Tracking-Tolerant Visual Cryptography

Ruofei Du*, Eric Lee†, and Amitabh Varshney‡ *Fellow, IEEE*

Augmentarium, Department of Computer Science, and University of Maryland Institute for Advanced Computer Studies
University of Maryland, College Park

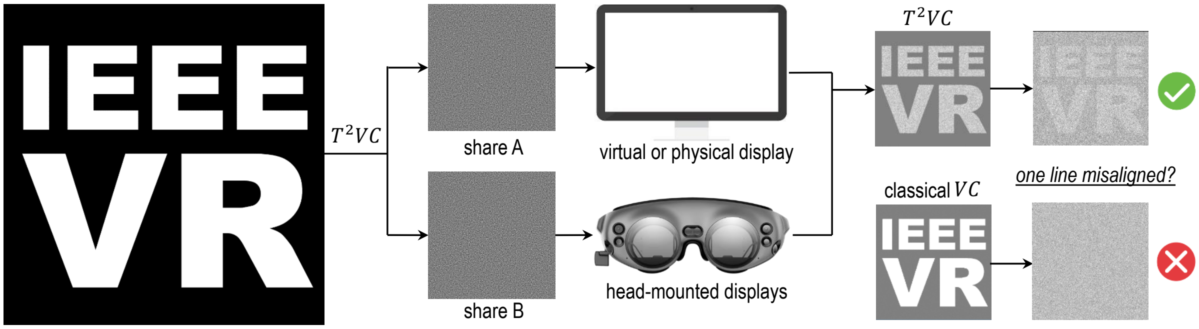


Figure 1: Results and overview of our system, T^2VC, which is able to split a confidential message into two shares of images and guarantees that the original information could not be revealed with either share of image alone. When the user looks through the two aligned images in the head-mounted display, the secret message is revealed directly to the user's human visual system. Nevertheless, head jittering may cause the two images to slightly misalign with each other even with visual tracking algorithms. Our algorithm outperforms the classic visual cryptography algorithm in presence of one or two rows of misalignment.

Abstract

We introduce a novel secure display system, which uses visual cryptography [4] with tolerance for tracking. Our system brings cryptographic privacy from text to virtual worlds [3]. Much like traditional encryption that uses a public key and a private key, our system uses two images that are both necessary for visual decryption of the data. The public image could be widely shared on a printed page, on a traditional display (desktop, tablet, or smartphone), or in a multi-participant virtual world, while the other private image can be exclusively on a user's personal AR or VR display. Only the recipient is able to visually decrypt the data by fusing both images. In contrast to prior art, our system is able to provide tracking tolerance, making it more practically usable in modern VR and AR systems. We model the probability of misalignment caused by head or body jitter as a Gaussian distribution. Our algorithm diffuses the second image using the normalized probabilities, thus enabling the visual cryptography to be tolerant of alignment errors due to tracking.

Keywords: visual cryptography, augmented reality (AR), tracking

Index Terms: H.5.1 [Information Interfaces and Presentation (e.g., HCI)]: Multimedia Information Systems—Artificial, augmented, and virtual realities I.3.3 [Computer Graphics]: Picture/Image Generation—Display algorithms

1 Introduction

We present T^2VC, a tracking-tolerant visual cryptography system for AR or VR head-mounted displays (HMDs). Our system presents a practical and robust cryptographic solution that eliminates every device from the trusted computing bases (TCB) and assumes no connection between the TCBs. First, T^2VC splits the confidential information (as an image) into two shares. One share of data is displayed on the ordinary screen while the other share of data is displayed on the HMD. The user decrypts the message by visually aligning the two shares of information. Our work is built upon the pioneering research by Andrabi *et al.* [1], which first demonstrates the potential of using AR HMD to reveal secret messages using the visual cryptography system induced by Naor and Shamir [4]. However, their system requires a chinrest and takes over ten seconds for the users to recognize a single character. This is largely due to head jitters when manually aligning the two images.

To solve the challenge of head jittering, T^2VC leverages the visual tracking modules in *Magic Leap One*[1]. While visual tracking algorithms may roughly align two images together, they may still suffer from one or two pixels of misalignment. Our system further models the misalignment of head jitter using a 2D Gaussian distribution. We have developed a novel algorithm to enhance the visibility of the classical visual cryptography via diffusion with Gaussian kernels, thus enabling the algorithm to be tolerant with misalignment.

