SIFT KEYPOINT REMOVAL VIA CONVEX RELAXATION

An Cheng, Yuanman Li and Jiantao Zhou

Department of Computer and Information Science, University of Macau
{mb25436, mb25510, jtzhou}@umac.mo

ABSTRACT

Due to the high robustness against various image transformations, Scale Invariant Feature Transform (SIFT) has been widely employed in many computer vision and multimedia security areas to extract image local features. Though SIFT has been extensively studied from various perspectives, its security against malicious attack has rarely been addressed. In this work, we demonstrate that the SIFT keypoints can be effectively removed, without introducing serious distortion on the image. This is achieved by formulating the SIFT keypoint removal as a constrained optimization problem, where the constraints are well-designed to suppress the existence of local extremum and prevent generating new keypoints within a local cuboid in the scale space. We show that such optimization problem in the ideal case is non-convex. To make the computation feasible, we propose a relaxation technique to convexify the original problem, while maximally preserving the solution space. As demonstrated experimentally, our proposed SIFT removal algorithm significantly outperforms the state-of-the-arts in terms of *keypoint removal rate-distortion (KRR-D)* performance. Our results imply that an authorization mechanism is required for SIFT-based systems to verify the validity of the input data, so as to achieve high reliability.

Index Terms— SIFT, keypoint removal, convex optimization, convex relaxation

1. INTRODUCTION

Image local feature extraction and description are important ingredients in many pattern recognition and multimedia security systems [1, 2]. Among various approaches for extracting image features, SIFT has become extremely popular, due to the excellent robustness against partial occlusion, clutter, noise, lighting changes, and geometric transformations [1]. In addition to the traditional content based image retrieval (CBIR) systems [3], SIFT was also employed in multimedia security systems, e.g., to detect the copy-move forgery [2]. The success of SIFT in these applications depends on the assumption that SIFT keypoints and the associated feature descriptors cannot be severely destroyed without seriously affecting the image quality.

Unfortunately, some recent studies showed that scale space image features, including SIFT ones, could be destroyed by suppressing the local extremum in the scale space. This could be a big threat to the reliability and security of many systems built upon the SIFT features. Imagine that a criminal may erase the SIFT keypoints of his ID photo, and then can pass the security checking system running over the SIFT feature domain.

The pioneer work [4] attempted to inhibit a SIFT keypoint by duplicating another local extremum in the detection region. Do *et al.* later argued that this attack was not enough to be a threat for a CBIR system, since new SIFT points will be generated around the original one [5]. It was found that the descriptors of these newly generated SIFT keypoints are very similar to the original one, and hence, can still be matched [5]. To improve the removal performance, Do *et al.* suggested two strategies [6]. The first one is to force the contrast value of each keypoint to be lower than the *contrast threshold*. These keypoints will be filtered out in the refinement stage. However, the problem of new keypoint generation (NKG) still remains unsolved. The second strategy utilized the local and global smoothing for erasing keypoints, at the cost of large distortion of the resulting image. In [7], Amerini *et. al.* first performed classification of SIFT keypoints, and then selected an existing removal strategy for each class, so as to improve the visual quality of the resulting image. To get better tradeoff between the keypoint removal and image distortion, Lu and Hsu constructed a constrained optimization framework [8]. As will be clear shortly, the constraints incorporated into their optimization framework are too restricted, which seriously narrows the solution space, leading to large distortion. Further, the NKG problem was not fully resolved.

In this work, we study the SIFT keypoint removal problem by proposing a new constrained optimization framework, where we design a set of novel constraints to specifically address the local extremum suppression and NKG problem simultaneously. We show that the ideal constraints are actually non-convex, making the whole optimization problem difficult to be solved. To make the computation feasible, we

This work was supported in part by the Macau Science and Technology Development Fund under grants FDCT/009/2013/A1, FDCT/046/2014/A1, in part by the Research Committee at University of Macau under grants MRG007/ZJT/2015/FST, MRG021/ZJT/2013/FST, MYRG2014-00031-FST, and in part by the National Science Foundation of China under grants 61402547.

adopt a convex relaxation technique to convexify the original non-convex problem, while maximally preserving the solution space. It can be shown that the framework in [8] can be treated as a restricted case of our proposed scheme. As verified experimentally, the KRR-D performance can be significantly improved. To further demonstrate the effectiveness, we provide a case study of defeating a SIFT-based image copy-move forgery detection system.

