover privacy-preserving SIFT in the encrypted area it uses homomorphic encryption. Notwithstanding, their answer is either computationally recalcitrant or insecure from the security viewpoint. While putting extraordinary exertion on the security or effectiveness viewpoint, one regular downside of this is that they all need comprehensive evaluations regarding the protection of the key characteristics of its relating original picture feature extraction algorithm. As such, regardless of whether these arrangements can very much ensured the important characteristics of the first SIFT as far as distinctiveness and robustness or not stays being referred. This system proposed an effective privacy preserving outsourcing protocol over huge encrypted picture information for the usual scale in variant feature transform (SIFT).

## II. LITERATURE REVIEW

In paper [1], proposed SIFT outsourcing convention which is rely upon implemented secure intuitive conventions BSMP also in BSCP. They likewise recognize and evaluate the security and furthermore adequacy of created framework.

In paper [2] have considered past framework and methods utilized, they demonstrated that past strategies are not secure or proficient. Additionally observed that none of the framework made the utilization of extremely indispensable attributes of the original SIFT in terms of distinctiveness and robustness. After that creator proposed a framework which is productive and in addition satisfies all the security needs without loss of its key qualities, by separating the original image information arbitrarily. Finally, composed two conventions for correlation and secure multiplication and sending of the feature extraction computation onto two independent cloud servers.

In paper [3] created privacy preserving dynamic medical content mining and furthermore image feature extraction conspire PPDM in cloud-assisted e-healthcare system. At begin they have built up an effective privacy preserving completely holomorphic information aggregation from any one way trapdoor strategy that will help the executed PPDM. After that outsourced disease modeling and additionally prior intervention is gained, as needs by conceiving a privacy preserving technique relationship coordinating PPDM1 utilizing dynamic medical content mining and in addition planning a privacy preserving medical image highlight extraction PPDM2.

In paper [4] implemented a privacy preserving and duplicate prevention content based image recovery framework in a cloud computing. For encryption of visual features the secure KNN calculation is utilized. Similarity scores can be registered with the encrypted features by the cloud server that lets the cloud server to rank the images maintaining a strategic distance from the correspondence stack.

In paper [5] proposed POP framework that will makes feasible for cloud servers to give privacy preserving photo sharing and scanning service for cell phone clients who needs to outsource photo management and at the time ensuring their privacy in photos. They are designed framework that gives security to the outsourced images for restricting access from unapproved clients, likewise makes conceivable to the clients to encode their image to look along these lines the search can be outsourced towards untrusted cloud server.

In paper [6] by utilizing greedy depth first search calculation built up an special tree based record structure for giving productive multi keyword ranked search. The developed framework conveys sub-linear search time tackles the deletion and insertion of document adaptably.

In paper [7] proposed security of the SIFT system relying upon holomorphic encryption. Author showed that examination of security relying upon the discrete logarithm issue and RSA which PPSIFT maintains a strategic distance from cipher text just attack also a portion of the plaintext attacks.

In paper [8] proposed halfway and developed inquiry strategy between the customer and the server. The inquiry performed in server with the expanded question rundown, and sends back every coordinating thing to the customer and it performs a hunt inside of the got coordinating arrangement of unique inquiry results. The off tunable idea security was utilized to modify the protection assurance level as per an arrangement. The hash-based piecewise altered indexing is to separation an element vector into pieces and file every piece with a sub hash esteem in amplified question.

