PRIVACY-PRESERVING PATTERN RECOGNITION WITH IMAGE COMPRESSION

Takayuki Nakachi[1] and Hitoshi Kiya[2]

[1]Nippon Telegraph and Telephone Corporation,Kanagawa, Japan
[2]Tokyo Metropolitan University, Tokyo, Japan

ABSTRACT

In this paper, we propose a privacy-preserving pattern recognition scheme that well supports image com- pression. The proposed scheme is based on secure sparse coding using a random unitary transform. It offers the following two prominent features: 1) It is capable of pattern recognition in the encrypted image domain. Even if data leaks, privacy can be maintained because data remains encrypted. 2) It realizes Encryption-then-Compression (EtC) systems, where image encryption is conducted prior to compression. The pattern recognition can be carried out in the compressed signal domain using a few sparse coefficients. Based on the pattern recognition result, it can compress the selected images with high quality by estimat- ing sufficient number of sparse coefficients. We use the INRIA dataset to demonstrate its performance in detecting humans. The proposal is shown to realize human detection with encrypted images and efficiently compress the images selected in the image recognition stage.

KEYWORDS

Surveillance Camera, Pattern Recognition, Secure Computation, Sparse Coding, Random Unitary Trans- form

1. INTRODUCTION

With the increase in threats and criminal activity, security is seen as a major public concern. Image/video surveillance is one approach to addressing this issue. Many image/video surveillance systems are now widely deployed in many public spaces such as airports, banks, shopping streets, public streets, etc., and they are recording huge amounts of image/video every day. Fortunately, edge/cloud computing offers an efficient way of handling and analyzing the huge amounts of image/video data. However, edge/cloud computing poses some serious issues for end users, such as unauthorized use, data leaks, and privacy failures due to the unreliability of providers and accidents [1].

Many studies have examined the processing of encrypted data; most proposals use homomorphic encryption (HE) and secure multiparty computation (MPC) [2]. Even though service providers cannot directly access the native content of the encrypted signals, they can still apply HE and MPC. In particular, fully homomorphic encryption (FHE) allows arbitrary computation on en- crypted data [3]. However, these methods impose high communication costs, high computation complexity or large cipher text size, so further advances are needed for attractive applications such as big data analysis and advanced image/video processing. We take the random unitary transform approach as we focus on secure image processing [4]. Random unitary transform based encryp- tion methods have lower communication costs, lower computation complexity or small cipher text size. We continue to study secure sparse coding for pattern recognition [5]-[8],

Natarajan Meghanathan et al. (Eds): SIPP, BIGML, DaKM, SOEN, AISC - 2020
pp. 01-12, 2020. © CS & IT-CSCP 2020 DOI: 10.5121/csit.2020.100201

Encryption- then-Compression (EtC) systems [9]-[11]. Orthogonal Matching Pursuit (OMP), a sparse coding algorithm, is executed in the encrypted signal domain.

Early work on sparse coding was based on the efficient coding hypothesis, which states that the goal of visual coding is to faithfully reproduce the visual input while minimizing the neural effort [12]. It effectively represents observed signals as the linear combination of a small number of atoms. Sparse dictionary learning has been successfully applied to various image/video and audio processing applications [13]-[16]. The effectiveness of sparse coding has been reported for pattern recognition [15], image compression [16]. For example, the experiments of Ref. [16] show that rate-distortion based sparse coding outperforms JPEG and JPEG2000 by up to 6+ dB and 2+ dB, respectively.

In this paper, we propose a privacy-preserving pattern recognition scheme that extends previously proposed EtC methods [9]-[11]. The secure pattern recognition methods and EtC systems mentioned above were proposed separately. This current proposal offers not only image pattern recognition but also image compression. The integrated system is realized by performing pattern recognition in the secure compressed domain. 1) It is capable of efficient pattern recognition in the encrypted image domain. Even if data leaks, privacy is maintained because the data remains encrypted. 2) It works as an EtC system. Pattern recognition and image compression can be carried out seamlessly in the same compressed signal domain. This means that the proposed secure OMP algorithm chooses the atoms sequentially and then calculates the sparse coefficients. Pattern recognition employs the few sparse coefficients. Based on the pattern recognition result, additional atoms are chosen and used to compress the selected images. Finally, we employ the INRIA person dataset to evaluate the human detection performance of the proposed method [17]. Detecting humans in images is essential for not only image/video surveillance but also many applications such as automatic driver assistance, etc.

The organization of this paper is as follows. In Sec. 2, we explain related work. Section 3 describes sparse coding for image modeling. In Sec. 4, we propose secure sparse coding for secure sparse coding for pattern recognition with image compression. Section 5 shows simulation results. Conclusions and future work are given in Sec. 5.