The rest of the paper is organized as follows. Section 2 gives the preliminary of the SIFT algorithm. In Section 3, we present our method for removing SIFT keypoints via convex relaxation technique. The experimental results on SIFT keypoint removal over the UCID database and a case study are reported in Section 4. We conclude in Section 5.

2. PRELIMINARY OF SIFT

SIFT is one of the most popular algorithms in computer vision to extract and describe image local features [1]. The SIFT algorithm can be divided into two stages: i) keypoint identification via extremum detection in the scale space, and ii) feature descriptor generation.

At the first stage, the input image $\mathbf{I}$ is convolved with Gaussian filters at multiple scales, to generate successive Gaussian-blurred images. The keypoints are then taken as the extreme points of the Difference of Gaussians (DoG) domain. For a specific scale s, a DoG image is defined as

$$D_{\mathbf{I}}(x,y,s) = L_{\mathbf{I}}(x,y,s+1) - L_{\mathbf{I}}(x,y,s) \tag{1}$$

where $L_{\mathbf{I}}(x,y,s)$ is the Gaussian blurred image given by

$$L_{\mathbf{I}}(x,y,s) = \mathbf{I}(x,y) \otimes G(x,y,\sigma_s) \tag{2}$$

Here $G(x,y,\sigma_s)$ is the Gaussian kernel with standard deviation σ_s, and the subscript $\mathbf{I}$ corresponds to the input image. The convolved images are grouped by octave. Letting $\mathbf{k} = (x,y,s)$, we can write $D_{\mathbf{I}}(\mathbf{k}) \triangleq D_{\mathbf{I}}(x,y,s)$. Each DoG value $D_{\mathbf{I}}(\mathbf{k})$ will be compared with its 26 neighbors within a $3 \times 3 \times 3$ cube centered by it. If $D_{\mathbf{I}}(\mathbf{k})$ is a local extremum (minimum or maximum), then it will be selected as a candidate keypoint. In the sequel, we interchangeably use the index $\mathbf{k}$ or the DoG value $D_{\mathbf{I}}(\mathbf{k})$ to refer to a point in the scale space. All the candidate keypoints will be further refined according to a contrast threshold and edge threshold. At the stage ii), a 128-dimensional descriptor will be assigned for each survived keypoint, encoding its surrounding information in the scale space. For more details, please refer to [1].

3. PROPOSED SIFT KEYPOINTS REMOVAL METHOD

In this section, we present our technique to effectively remove the SIFT keypoints. The design goals are two fold: 1) significantly reduce the number of SIFT keypoints, and 2) maintain high quality of the resulting image. These two goals are fundamentally conflicting with each other. More severe SIFT keypoint removal generally leads to larger distortion. To achieve the first goal, we attempt to suppress local extremum in the scale space, and avoid to generate new SIFT keypoints in a local cuboid. For the sake of simpler presentation, we give our SIFT keypoint removal algorithm by taking the first octave of the scale space as an example. The derived technique can be naturally extended to remove the SIFT keypoints in all the octaves.

The SIFT keypoints can be straightforwardly localized in the scale space, because the SIFT algorithm is completely transparent even for a malicious attacker. Let $\mathbf{k}_o = (x_o, y_o, s_o)$ be the index of a generic SIFT keypoint in the scale space. Define

$$\mathcal{S}_o = \Big\{(x,y,s)\Big| |x-x_o| \le 1, |y-y_o| \le 1, |s-s_o| \le 1, \; x,y,s \in \mathcal{Z}\Big\} \tag{3}$$

as the index set of all the 27 points in the $3 \times 3 \times 3$ cube centered by $\mathbf{k}_o$. As $\mathbf{k}_o$ is a keypoint, one of the following two inequalities holds

$$D_{\mathbf{I}}(\mathbf{k}_o) > D_{\mathbf{I}}(\mathbf{k}), \forall \mathbf{k} \in \mathcal{S}_o \setminus \{\mathbf{k}_o\} \tag{4}$$

$$D_{\mathbf{I}}(\mathbf{k}_o) < D_{\mathbf{I}}(\mathbf{k}), \forall \mathbf{k} \in \mathcal{S}_o \setminus \{\mathbf{k}_o\} \tag{5}$$