## III. PROPOSED APPROACH

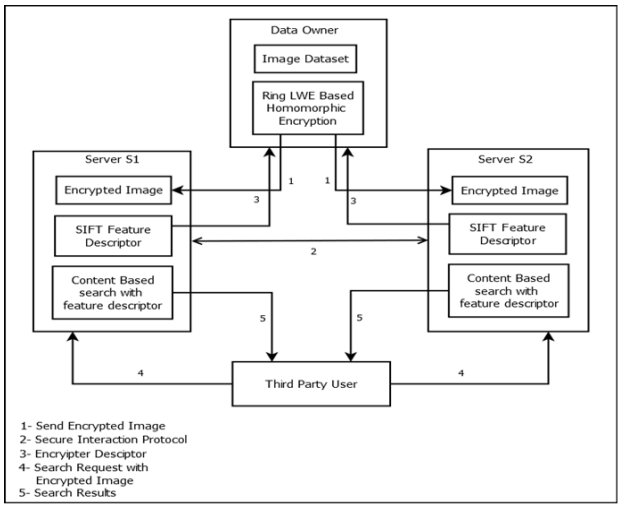


Figure 1.Proposed System Architecture

### A. Proposed System Overview

Detailed descriptions of the proposed system are as follows:

In recent most of the data owners interested in outsourcing their huge amount of personal multimedia data onto the cloud as it is the cost efficient and flexible solution. Such data is used by most of the service providers or any other applications for various purpose such as learning, searching or for behavioral advertising. In such cases sometimes this outsourced multimedia data may disclose the data owner's private information.

The protocol of privacy preserving computation over outsourced multimedia data when Scale-Invariant Feature Transform (SIFT) feature extraction method apply on encrypted image data is used. With the help of these extracted feature descriptors content based search is done by third party users with high level of

security over encrypted image data. To test the performance of system, used Caltech256 data-set and analyzed proposed system.

Consider a server based image feature extraction outsourcing calculation framework including three users: the information proprietor O, the server S1, server S2. In this system, consider that the O holding a huge volume of delicate image records is resource constrained. Consequently, might want to outsource both the image data set and the calculation intensive SIFT assignment to the cloud by utilizing its inexhaustible storage and calculation resources. By assuming that S1 and S2 are independent with each other and can be considered to have a place with two independent cloud service co-ops like Amazon EC2 and Microsoft Azure. To ensure the security, O will initially encrypt each image set and after that distribute the cipher text S1 and S2. After implementing SIFT calculation in the cipher content area by means of secure interaction conventions, S1 and S2 will restore the encrypted highlight descriptors to the information proprietor, who can in the end recoup the real feature descriptors from their encrypted versions.

In the proposed system the data owner is responsible to encrypt the images from the data set by using the blow fish algorithm, which is block cipher and it process in bytes thus it is send to server1 and server2 .At the server side the server has the encrypted images and a key which is send by the data owner and the decryption process is carried out at the server side. After the image is decrypted the SIFT is applied thus the feature are extracted and the feature are shown on the image and send to owner.

### A. Algorithm

I) SIFT(Scale Invariant Feature Transform)

Scale Space Extreme Detection

In this stage identify the locations and scales that are identifiable from different views of the same object. For this scale function is used, defined as:

$L(xi, yi, \sigma) = G(xi, yi, \sigma) * I(xi, yi)$

Where * is the convolution operator, $G(xi, yi, \sigma)$ is a variable-scale Gaussian and I(xi, yi) is the input image. Difference of Gaussians is one such technique, locating scale-space extreme, $D(xi, yi, \sigma)$ by computing the difference between two images, one with scale k times the other. D(x, y, $D(xi, yi, \sigma) = L(xi, yi, k \sigma) - L(xi, yi, \sigma)$) is then given by:

$D(xi, yi, \sigma) = L(xi, yi, k \sigma) - L(xi, yi, \sigma)$

To detect the local maxima and minima of D(xi, yi, $D(xi, yi, \sigma) = L(xi, yi, k \sigma) - L(xi, yi, \sigma)$) each point is compared with its 8 neighbors at the same scale, and its 9 neighbors up and down one scale. If this value is the minimum or maximum of all these points then this point is an extreme.

Keypoint Localization

This technique is used for eliminating more points from the list of key points by discovering those that have low contrast or are poorly localized on an edge. Laplacian function is used for this. If the function value at z is below a threshold value then this point is excluded.