2. RELATED WORK

In this section, we review the conventional secure pattern recognition methods and Encryption-then-Compression (EtC) systems.

2.1. Secure Pattern Recognition

We have proposed secure sparse coding for pattern recognition [5]-[8]. Feeding the encrypted images into the secure OMP computation yields the sparse coefficients used for pattern recognition. We verified that by adopting the random unitary transform, the pattern recognition performance is not degraded, which proves that the proposed framework operates securely with no performance degradation. Furthermore, compared with deep-learning based methods such as SPCANet [18], the sparse coding based method has several prominent advantages such as 1) low computational complexity and less data needed for training, 2) transparent machine learning: the algorithm is interpretable as the optimization problem is written in closed form. Refs. [6][7] detail the experiments and results.

2.2. Encryption-then-Compression (EtC) Systems

Encryption-then-Compression (EtC) systems [9]-[11] [19]-[21] have been proposed to securely transmit and compress images through an untrusted channel provider; the traditional technique is to use Compression-then-Encryption (CtE) systems. EtC systems allow us to close non-encrypted

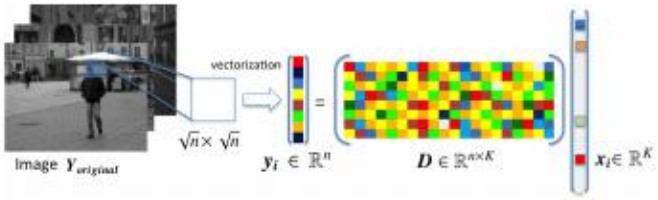


Figure 1: Sparse coding for image patches

images to SNS providers, because encrypted images can be directly compressed even when the images are multiply recompressed by SNS providers. Well-known EtC systems are block scrambling-based encryption schemes that are compatible with international standards, e.g. JPEG, JPEG2000, etc [19]-[21]. While the sparse coding based EtC systems [9]-[11] are not compatible with international compression standards, they do provide high coding performance because they form dictionaries that fit the observed signals.

3. SPARSE CODING FOR IMAGE MODELING

In this section, we overview sparse coding for image modeling which is the basis of secure pattern recognition and EtC systems.

3.1. Sparse Coding for Image Patches

We consider image patches of size $\sqrt{n} \times \sqrt{n}$ pixels that are ordered lexicographically as column vectors $\boldsymbol{y}_i = \{y_1, , ..., y_n\}^T \in \mathbb{R}^n$. The patches are extracted from image $\boldsymbol{Y}_{original}$ as shown in Fig. 1. We assume that every image patch $\boldsymbol{y}_i$ can be represented sparsely given the over-complete dictionary $\boldsymbol{D} = \{\boldsymbol{d}_1, ..., \boldsymbol{d}_K\} \in \mathbb{R}^{n\times K}$ whose columns contain K prototype atoms $\boldsymbol{d}_i$:

$$\boldsymbol{y}_i = \boldsymbol{D}\boldsymbol{x}_i, \tag{1}$$

where $\boldsymbol{x}_i = \{x_1, ..., x_K\}^T \in \mathbb{R}^K$ are sparse coefficients, $i = 1, \cdots, N$, and N is the total number of patches. In advance, dictionary $\boldsymbol{D}$ is designed for the images by training algorithms such as MOD [23] and K-SVD [24].

If n < K and D is a full-rank matrix, an infinite number of solutions to the representation problem are available. The solution with the fewest number of nonzero coefficients is certainly an appealing representation. This sparsest representation is the solution given by

$$(P_0) \quad \min_{\boldsymbol{x}_i} \|\boldsymbol{x}_i\|_0 \quad \text{subject to} \quad \boldsymbol{y}_i = \boldsymbol{D}\boldsymbol{x}_i, \tag{2}$$

where $\|\cdot\|_0$ is the l_0-norm, counting the nonzero entries of the vector. Extraction of the sparsest representation is, however, an NP-hard problem [25].

3.2. Selection of Dictionary Atoms

Dictionary atoms are typically estimated by a ”pursuit algorithm” that finds the following approximate solution:

$$\boldsymbol{x}_i = \arg\min_{\boldsymbol{x}_i} \left\|\boldsymbol{y}_i - \boldsymbol{D}\boldsymbol{x}_i\right\|_2^2 \quad \text{subject to} \quad \|\boldsymbol{x}_i\|_0 < \epsilon_i. \tag{3}$$

We assume dictionary **D** is fixed. Well-known pursuit algorithms include Orthogonal Matching Pursuit (**OMP**) [22]. **OMP** is a greedy, step-wise regression algorithm. At each stage, **OMP** selects the dictionary atom having the maximal projection onto the residual signal. After each selection, the representation coefficients w.r.t. the atoms selected so far are found via least-squares search.