To eliminate the keypoint $\mathbf{k}_o$, an effective way is to violate the above two inequalities defined over the DoG domain simultaneously. We should also bear in mind that such process needs to be conducted with minimum affection of the image quality, which is measured over the pixel domain. We first extract a local image patch $\mathbf{p}_o$ of size $P \times P$ centered by (x_o, y_o) from the input image $\mathbf{I}$, where $P = 7$ in our implementation. Letting $\mathbf{E}_o$ be the operator for extracting the image patch $\mathbf{p}_o$, we can write $\mathbf{p}_o = \mathbf{E}_o \circ \mathbf{I}$. Our target is to generate a new image patch $\hat{\mathbf{p}}_o$ such that, in the new image $\hat{\mathbf{I}}$ accommodating $\hat{\mathbf{p}}_o$, $\mathbf{k}_o$ is not a keypoint in the scale space, and no new keypoints are generated in its surroundings. Clearly, we have $\hat{\mathbf{p}}_o = \mathbf{E}_o \circ \hat{\mathbf{I}}$, and all the remaining pixels in $\hat{\mathbf{I}}$ are copied from the available $\mathbf{I}$. We then can formulate the following *generic* constrained optimization problem

$$\begin{aligned} &\min_{\hat{\mathbf{p}}_o} ||\mathbf{p}_o - \hat{\mathbf{p}}_o||_2^2 \\ \text{s.t.} \quad &(C.1): \mathbf{k}_o \text{ is not an extremum in } \mathcal{S}_o \text{ of } \hat{\mathbf{I}} \\ &(C.2): \text{no new keypoints generated} \end{aligned} \tag{6}$$

The objective function is straightforward, aiming at minimizing the distortion between the resulting patch and the original one. Our contribution in this work primarily lies in the

determination of the two constrains $(C.1)$ and $(C.2)$ in an appropriate manner. Certainly, to make the above optimization problem tractable, it is desired to make both $(C.1)$ and $(C.2)$ convex, which permits efficient numerical implementations.

3.1. Determination of $(C.1)$

The purpose of imposing the constraint $(C.1)$ is to ensure that $\mathbf{k}_o$ is not a keypoint in the new image $\hat{\mathbf{I}}$. An effective way to this end is to make the inequalities given in (4) and (5) invalid simultaneously. Let

$$\alpha_o = \min_{\mathbf{x}\in\mathcal{S}_o\setminus\{\mathbf{k}_o\}} D_{\hat{\mathbf{I}}}(\mathbf{x}), \tag{7}$$

$$\beta_o = \max_{\mathbf{x}\in\mathcal{S}_o\setminus\{\mathbf{k}_o\}} D_{\hat{\mathbf{I}}}(\mathbf{x}) \tag{8}$$

The condition that $\mathbf{k}_o$ is no longer an extremum (and hence a keypoint) within $\mathcal{S}_o$ can be expressed as the following inequality

$$\alpha_o \le D_{\hat{\mathbf{I}}}(\mathbf{k}_o) \le \beta_o, \tag{9}$$

where the two inequalities hold because $\mathbf{k}_o$ is not identified as a keypoint if more than one extremum exists within the same $\mathcal{S}_o$. Unfortunately, the constraint (9) is non-convex (formal proof will be given in our future work), which makes the resulting optimization problem difficult to solve. To tackle this challenge, we resort to a convex relaxation technique to approximate the non-convex constraint in (9). Clearly, the approximated constraint should be convex. Meanwhile, we need to preserve the solution space as much as we can. This is because too strong relaxation, though capable of retaining convex constraint, may cause the solution space far away from the truly optimal solution, leading to severe distortion.

We notice that the difficulty of evaluating α_o and β_o is because the minimization/maximization is taken over the unknown $\hat{\mathbf{I}}$. As only $\mathbf{I}$ is available, we need to approximate both α_o and β_o from $\mathbf{I}$. It is observed that when $\hat{\mathbf{I}}$ and $\mathbf{I}$ are close to each other, the order in the scale space (relative relationship between one DoG value and its surroundings) tends to be preserved, though the DoG values may change with large magnitudes.