Orientation Assignment

Consistent orientation is assign to the key points on the basis of properties of local image.

Keypoint Descriptor

The local gradient data is used for creating keypoint descriptors. Information of gradient is rotated for line up with orientation of keyword after that it weighted by a gaussian variance of 1.5 * keypoint scale. Result of this is then utilize for creating a set of histogram over a window centered on the keypoint.

II) KNN Search Algorithm

**Input**: dataset of encrypted image features

1. In knn algorithm we form k-number of cluster in which every clusters follow some property.
2. In our system all the image features are clustered according to nearest neighbor of centroid point.
3. In every block feature having max no of images will be consider for the cluster
4. At this stage we form the k number of clusters where every cluster has unique image feature.
5. At testing stage we get one image and find its feature
6. We pass these feature to knn search algorithm
7. Knn search will search these testing features in our cluster block. After matching the cluster we will get all the image features of that cluster as a result.
8. We form the images of that feature and get output result.

### B. Mathematical Model

S = {Input, Process, Output}, where S is the system.

Input: Encrypted Image Dataset

Output: Content based image search over encrypted image data with feature descriptor

Process:

1) Image Encryption:

Image I(xi, yi) is encrypted by computing I1(xi,yi) as:

$$I1(xi, yi) = I(xi, yi) + I2(xi, yi)$$

2) Keypoint Localization:

The locations are disclosed to the two servers.

S1 will compute:

$$L_1(xi, yi, \sigma) = G(xi, yi, \sigma).\, I1(xi, yi)$$

Variable scale Gaussian:

$$G(xi, yi, \sigma) = \frac{1}{2\pi\sigma^2}.\, e^{\frac{-(xi^2+yi^2)}{2\sigma^2}}$$

S1 further compute:

$$D_1(xi, yi, \sigma) = L_1(xi, yi, r6) - L_1(xi, yi, \sigma)$$

S2 will do same operations parallel.

3) Orientation Assignment:

$$\begin{aligned}\Delta Diff1 = Diff_{1y} - kDiff_{1xi} \\ = \alpha.\,(L1(xi, yi + 1) \\ - L1(xi, yi + 1) \\ - k\beta.\,(L1(xi + 1, yi) - L1(xi \\ - 1, yi))\end{aligned}$$

4) Descriptor Generation:

5) Content based search over encrypted descriptor:

## IV. RESULTS AND DISCUSSION

### *A.* Experimental Setup

a) Software Requirement and Specification

The system used the following software for implementation:

- JDK 1.8
- Netbeans

### *B.* Expected Result

In this section discussed the experimental result of the proposed system.

Table 1: Accuracy Comparison

| System | Time in ms | Accuracy in % |
|---|---|---|
| Existing System (with SIFT) | 1300 | 80% |
| Proposed System (with KNN Search) | 900 | 90% |

Figure 2 shows that Time Comparison graph. Comparison graph shows that the time required for implementing the proposed system is less than the time required for implementing the existing system.

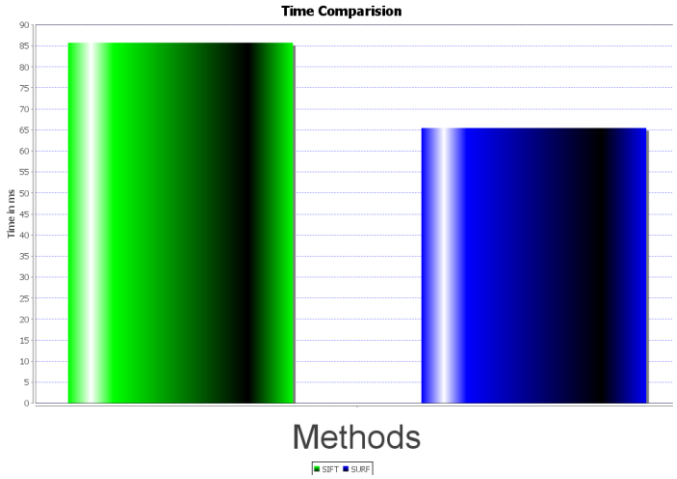


Fig 2: Time Comparision

Figure 3 shows the accuracy of the proposed system with the existing system. it shows that the proposed system is more accurate than existing system.