3.3. Dictionary Learning

An over-complete dictionary $\boldsymbol{D}$ is designed by adapting its content to fit a given set of images. Given the set $\boldsymbol{Y}=\{\boldsymbol{y}_i\}_{i=1}^N$, we assume that there exists a dictionary, $\boldsymbol{D}$, that can recreate the given images via sparse combinations. The overall mean square error of a representation is given by

$$E = \|\boldsymbol{Y} - \boldsymbol{D}\boldsymbol{X}\|_2^2 . \tag{4}$$

MOD (Method of Optimal Direction) [23] and K-SVD (K-Singular Value Decomposition) [24] are well-known dictionary learning algorithms. Assuming that $\boldsymbol{X}=\{\boldsymbol{x}_i\}_{i=1}^N$ is fixed, the MOD algorithm allows us to seek an update to $\boldsymbol{D}$ such that the above error is minimized. Taking the derivative of (4) with respect to $\boldsymbol{D}$, yields

$$\boldsymbol{D} = \arg\min_{\boldsymbol{D}} \|\boldsymbol{Y} - \boldsymbol{D}\boldsymbol{X}\|_F^2 = \boldsymbol{Y}\boldsymbol{X}^T(\boldsymbol{X}\boldsymbol{X}^T)^{-1}. \tag{5}$$

K-SVD is an iterative method that uses singular value decomposition; it alternates between sparse Coding based on the current dictionary and the process of updating the dictionary atoms to better Fit the data.

4. Secure Sparse Coding for Pattern Recognition With Image Compression

In this section, we propose a privacy-preserving pattern recognition system that offers image compression as an integrated component. The integrated system is realized by performing pattern recognition in the secure compressed domain.

4.1. Secure Computation Architecture

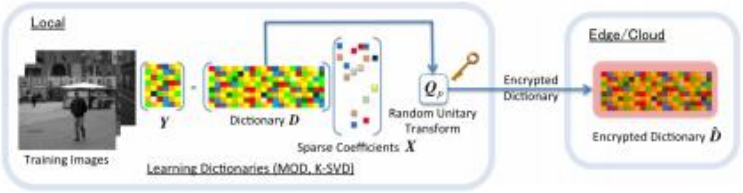
Figure 2 illustrates the architecture of privacy-preserving pattern recognition with image compression. It is based on the sparse coding of image patches. Figure 2(a) shows the training step. Dictionary $\boldsymbol{D}$ is designed by MOD or K-SVD algorithm at the local site. Feeding the training images to the learning algorithm yields dictionary $\boldsymbol{D}$. Next, we apply random unitary transform function $T(\cdot)$ to dictionary $\boldsymbol{D}$ to generate encrypted dictionary $\hat{\boldsymbol{D}}$. Encrypted dictionary $\hat{\boldsymbol{D}}$ is sent to the appropriate edge/cloud site and stored in a database.

Figure 2(b) shows the running step. The local site applies the same random unitary transform function $T(\cdot)$ to test image $\boldsymbol{Y}$ to generate encrypted image $\hat{\boldsymbol{Y}}$. Then encrypted image $\hat{\boldsymbol{Y}}$ is sent to the edge/cloud site. The edge/cloud site uses encrypted image $\hat{\boldsymbol{Y}}$ and encrypted dictionary $\hat{\boldsymbol{D}}$ to perform secure OMP computation. Secure OMP chooses the atoms sequentially and calculates the sparse coefficients $\boldsymbol{X}$ from the encrypted $\hat{\boldsymbol{Y}}$ and $\hat{\boldsymbol{D}}$. At first, pattern recognition is carried out in the compressed signal domain using a few sparse coefficients. $\hat{\boldsymbol{X}}^P$ is the set of the few sparse coefficients used for pattern recognition. Then the images selected by the pattern recognition stage are compressed. For this compression, additional atoms are chosen and calculates the sparse coefficients by secure OMP computation. $\hat{\boldsymbol{X}}^C$ is a set of sparse coefficients used for compression.