More specifically, we can easily get

$$\mathbf{x}_{\min} = \arg\min_{\mathbf{x}\in\mathcal{S}_o\setminus\{\mathbf{k}_o\}} D_{\mathbf{I}}(\mathbf{x}), \tag{10}$$

in which the minimization is carried out in the available $\mathbf{I}$, rather than in the unknown $\hat{\mathbf{I}}$. Similarly, we can retain

$$\mathbf{x}_{\max} = \arg\max_{\mathbf{x}\in\mathcal{S}_o\setminus\{\mathbf{k}_o\}} D_{\mathbf{I}}(\mathbf{x}) \tag{11}$$

According to the aforementioned order-preserving strategy, we now can estimate α_o and β_o by

$$\hat{\alpha}_o = D_{\hat{\mathbf{I}}}(\mathbf{x}_{\min}) \tag{12}$$

$$\hat{\beta}_o = D_{\hat{\mathbf{I}}}(\mathbf{x}_{\max}) \tag{13}$$

where the indexes $\mathbf{x}_{\min}$ and $\mathbf{x}_{\max}$ given in (10) and (11) are derived from $\mathbf{I}$. It should be emphasized that here the subscript of the DoG function is $\hat{\mathbf{I}}$, instead of $\mathbf{I}$. In other words, we only exploit the relative order information, while not the exact DoG values of $\mathbf{I}$.

Therefore, the condition $(C.1)$ is chosen to be the relaxed version of (9), and can be written as

$$\hat{\alpha}_o \le D_{\hat{\mathbf{I}}}(\mathbf{k}_o) \le \hat{\beta}_o \tag{14}$$

As $\mathbf{I}$ is available, both $\mathbf{x}_{\min}$ and $\mathbf{x}_{\max}$ are fixed. Noticing that the DoG function $D_{\hat{\mathbf{I}}}(\mathbf{x})$ is linear with respect to $\hat{\mathbf{I}}$, the above constraint $(C.1)$ is also linear (and hence convex).

Let us further check the validity of ensuring that $\mathbf{k}_o$ is no longer a keypoint. When the condition (14) holds, at least one $\mathbf{x} \in \mathcal{S}_o \setminus \{\mathbf{k}_o\}$ exist such that $D_{\hat{\mathbf{I}}}(\mathbf{k}_o) \ge D_{\hat{\mathbf{I}}}(\mathbf{x})$. This implies that $\mathbf{k}_o$ is not a unique minimum. Similarly, when looking at the righthand side of the inequality (14), we conclude that at least one $\mathbf{x} \in \mathcal{S}_o\setminus\{\mathbf{k}_o\}$ exist such that $D_{\hat{\mathbf{I}}}(\mathbf{k}_o) \le D_{\hat{\mathbf{I}}}(\mathbf{x})$. This implies that $\mathbf{k}_o$ is not a unique maximum neither. Therefore, when (14) is satisfied, $\mathbf{k}_o$ will not be detected as a keypoint in the scale-space domain.

3.2. Determination of the condition $(C.2)$

Though the constraint $(C.1)$ eliminates the keypoint $\mathbf{k}_o$, there is no guarantee that new SIFT keypoints will not appear. Such new keypoint generation (NKG) problem is very serious because the new keypoints generated around the original one can still be matched with high probability. To solve this problem, we design the constraint $(C.2)$ to make sure that no SIFT keypoints exist in a local cuboid centered by $\mathbf{k}_o$ in the scale space, as shown in Fig. 1. Same with the setting of [1], we assume that there are 5 scales within each octave. The size of the cuboid is set to be $U \times U \times 3$, where $U = 7$ in our experiment. It should be noted that within the cuboid, there may exist some original SIFT keypoints, besides the newly generated ones. If this is the case, we also remove them by imposing the following condition $(C.2)$. More specifically, let us first construct the cuboid centered by $\mathbf{k}_o$

$$\begin{aligned}\mathcal{T}_o = \quad & \left\{(x,y,s)\middle| |x-x_o| \le \frac{U-1}{2}, |y-y_o| \le \frac{U-1}{2},\right. \\ & \left. |s-s_o| \le 1, x,y,s \in \mathcal{Z}\right\} \setminus \left\{\mathbf{k}_o\right\}\end{aligned} \tag{15}$$