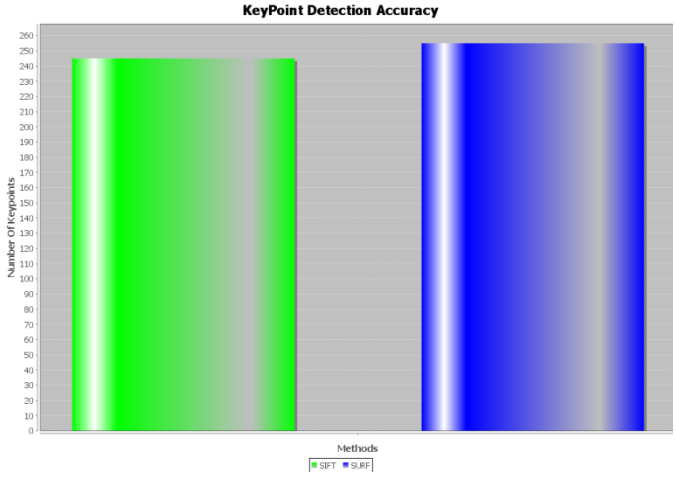


Fig 3: Accuracy Comparison

## V. CONCLUSION AND FUTURE SCOPE

In proposed system novel privacy-preserving SIFT outsourcing protocol is proposed by considering existing secure interactive protocols BSMP and BSCP for feature descriptor retrial from outsourced picture dataset. It fulfill the security of search image as well as content based search on encrypted multimedia information after applying content based search. Both examine and considerably examine efficiency and security of the system. From the experimental results it is conclude that the system gives more accurate result by using SURF algorithm as compared to the existing system.

## VI. REFERENCES

1. QWang, S. Hu, J.Wang, Z.Wang, K. Ren, and M. Du "Computation of Feature Extractions Over Encrypted Image Data" 2016.
2. S Hu, Q. Wang, J. Wang, Z. Qin and K. Ren "Securing SIFT: Privacy-Preserving Outsourcing Computation of Feature Extractions Over Encrypted Image Data" 2016.
3. J Zhou, Z. Cao, X. Dong and X. Lin "PPDM: A Privacy-Preserving Protocol for Cloud-Assisted e-Healthcare Systems" 2015.
4. Z Xia, X. Wang, L. Zhang, Z. Qin, X. Sun and K. Ren "A Privacy-Preserving and Copy-Deterrence Content-Based Image Retrieval Scheme in Cloud Computing" 2016.
5. X Ding, T. Jung, C. Liu, X. Y. Li , L. Zhang, and Y. Liu, "POP: Privacy-Preserving Outsourced Photo Sharing and Searching for Mobile Devices" 2015 .
6. XWang, Q.Wang, X. Sun, Z. Xia, "A secure and dynamic multi-keyword ranked search scheme over encrypted cloud data" 2015.
7. S-C. Pei and C.-S. Lu, "Image feature extraction in encrypted domain with privacy-preserving sift 2012.
8. L Amsaleg, Weng, A. Morton, and S. Marchand Maillet, "Aprivacy preserving framework for large-scale content-based information retrieval" 2015.
9. -Y. Hsu, C.-S. Lu, and S.-C. Pei "Image feature extraction in encrypted domain with privacy-preserving sift" 2012.
10. Y. Huang, D. Evans, J. Katz, and L. Malka "Faster secure two-party computation using garbled circuits 2011.
11. M J. Atallah and K. B. Frikken, "Securely outsourcing linear algebra computations" 2010.