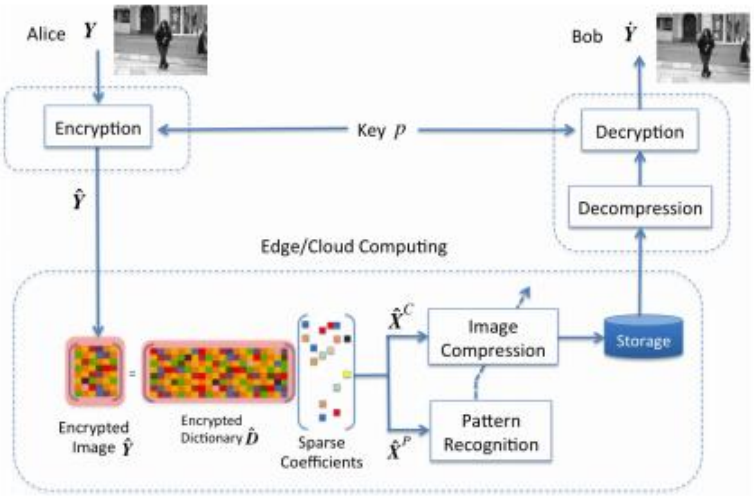
4.2. Random Unitary Transform

The encrypted images and dictionary are generated by using the random unitary transform approach. A vector $\boldsymbol{f}_i$ $(i = 1, \cdots, L) \in \mathbb{R}^N$ is encrypted by a random unitary matrix $\boldsymbol{Q}_p \in \mathbb{C}^{N \times N}$ with a private key p as follows

$$\hat{\boldsymbol{f}}_i = T(\boldsymbol{f}_i, p) = \boldsymbol{Q}_p \boldsymbol{f}_i, \tag{6}$$



(a) Training: generating encrypted dictionary



(b) Running: pattern recognition with image compression

Figure 2: Architecture of privacy-preserving pattern recognition with secure OMP computation.

where $\hat{\boldsymbol{f}}_i$ is an encrypted vector; L is the number of vectors. Note that the random unitary matrix $\boldsymbol{Q}_p$ satisfies

$$\boldsymbol{Q}_p^* \boldsymbol{Q}_p = \boldsymbol{I} \tag{7}$$

where $[\cdot]^*$ and I mean the Hermitian transpose operation and the identity matrix, respectively. In addition to unitarity, Q_p must have randomness for generating the encrypted signal. GramSchmidt orthogonalization is a typical method for generating Q_p. Furthermore, the encrypted vector has the following properties.

· Property 1: Conservation of Euclidean distances.

$$\left\| \boldsymbol{f}_i - \boldsymbol{f}_j \right\|_2^2 = \left\| \hat{\boldsymbol{f}}_i - \hat{\boldsymbol{f}}_j \right\|_2^2 \tag{8}$$

· Property 2: Norm isometry

$$\|\hat{\boldsymbol{f}}_i\|_2^2 = \|\boldsymbol{f}_i\|_2^2 \tag{9}$$

· Property 3: Conservation of inner products.

$$\boldsymbol{f}_i^* \boldsymbol{f}_j = \hat{\boldsymbol{f}}_i^* \hat{\boldsymbol{f}}_j \tag{10}$$

4.3. Secure OMP Computation

The proposed secure sparse coding computation generates encrypted signal $\hat{\boldsymbol{y}}_i$ and dictionary $\hat{\boldsymbol{D}}$ by the following transforms:

$$\hat{\boldsymbol{y}}_i = T(\boldsymbol{y}_i, p) = \boldsymbol{Q}_p \boldsymbol{y}_i \tag{11}$$

$$\hat{\boldsymbol{D}} = T(\boldsymbol{D}, p) = \boldsymbol{Q}_p \boldsymbol{D}. \tag{12}$$

The sparse coefficient $\hat{\boldsymbol{x}}_i$ is estimated for each image patch $\hat{\boldsymbol{y}}_i$. Instead of Eq. (3), we consider the following optimization problem in which $\hat{\boldsymbol{y}}$ and $\hat{\boldsymbol{D}}$ are assumed to be given:

$$\hat{\boldsymbol{x}}_i = \underset{\boldsymbol{x}}{arg\ min} \| \hat{\boldsymbol{y}}_i - \hat{\boldsymbol{D}}\boldsymbol{x}_i \|_2^2 \quad subject\ to \quad \|\boldsymbol{x}_i\|_0 < \epsilon. \tag{13}$$

The sparse coefficient $\hat{\boldsymbol{x}}_i$ yielded by secure OMP computation is the same result as that created by the non-encrypted version [9]-[11]. The algorithm is shown below (prefix i of $\hat{\boldsymbol{x}}_i$ and $\hat{\boldsymbol{y}}_i$ is omitted for notation simplicity):

Secure OMP Computation Algorithm

Initialization: $k = 0$, and set

· The initial solution $\boldsymbol{x}^0 = \boldsymbol{0}$

· The initial residual $\hat{\boldsymbol{r}}^0 = \hat{\boldsymbol{y}} - \hat{\boldsymbol{D}}\mathrm{x}^0 = \hat{\boldsymbol{y}} = \boldsymbol{Q}_p \boldsymbol{y}$

· The initial solution supports $\boldsymbol{S}^0 = \emptyset$.