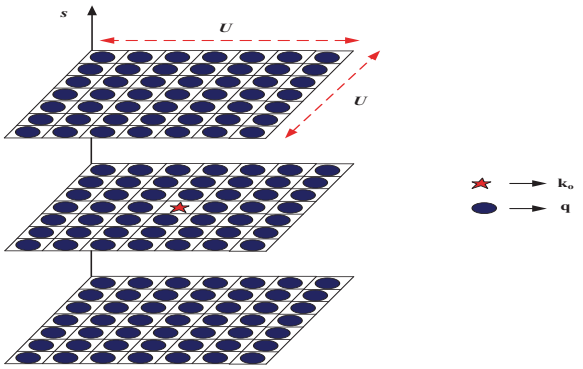


Fig. 1. The $U \times U \times 3$ sized cuboid centered by $\mathbf{k}_o$.

For each point $\mathbf{q} \in \mathcal{T}_o$, there are two possibilities: 1) it is not a keypoint in $\mathbf{I}$; and 2) it is a keypoint in $\mathbf{I}$. The two cases will be discussed separately below.

If Case 1 happens, then $\mathbf{q}$ is not an extremum in the $3 \times 3 \times 3$ cube $\mathcal{S}_q$ in the scale space centered by $\mathbf{q}$, where $\mathcal{S}_q$ can be similarly defined as (3). Let

$$\mathbf{x}^q_{\min} = \arg\min_{\mathbf{x}\in\mathcal{S}_q} D_{\mathbf{I}}(\mathbf{x}), \tag{16}$$

$$\mathbf{x}^q_{\max} = \arg\max_{\mathbf{x}\in\mathcal{S}_q} D_{\mathbf{I}}(\mathbf{x}) \tag{17}$$

Adopting a similar strategy of deriving (14), the condition to ensure that $\mathbf{q}$ will not become a new keypoint can be written as

$$D_{\hat{\mathbf{I}}}(\mathbf{x}^q_{\min}) \le D_{\hat{\mathbf{I}}}(\mathbf{q}) \le D_{\hat{\mathbf{I}}}(\mathbf{x}^q_{\max}) \tag{18}$$

Clearly, this constraint is convex with respect to $\hat{\mathbf{I}}$.

For the Case 2 where $\mathbf{q}$ itself is also a keypoint in $\mathbf{I}$, we need to exclude it from the set $\mathcal{S}_q$ when calculating $\mathbf{x}^q_{\min}$ and $\mathbf{x}^q_{\max}$ using (16) and (17). Otherwise, (18) is always satisfied, as one of the inequalities must hold with equality. More specifically, we now compute

$$\mathbf{x'}^q_{\min} = \arg\min_{\mathbf{x}\in\mathcal{S}_q\setminus\{\mathbf{q}\}} D_{\mathbf{I}}(\mathbf{x}), \tag{19}$$

$$\mathbf{x'}^q_{\max} = \arg\max_{\mathbf{x}\in\mathcal{S}_q\setminus\{\mathbf{q}\}} D_{\mathbf{I}}(\mathbf{x}) \tag{20}$$

In this case, the condition of making $\mathbf{q}$ non-keypoint can be expressed as

$$D_{\hat{\mathbf{I}}}(\mathbf{x'}^q_{\min}) \le D_{\hat{\mathbf{I}}}(\mathbf{q}) \le D_{\hat{\mathbf{I}}}(\mathbf{x'}^q_{\max}) \tag{21}$$

