Private Image Computation: the Case of Cloud based Privacy-preserving SIFT

Zhan Qin, Jingbo Yan, Kui Ren
Department of Electrical and Computer Engineering
SUNY at Buffalo, Buffalo, NY 14260, USA
Email: zhanqin@buffalo.edu, jbyan@xidian.edu.cn, kuiren@buffalo.edu

***Abstract*—In this paper, we present SecSIFT, a high-performance cloud based image feature detection system for performing Scalar Invariant Feature Transform (SIFT) over private image data without compromising the privacy. In contrast to previous works, we outsource the computation of image feature detection to a set of independent, co-operative cloud servers, and keep the outsourced computation procedures as simple as possible. Using this framework, we are not restricted by efficiency limitations of homomorphic encryption scheme and thus are able to implement applications such as social discovery or behavior prediction with less complexity on computation and communication.**

I. Introduction

There is an emerging research field on discovering solutions to enable various image feature detection algorithms over outsourced image data while still preserving its privacy [1]–[5]. Most existing works related to privacy in image feature extraction focus on different problem settings and aspects: In secure multimedia data search, the authors propose a design to protect the privacy of the query image when searching over a public non-encrypted database in [2]. The oblivious retrieval is utilized to protect query in [3]. Another method to generate the feature vectors without leaking private information is proposed in [5]. In [4], a scheme is proposed to utilize Paillier encryption scheme from [6] to enable a secure SIFT computation over the ciphertext. Nevertheless, their design has a serious security flaw in secure comparison and requires unpractical computational resources: the proposed homomorphic comparison algorithm is neither secure nor efficient.

However, to our knowledge, none of the proposals cited above have made a noticeable impact on commercially deployed systems. Regarding the research on private image feature detection systems, we believe that all of the previous proposals have two common shortcomings that make them unsuitable for practical acceptance: ambiguous *privacy definition* and *scalability*. Regarding *privacy definition*, all of the previous works on local image features extraction ignores the privacy of the feature point's location and leaks them to cloud server. Regarding *scalability*, existing solutions ignore or cannot solve the problem of efficiency: in real world, the cloud computing service still runs in a Pay-per-User model. In other words, the cloud can only provide economical but not unlimited computing resources and is not able to accommodate existing massive resource-demanding solutions. All of the previous works require at least public-key operations or something even more expensive. Moreover, they also require roughly one kilobit of bandwidth per single bit of outsourced image data.

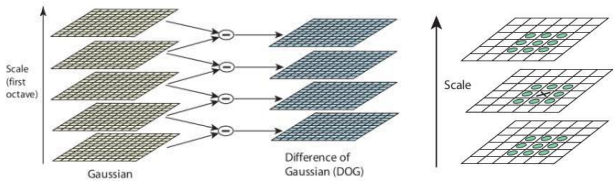

Fig. 1. **SIFTOverview** : The image forms a series of octave via convolutions and subductions

This paper presents SecSIFT, a system for performing privacy-preserving image feature detection over private user image data that scales substantially better than previous systems. SecSIFT operates under similar trust model as several recent private proposals, i.e., honest-but-curious servers that do not collude. What differentiate our proposal from other homomorphic based privacy-preserving image feature extraction solutions (such as [3] and [4]) are two points: How the privacy of local feature's value and location are defined, and How different procedures in computation are divided between cloud servers. Prior works only protect the image data by encrypt them in a homomorphic encryption scheme. They either leaks the location of feature point to cloud [4] or requires unpractical computational resources [3]. In our proposal, however, the location and value of the feature points and image are protected by dividing the whole algorithm into different computation tasks and allocating them to different cloud servers. The *scalability* is derived from the fact that SecSIFT uses simple noise addition and dummy insertion operations as its crypto primitive in different component of the system.

II. SecSIFT Overview

A. Brief Intro to SIFT

Scale Invariant Feature Transform (SIFT) is one of the most popular algorithm in computer vision to extract and describe local features in images. We separate it into mainly 2 steps to describe the algorithm briefly:

- 1. Scale-space Extrema Detection: the Difference of Gaussian (DoG) values are first extracted from a set of reference images with various σ. Then these DoG values compose a series of octave over scale and space. For

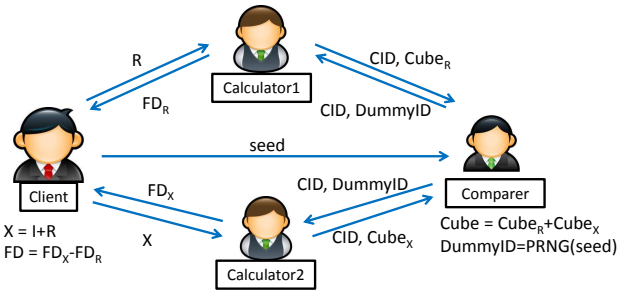

Fig. 2. **SystemOverview** : Client uploads image I and seed to Calculator1, Calculator2, and Comparer, receives Feature Descriptor (FD) from cloud components.

example, one pixel is compared with its 26 neighbors to find a local extrema, which is a keypoint (feature point).