Main Iteration:

Increment k by 1 and perform the following steps:

· Sweep: Compute the errors

$$\hat{\epsilon}(i) = \left\|\boldsymbol{r}^{k-1}\right\|_2^2 - \frac{(\boldsymbol{d}_i \cdot \boldsymbol{r}^{k-1})^2}{\|\boldsymbol{d}_i\|_2^2}. \tag{14}$$

· Update Support: Find the minimizer

$$\begin{aligned} i_0 &= \underset{i \notin \boldsymbol{S}^{k-1}}{arg\ min} \{\hat{\epsilon}(i)\} \\ &= \underset{i \notin \boldsymbol{S}^{k-1}}{arg\ min} \{\epsilon(i)\}, \boldsymbol{S}^k = \boldsymbol{S}^{k-1} \cup \{i_0\}. \end{aligned} \tag{15}$$

· Update Provisional Solution: compute

$$\hat{\boldsymbol{x}}^k = \{(\hat{\boldsymbol{D}}_{\boldsymbol{S}^k})^T \hat{\boldsymbol{D}}_{\boldsymbol{S}^k}\}^{-1} \{(\hat{\boldsymbol{D}}_{\boldsymbol{S}^k})^T \hat{\boldsymbol{y}}\}. \tag{16}$$

· Update Residual: compute

$$\hat{\boldsymbol{r}}^k = \boldsymbol{Q}_p \boldsymbol{r}^k. \tag{17}$$

· Stopping Rule:
If $\|\hat{r}^k\|_2 < \epsilon$, stop. Until satisfaction is achieved, commence another iteration. Alternative stopping rule is given by

$$k = T_k, \tag{18}$$

where T_k is the number of specified atoms. Iteration is repeated until the number of selected atoms reaches T_k.

Output: The proposed solution $\hat{x}$ is obtained after k iterations.

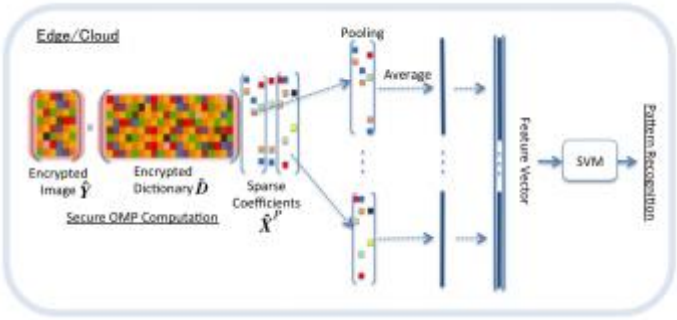


Figure 3: Feature extraction and classification.

4.4. Feature Extraction and Classification

The secure OMP algorithm select atoms sequentially and calculates the corresponding sparse coefficients for each image patch. We use just a few sparse coefficients (calculated using only $k = 1$ *or* 2 iterations) for pattern recognition. Figure 3 shows the procedure of feature extraction and classification. The sparse coefficients at each image patch are used for formatting the feature vector. In order to reduce the dimension of the feature, we take the statistics of the spatially local sparse coefficients of atoms as the feature, which corresponds to local spatial pooling. Multiple sparse coefficients $\boldsymbol{x}_i$, which correspond to local $B \times B$ image patch $\boldsymbol{y}_i$, are grouped into the averaged sparse coefficient $\bar{\boldsymbol{x}}_j (j =, 1, 2, \cdots, N/B^2)$, where B is block size. The averaged sparse coefficients $\bar{\boldsymbol{x}}_j$ are vectorized to produce feature vector $\vec{\boldsymbol{x}}$.

SVM is a supervised machine learning algorithm that can be used for both classification or regression tasks, but it is mostly used for the former. In SVM, we input a feature vector $\vec{\boldsymbol{x}}$ to the discriminant function as

$$(\vec{\boldsymbol{x}}) = sign(\omega^T \vec{\boldsymbol{x}} + b) \tag{19}$$

with

$$sign(u) = \begin{cases} 1(u > 1) \\ -1(u \leq 1), \end{cases} \tag{20}$$

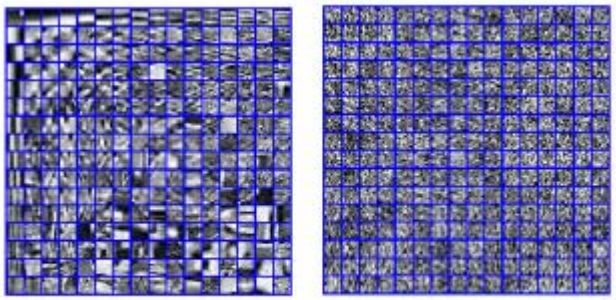
where ω is a weight parameter and b is a bias. SVM also has a technique called the kernel trick, which is a function that takes a low dimensional input space and transforms it into a higher dimensional space. This can be used for non-linear classification. For the pattern recognition task, classification is performed using a linear SVM. The SVM is trained using task data from training subjects.