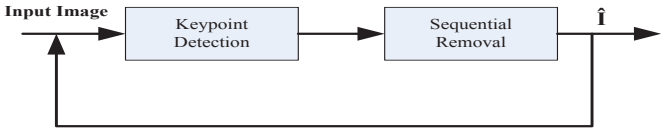


Fig. 2. Iterative process of removing SIFT keypoints

Incorporating all the conditions $(C.1)$ and $(C.2)$, we finally arrive at the convex optimization problem for removing the SIFT keypoints

$$\begin{aligned} &\min_{\hat{\mathbf{p}}_o} ||\mathbf{p}_o - \hat{\mathbf{p}}_o||_2^2 \\ \text{s.t} \quad &\hat{\alpha}_o \le D_{\hat{\mathbf{I}}}(\mathbf{k}_o) \le \hat{\beta}_o \\ &D_{\hat{\mathbf{I}}}(\mathbf{x}^q_{\min}) \le D_{\hat{\mathbf{I}}}(\mathbf{q}) \le D_{\hat{\mathbf{I}}}(\mathbf{x}^q_{\max}), \forall \mathbf{q} \in \mathcal{T}_o \bigcap \mathcal{K} \\ &D_{\hat{\mathbf{I}}}(\mathbf{x'}^q_{\min}) \le D_{\hat{\mathbf{I}}}(\mathbf{q}) \le D_{\hat{\mathbf{I}}}(\mathbf{x'}^q_{\max}), \forall \mathbf{q} \in \mathcal{T}_o \bigcap \mathcal{K}^c \end{aligned} \tag{22}$$

where

$$\mathcal{K} = \Big\{(x, y, s)\Big|(x, y, s) \text{ is a keypoint of } \mathbf{I}\Big\} \tag{23}$$

and $\hat{\mathbf{p}}_o = \mathbf{E}_o \circ \hat{\mathbf{I}}$.

3.3. Iterative process of removing SIFT keypoints

In the above two subsections, we have presented our approach for removing one SIFT keypoint $\mathbf{k}_o$, and prohibit the existence of any SIFT keypoints in a local cuboid centered by $\mathbf{k}_o$. We can sequentially apply such removal operations to process all the SIFT keypoints. Nevertheless, after one round of processing, there may still exist many unremoved SIFT keypoints. This is because we can only ensure that no keypoints exist in a local cuboid of size $U \times U \times 3$. While outside that cuboid, no guarantee can be provided. More seriously, when multiple keypoints are tightly clustered, different patches $\hat{\mathbf{p}}_o$'s may be overlapped spatially. The interference among different patches could also generate some new SIFT keypoints.

To solve these two problems, we propose an iterative strategy, as illustrated in Fig. 2. In each round t, we sequentially process all the identified SIFT keypoints, and eventually get the resulting image $\hat{\mathbf{I}}_t$. In the next round, we use $\hat{\mathbf{I}}_t$ as input, and apply the removal operations for all the identified SIFT keypoints again. Such process will be iteratively performed for several rounds until the KRR is desired or the iteration number exceeds a threshold.

3.4. Remarks on the comparison with [8]

In this subsection, we give a brief comparison of our proposed SIFT keypoint removal method with the one in [8], in which a constrained optimization framework was also developed. As

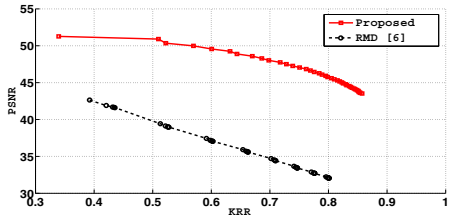
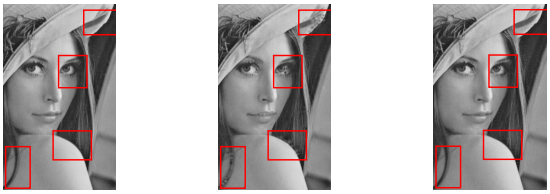


Fig. 3. Comparison with RMD method [6] in terms of averaged KRR-D performance.



(a) Original (b) [6] (c) Proposed

Fig. 4. Visual quality comparison with RMD method [6]. (a) original, (b) RMD (KRR 77.18%, PSNR 37.75 dB), and (c) Proposed (KRR 83.08%, PSNR 45.10 dB)

the objective functions in both methods are straightforward ℓ_2 distortion terms, the key differences lie in the constraints.

In [8], the constraint to suppress the SIFT keypoint is to make the minimum point and the second minimum point equal, and hence generating two minima in the same $3 \times 3 \times 3$ cube. This condition can be regarded as a special case of our constraint $(C.1)$ given in (14) when the left inequality is satisfied with equality. Certainly, the solution space of the framework in [8] is much restricted, which would lead to much more severe distortions. The results in the next section will verify our conclusion experimentally.