- 2. Feature Descriptor Generation: Once we find the keypoint, a 16×16 neighborhood around the keypoint is taken. It is devided into 16 sub-blocks of 4×4. For each sub-block, 8 bin orientation histogram is created. So a total of 128 bin values are available. It is represented as a vector to form Feature Descriptor.

B. System Components and Workflow

Figure 2 shows a high-level overview of SecSIFT, which consists of the following system components:

Client: The *client* stores image data, and intend to outsource their image data to cloud to perform SIFT algorithm to extract feature descriptors (FD).

Calculators and Comparer: The *comparer* and two *calculators* compose the cloud server. They work together to ensure that the image from the clients and the extracted feature descriptors to the *clients* are handled in a way that satisfies our privacy requirements.

Workflow: We give a in a step-by-step brief description of our design as follows:

- 1) At the beginning, the *client* generates the pseudo-random noise R to hide the image I pixel by pixel, which forms another matrix R and $X = R + I$.
- 2) After receiving encrypted images R and X from *client*, the *calculator* performs convolution and subduction as the original SIFT algorithm to get octaves and form image cubes composed by 27 neighboring pixels. Then, it encrypts all pixels in the processed image cubes through Order Preserving Encryption (OPE) scheme, and shuffles the surrounding pixels in the cube randomly. After that it and pass each cube to the *comparer* with a pre-defined cube ID (CID).
- 3) After receiving encrypted image cubes $Cube_R$ and $Cube_X$ from *calculators*, the *comparer* decrypt them easily by additions and locates the cubes that include extrema by a series of comparisons. After that, it first generates a group of $dummyIDs$ with the $seed$ from the *client*, then mix them with the $CIDs$ connected with extrema and send them back to *calculators*.
- 4) The *calculators* get the polluted $CIDs$ and then perform the adjusted feature descriptor generation procedures similar to the original SIFT algorithm. The *client* can simply recover the encrypted Feature Descriptor FD by subduction and discard the dummy feature descriptors.

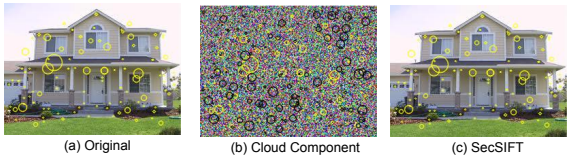

Fig. 3. **SampleExperiment** :

C. Privacy Analysis

The SecSIFT achieves two privacy properties:

Confidentiality: No system component can get any information of the pixel value or its location in the image matrix at the same time. According to the accepted security standard in image encryption [7], the content of the image is confidential.

Delocalization: To each component in the system, it is computational hard to get the location of neighboring feature points in the image, which means the feature point location of the image is protected. Breaking the delocalization property could enable system components to launch deduction attacks over the private images.

In summary, all pixel values and feature points locations are unrecoverable to all system components, i.e., the input and the output of the system do not violate the privacy of any client information.

III. Sample Experiment

In Figure 3, (a) shows the original results of SIFT algorithm, (b) shows the view of *calculator* (Regarding *comparer*, it is computational hard to recover the pixels' original locations). The yellow circle represents the true feature point and the black circle represents dummy feature points. From the view of cloud, all the processed image pixels are its features are confidential and delocalized to all components. To conclude, the preliminary results are promising and confirm the feasibility of our design on privacy-preserving image feature detection system.

References

[1] Y. Ke and R. Sukthankar, "Pca-sift: A more distinctive representation for local image descriptors," in *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, vol. 2. IEEE, 2004, pp. II–506.

[2] J. Shashank, P. Kowshik, K. Srinathan, and C. Jawahar, "Private content based image retrieval," in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*. IEEE, 2008, pp. 1–8.

[3] W. Lu, A. L. Varna, A. Swaminathan, and M. Wu, "Secure image retrieval through feature protection," in *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on*. IEEE, 2009, pp. 1533–1536.

[4] C.-Y. Hsu, C.-S. Lu, and S.-C. Pei, "Secure and robust sift," in *Proceedings of the 17th ACM international conference on Multimedia*. ACM, 2009, pp. 637–640.

[5] E. M. Newton, L. Sweeney, and B. Malin, "Preserving privacy by de-identifying face images," *Knowledge and Data Engineering, IEEE Transactions on*, vol. 17, no. 2, pp. 232–243, 2005.

[6] S. D. Galbraith, "Elliptic curve paillier schemes," *Journal of Cryptology*, vol. 15, no. 2, pp. 129–138, 2002.

[7] Y. Wang, K.-W. Wong, X. Liao, and G. Chen, "A new chaos-based fast image encryption algorithm," *Applied soft computing*, vol. 11, no. 1, pp. 514–522, 2011.