4.5. Quality Control for Image Compression

Feeding the encrypted dictionary and the encrypted image into the secure OMP computation yields the sparse coefficients $\hat{\boldsymbol{x}}_i$ for each image patch $\boldsymbol{y}_i$. The decoded image $\dot{\boldsymbol{y}}_i$ can be obtained by $\dot{\boldsymbol{y}}_i = \boldsymbol{Q}_p^* \hat{\boldsymbol{D}} \hat{\boldsymbol{x}}_i$. This means that the proposed scheme can work as an EtC system. The image quality of the decoded image $\dot{\boldsymbol{y}}_i$ can be controlled by threshold ϵ_i, which determines the stopping condition of the secure OMP algorithm, i.e. $\left\| \boldsymbol{r}_i^k \right\|_2 < \epsilon_i$. In order to keep the image quality of each image patch, the same threshold is set: $\epsilon_i = \text{constant}$ $(i = 1, \cdots, N)$. An alternative stopping rule is $k = T_k$. In this case, the number of atoms in each patch is set to be the same.

5. EXPERIMENTAL RESULTS

We carried out experiments on detecting humans in images from the INRIA person dataset [17]. Here we assume that we compress only those that include human(s) captured by surveillance systems.

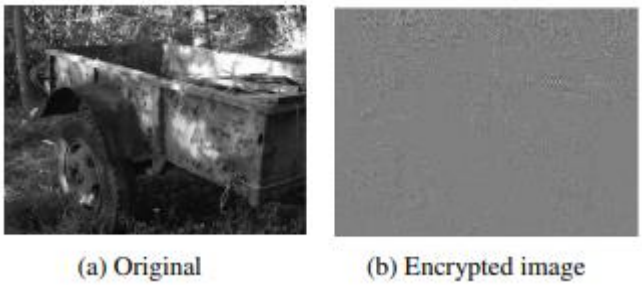


(a) Dictionary $\boldsymbol{D}$ (b) Encrypted dictionary $\hat{\boldsymbol{D}}$

Figure 4: A trained dictionary and corresponding encrypted dictionary for human images.

(a) Original (b) Encrypted image

Figure 5: A sample of original and encrypted human images.



(a) Original (b) Encrypted image

Figure 6: A sample of original and encrypted non-human images.

5.1. INRIA Person Dataset and Parameters

The INRIA person dataset is one of the most popular and widely used pedestrian detection benchmark datasets. The INRIA person dataset contains images of various sizes with and without humans. We evaluated the performance of the proposed method by challenging it with 480 × 640 pixels human and non-human images. The parameters settings are as follows:

1) Designing K-SVD: We applied K-SVD and trained a dictionary of size 64 × 256. The training data consisted of a set of image patches of size 8 × 8 pixels, randomly taken from 20 human images.
2) Creating the random unitary transform: We generated a 64 × 64 random unitary transform by the Gram-Schmidt orthogonalization method.
3) Designing and running the SVM: block size B=20 for local pooling of the sparse coefficients. For the human detection task, two-class classification is performed using a linear SVM. In the training step, the SVM is trained using 100 images (50 human images and 50 non-human images).

In the evaluation, we used 10-fold cross-validation. 100 images were partitioned into 10 subsamples (a single sub-sample contains 5 human and 5 non-human images). Of the10 sub-samples, a single sub-sample is retained as the validation data for testing, and the remaining 9 subsamples

Table 1: Detection Rate (DR) [%] of the proposed method.

(a) Number of atoms: $L = 1$

Test	1	2	3	4	5	6	7	8	9	10	Ave.
DR	100	70	80	70	90	90	80	60	90	70	80

(b) Number of atoms: $L = 5$

Test	1	2	3	4	5	6	7	8	9	10	Ave.
DR	90	60	90	70	90	90	80	50	100	70	79

Table 2: Detection Rate (DR) [%] of the non-encrypted method.

(a) Number of atoms: $L = 1$

Test	1	2	3	4	5	6	7	8	9	10	Ave.
DR	100	70	80	70	90	90	80	60	90	70	80

(b) Number of atoms: $L = 5$

Test	1	2	3	4	5	6	7	8	9	10	Ave.
DR	90	60	90	70	90	90	80	50	100	70	79

are used as training data. The cross-validation process is then repeated 10 times, with each of the 10 subsamples used exactly once as the validation data. The 10 results were then averaged to produce a single estimate.