Furthermore, the strategy to prevent from generating new keypoints in [8] is to force all the points in the $3 \times 3 \times 3$ cube, except $\mathbf{k}_o$ to be no smaller than $D_{\hat{\mathbf{I}}}(\mathbf{k}_o)$. However, this cannot fully solve the NKG problem. To determine whether a point in the scale space is a keypoint or not, we should compare all the 27 points in the cube centered by itself, rather than by $\mathbf{k}_o$. In contrast, when designing the condition $(C.2)$, we can ensure that no SIFT keypoints exist in a $U \times U \times 3$ cuboid, where U is a parameter balancing the removal performance and incurred distortion.

4. EXPERIMENTAL RESULTS

In this section, we evaluate the performance of our proposed SIFT keypoint removal algorithm in terms of KRR-D metric, and further demonstrate its effectiveness through a case study of defeating the copy-move forgery detection system. More specifically, the KRR, denoted by τ, is defined as

$$\tau = 1 - \frac{\#\text{ correctly matched keypoints after removal}}{\#\text{ number of original keypoints}} \quad (24)$$

We first compare the KRR-D performance of our proposed scheme with the state-of-the-art RMD method [6]. The experiments are conducted over the UCID image database [9] consisting of 1338 uncompressed images with various characteristics. All the images are converted into gray scale, and the SIFT keypoints are extracted using the popular open-source SIFT-VLFEAT [10] with peak threshold=4 and the edge threshold=12. A can be seen from Fig. 3, the KRR-D performance of our approach is significantly better than that of RMD [6]. The PSNR gain of our method over RMD becomes increasingly larger as KRR improves. For instance, the gain can be over 12 dB when KRR=0.8, which is quite remarkable. In addition to the KRR-D curves, we also present the visual quality comparison in Fig. 4 between these two methods. Our proposed technique achieves more visually pleasing result than that of the RMD approach.

We also perform the comparison among our method, RMD [6], and the one in [8] for 8 standard test images in Fig. 5. As the source code of [8] is not available, the results represented by the dotted lines ($\triangle$) are extracted from Table 1 of [8]. Our method outperforms both [6] and [8] by a big margin. When the KRR is low, the PSNR gap between our solution and [8] is relatively small, because the distortion difference is less if the number of removed SIFT keypoints becomes lower. Nevertheless, for large KRR, our approach achieves very remarkable improvement. This experimentally explains the importance of appropriately determining the constraints in the optimization framework, which is one of our major contributions in this work.

Furthermore, we demonstrate that our SIFT keypoint removal algorithm is effective in defeating a SIFT-based image copy-move forgery detection system [2], while still maintaining high quality of the image. Copy-move forgery is a popular means of making doctored images by cloning an area of the image onto another zone, probably accompanying with some appropriate geometric transformations. The state-of-the-art technique [2] has been proven to be quite effective in robustly detecting cloned region via SIFT matching. To malfunction the detection system, we first perform the copy-move forgery on 10 images randomly selected from UCID database. Some appropriate geometric transformations are conducted such that the forged images are visually more realistic and better fit the background. The 10 images and their forged versions are available `https://www.dropbox.com/sh/tm7tqfs5d8bmizp/AABy9ewmdPFSofsqQsGYkrqVa`. We apply our removal technique and the RMD [6] on the cloned region respectively. Fig. 6 shows the result for one of the images, in which a bear head is copied to cover another one. Before the removal, the clones region can be successfully detected by the approach

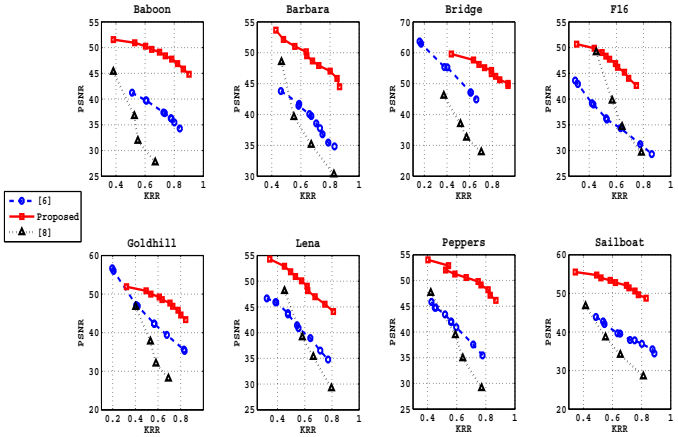
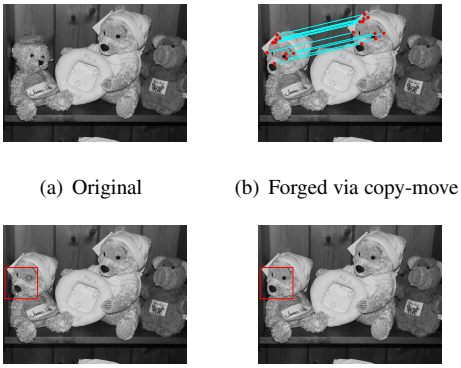


Fig. 5. Comparison with [6] and [8] in terms of KRR-D performance for 8 test images.



(a) Original (b) Forged via copy-move

(c) [6] (d) Proposed

Fig. 6. Before and after the SIFT keypoint removal.

of [2], through SIFT matching. After removing the SIFT keypoints, the number of matched SIFT descriptors drops to 0, meaning that both approaches are effective in defeating the forgery detection system [2]. However, the distortion incurred by the RMD method is much more severe and noticeable (see e.g., the bear's eyes enclosed in the red box).

5. CONCLUSION

In this work, we have investigated the security of SIFT against malicious attack. We have designed an effective strategy to remove the SIFT keypoints while still maintaining high quality of the image. This has been cast as a constrained optimization problem, where the constraints are well-designed to suppress the existence of local extremum and prevent generating new keypoints in a local cuboid in the scale space. Unfortunately, the optimization problem in the ideal case has been shown to be non-convex. To tackle such challenge, we have proposed a convex relaxation technique to approximate the original problem. Experimental results, including KRR-D performance comparison and a case study of malfunctioning a copy-move forgery detection, have been provided to demonstrate the superior performance of our scheme.

An implication of our results is that we cannot fully trust the image input to the SIFT-based system, especially those security systems. An authentication mechanism is needed to verify the data validity, before they are sent to SIFT extraction module. Otherwise, the decision made upon the extracted SIFT features could be groundless.

6. REFERENCES

[1] D.G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. of Comp. Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[2] I. Amerini, L. Ballan, R. Caldelli, A. Del Bimbo, and G. Serra, "A sift-based forensic method for copy–move attack detection and transformation recovery," *IEEE Trans. on Inf. Forensics and Security*, vol. 6, no. 3, pp. 1099–1110, 2011.

[3] H. Lejsek, F.H. Asmundsson, B.T. Jonsson, and L. Amsaleg, "NV-Tree: An efficient disk-based index for approximate search in very large high-dimensional collections," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 31, no. 5, pp. 869–883, 2009.

[4] C.Y. Hsu, C.S. Lu, and S.C. Pei, "Secure and robust SIFT," *Proc. ACM MM*, pp. 637–640, 2009.

[5] T.T. Do, E. Kijak, T. Furon, and L. Amsaleg, "Understanding the security and robustness of SIFT," *Proc. ACM MM*, pp. 1195–1198, 2010.

[6] T.T. Do, E. Kijak, T. Furon, and L. Amsaleg, "Deluding image recognition in SIFT-based CBIR systems," *Proc. ACM workshop on Multimedia in Forensics, Security and Intell.*, pp. 7–12, 2010.

[7] I. Amerini, M. Barni, R. Caldelli, and A. Costanzo, "Counter-forensics of SIFT-based copy-move detection by means of keypoint classification," *EURASIP J. on Image and Video Proc.*, vol. 2013, no. 1, pp. 1–17, 2013.

[8] C.S. Lu and C.Y. Hsu, "Constraint-optimized keypoint inhibition/insertion attack: security threat to scale-space image feature extraction," *Proc. ACM MM*, pp. 629–638, 2012.

[9] G. Schaefer and M. Stich, "UCID: an uncompressed color image database," 2003.

[10] A. Vedaldi and B. Fulkerson, "VLFeat: An open and portable library of computer vision algorithms," 2012.