5.2. Results

The trained dictionary and corresponding encrypted dictionary are shown in Fig. 4. Figures 5 and 6 show the original and corresponding encrypted images for a sample of human and non-human images, respectively. Feeding the encrypted dictionary and the encrypted images into the secure OMP computation yielded the sparse coefficients $\hat{x}_i$ for each image patch y_i.

Detection rate of the proposed privacy-preserving pattern recognition method is shown in Table 1. We evaluated two cases: the number of atoms L = 1 and L = 5. Detection rate is calculated by

$$Detection\ rate = \frac{Number\ of\ images\ correctly\ detected}{Total\ number\ of\ test\ images}. \quad (21)$$

Table 1 shows that the proposed method achieves a detection rate of around 80 [%]. Note that the results were obtained from encrypted images. Setting the number of atoms at L = 1 or L = 5 yielded almost the same performance. For comparison, we evaluated a pattern recognition method with the input being the non-encrypted version of OMP. Detection rate of the non-encrypted version is shown in Table 2. The 10-fold cross-validation used the same training and testing datasets for non-encrypted version of OMP and the secure OMP. The results show that the proposal has exactly the same detection performance as the non-encrypted version of the pattern recognition method.

Figure 7 plots coding efficiency (number of atoms vs. decoded image quality PSNR [dB]) for the selected human images. We controlled the image quality of the human images at each patch by setting number of atoms L = {1, 2, 3, 4, 5}. This figure shows that proposed method increases decoded image quality by adding the atoms sequentially. Note that there is no need to decompress and decrypt images when running the secure OMP algorithm.

6. CONCLUSIONS AND FUTURE WORK

In this paper, we proposed privacy-preserving pattern recognition with image compression. The pattern recognition can be carried out in the compressed signal domain. It can efficiently compress the images selected by the pattern recognition stage. We confirmed its performance by

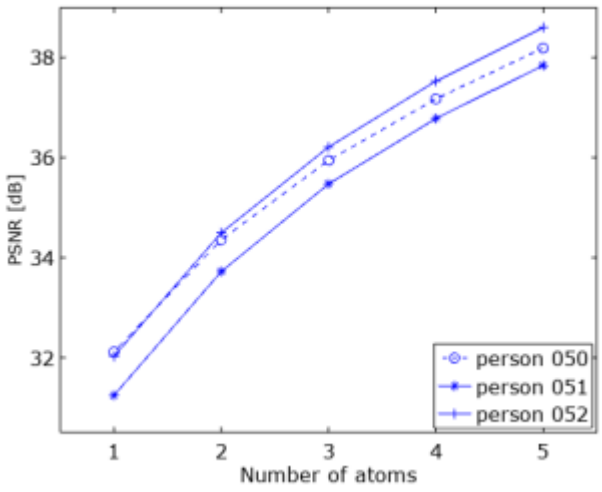


Figure 7: Coding efficiency (Number of atoms L vs. decoded image quality).

detecting humans in the INRIA dataset. In terms of estimation accuracy for pattern recognition, these experiments are merely the first step. Further study is required to enhance the proposal's performance.

REFERENCES

[1] C. T. Huang, L. Huang, Z. Qin, H. Yuan, L. Zhou, V. Varad-harajan, and C-C. J. Kuo,"Survey on securing data storage in the cloud,"APSIPA Transactions on Signal and Information Processing, vol. 3, e7, 2014.

[2] R. L. Lagendijk, Z. Erkin, and M. Barni, "Encrypted signal processing for privacy protection: Conveying the utility of homomorphic encryption and multiparty computation," IEEE Signal Processing Magazine, vol. 30, no. 1, pp. 82-105, Jan. 2013.

[3] Z. Brakerski, "Fundamentals of fully homomorphic encryption - A survey," Electronic Colloquium on Computational Complexity, report no. 125, 2018.

[4] I. Nakamura, Y. Tonomura, and H. Kiya,"Unitary transform-based template protection and its application to l2-norm minimization problems,"IEICE Transactions on Information and Systems, vol. E99-D, no.1, pp. 60-68, Jan. 2016.

[5] Y. Wang, T. Nakachi, and H. Ishihara, "Edge and cloud-aided secure sparse representation for face recognition," 27th European Signal Processing Conference (EUSIPCO 2019), Sep. 2019.

[6] Y. Wang and T. Nakachi, "Towards secured and transparent AI technologies in hierarchical computing networks," NTT Technical Review, <https://www.nttreview.jp/archive/2019/201909.html>, vol. 9, 2019.

[7] Y. Wang and T. Nakachi, "Secure face recognition in edge and cloud networks: from the ensemble learning perspective," IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP2020), to be presented.

[8] T. Nakachi, Y. Wang, and H. Kiya, "Privacy-preserving pattern recognition using encrypted sparse representations in L0 norm minimization," IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP2020), to be presented.

[9] T. Nakachi and H. Kiya, "Practical secure OMP computation and its application to image modeling," Proceedings of the 2018 International Conference on Information Hiding and Image Processing (IHIP2018), Sep. 2018.

[10] T. Nakachi, Y. Bandoh, and H. Kiya, "Secure dictionary learning for sparse representation," 27th European Signal Processing Conference (EUSIPCO 2019), Sep. 2019.

[11] T. Nakachi and H. Kiya, "Secure sparse representations in L0 norm minimization and its application to EtC systems," 13th International Conference on Signal Processing and Communication Systems (ICSPCS2019), d13, pp. 61-67, Dec. 2019.

[12] H. B. Barlow, "Possible principles underlying the transformation of sensory messages," Sensory Communication, pp. 217-234,1961.

[13] M. Elad, "Sparse and redundant representations: from theory to applications in signal and image processing," Springer, 2010.

[14] M. Elad, " Sparse and redundant representation modeling - what next?, "IEEE Signal Processing Letters, vol. 19, no. 12, pp. 922-928, Dec. 2012.

[15] Z. Jiang, Z. Lin, and L. S. Davis, "Label consistent K-SVD: learning a discriminative dictionary for recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35, no. 11, pp. 2651-2664, Nov. 2013.

[16] X. Zhang, W. Lin, Y. Zhang, S. Wang, S. Ma, L. Duan, and W. Gao, ”Rate-distortion optimized sparse coding with ordered dictionary for image set compression,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 28, no. 12, pp. 3387-3397, Dec. 2018.

[17] ”INRIA Person Dataset,” http://pascal.inrialpes.fr/data/ human/.

[18] L. Tian, C. Fan, Y. Ming, and Y. Jin,“ Stacked PCA network (SPCANet): an effective deep learning for face recognition,”2015 IEEE International Conference on Digital Signal Processing (DSP), pp. 1039-1043, Jul. 2015.

[19] K. Kurihara, M. Kikuchi, S. Imaizumi, S. Shiota, and H. Kiya, ”An encryption-thencompression system for JPEG/Motion JPEG standard,” IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, E98-A, no. 11, pp. 2238- 2245, 2015.

[20] W. Sirichotedumrong and H. Kiya,“ Grayscale-based block scrambling image encryption using YCbCr color space for Encryption-then-Compression systems, ”APSIPA Trans. Signal and Information Processing, vol. 8, no. E7, February 2019.

[21] T. Chuman, W. Sirichotedumrong, and H. Kiya, ”Encryption-then-compression systems using grayscale-based image encryption for JPEG images,” IEEE Transactions on Information Forensics and Security, vol. 14, no. 6,, pp. 1515-1525, June 2019.

[22] Y. C. Pati, R. Rezaiifar, Y. C. P. R. Rezaiifar, and P. S. Krishnaprasad,“Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition,” Proceedings of 27th Asilomar Conference on Signals, Systems and Computers, pp. 40-44, 1993.

[23] K. Engan, S. O. Aase, and J. Hakon Husoy, “ Method of optimal directions for frame design,”1999 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP1999), pp. 2443-2446, 1999.

[24] M. Aharon, M. Elad, and A. Bruckstein,“ K-SVD: An algorithm for designing overcomplete dictionary for sparse representation, ”IEEE Transactions on Signal Processing, vol. 54, no. 11, pp. 4311-4322, Nov. 2006.

[25] B. K. Natarajan,“Sparse approximate solutions to linear systems,”SIAM Journal on Computing, 24, 2, pp. 227-234, 1995.

AUTHORS

Takayuki Nakachi received the Ph.D. degree in electrical engineering from Keio University, Tokyo, Japan, in 1997. Since joining Nippon Telegraph and Telephone Corporation (NTT) in 1997, he has been engaged in research on super-high-definition image/video coding and media transport technologies. From 2006 to 2007, he was a visiting scientist at Stanford University. Dr. Nakachi is a member of the Institute of Electrical and Electronics Engineers the Institute of Electronics (IEEE) and the Information and Communication Engineers (IEICE) of Japan.

Hitoshi Kiya received his B.E and M.E. degrees from Nagaoka University of Technology, in 1980 and 1982, respectively, and his Dr. Eng. degree from Tokyo Metropolitan University in 1987. In 1982, he joined Tokyo Metropolitan University, where he became a Full Professor in 2000. From 1995 to 1996, he attended the University of Sydney, Australia as a Visiting Fellow. He is a Fellow of IEEE, IEICE and ITE. He currently serves as President-Elect of APSIPA, and he served as Inaugural Vice President (Technical Activities) of APSIPA from 2009 to 2013, and as Regional Director-at-Large for Region 10 of the IEEE Signal Processing Society from 2016 to 2017.