2 Algorithm

The main idea behind T^2VC is: for each pixel p in one share, we model the probability of misalignment on another pixel q as a 2D Gaussian distribution centered at the pixel p. In this way, we sacrifice a little contrast in the fused result for better clarity when one or two rows of misalignment occurs.

2.1 Preprocessing

Following [1] and [4], given a confidential visual image $\boldsymbol{I}$, we first generate a binary image $\hat{\boldsymbol{I}}$ by thesholding every 2×2 block of pixels in $\boldsymbol{I}$. Here, we denote $\mathscr{F}(\hat{\boldsymbol{I}})$ and $\mathscr{B}(\hat{\boldsymbol{I}})$ as the set of foreground (white) and background (black) pixels of $\hat{\boldsymbol{I}}$, respectively.

Next, we model the range of misalignment as an $s \times s$ square and generate an $s \times s$ 2D Gaussian kernel $\mathscr{G}(x,y,\sigma)$ at scale σ:

$$\mathscr{G}(x,y,\sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \qquad (1)$$

In our experiments, we choose $s = 3, \sigma = 1.0$ and $s = 5, \sigma = 2.0$.

*e-mail: me@duruofei.com, now at Google LLC, San Francisco.
†e-mail: ericlee@umiacs.umd.edu
‡e-mail: varshney@cs.umd.edu

[1] Magic Leap One: `https://magicleap.com/magic-leap-one`

ALGORITHM 1: Tracking-Tolerant Visual Cryptography

```
Input: a binary secret image $\hat{I}$
Output: two shares of information $I^{\alpha}$ and $I^{\beta}$
Generate a random share of $I^{\alpha}$;
for each $2 \times 2$ block $b_{rc}$ of $\hat{I}$ do
    for each $2 \times 2$ block $b_{ij}$ of $\hat{I}$, where $|r-i| \leq \frac{s}{2}, |c-j| \leq \frac{s}{2}$ do
        if $b_{ij} \in \mathscr{F}(IB)$ or AllowBackgroundDiffusion then
            Look up the probability of misaligning $b_{rc}$ with
              $b_{ij}$ from the from the Gaussian kernel $\mathscr{G}(x,y,\sigma)$:
              $\mathscr{P}(b_{rc}, b_{ij}) \leftarrow \mathscr{G}(r-i, c-j, \sigma)$ ;
            Increase the normalization factor of $b_{rc}$:
              $\mathscr{N}_{rc} \leftarrow \mathscr{N}_{rc} + \mathscr{P}(b_{rc}, b_{ij})$;
        end
    end
    Normalize probabilities: $\mathscr{P}(b_{rc}, b_{ij}) \leftarrow \mathscr{P}(b_{rc}, b_{ij}) / \mathscr{N}_{rc}$;
    Generate a random uniform sample: $r \in [0,1]$;
    Set the accumulated probabilities: $\mathscr{A}_{rc} \leftarrow 0$
    for each $2 \times 2$ block $b_{ij}$ of $\hat{I}$, where $|r-i| \leq \frac{s}{2}, |c-j| \leq \frac{s}{2}$ do
        if $b_{ij} \in \mathscr{F}(IB)$ or AllowBackgroundDiffusion then
            $\mathscr{A}_{rc} \leftarrow \mathscr{A}_{rc} + \mathscr{P}(b_{rc}, b_{ij})$;
            if $r \leq \mathscr{A}_{rc}$ then
                $(p,q) \leftarrow (i,j)$   break;
            end
        end
    end
    $I^{\beta}_{rc} \leftarrow b_{pq} \in \mathscr{B}(\hat{I})$ ? $I^{\alpha}_{pq}$ : $WHITE - I^{\alpha}_{pq}$ ;
end
```

2.2 Generation of Two Shares

T^2VC generates the first share as the classical VC approach. For each 2×2 block of pixels in the first share, we randomly choose one of the six VC patterns. Next, we carry out two solutions to deal with the possible misalignment: 1) T^2VC^*: for the second share, we only diffuse the foreground pixels: each foreground pixel has a probability to be misaligned with one of its surrounding pixels; in this way, when the two shares match perfectly, the background is unchanged, but the foreground becomes darker. 2) T^2VC: for the second share, we diffuse both the background and foreground pixels to enhance the contrast: every pixel has a probability to be misaligned with one of its surrounding pixels. Please refer to the source code is provided in the *supplementary material*.

3 Experimental Results

To valid the effectiveness of our algorithm, we conduct preliminary experiments via both simulation and physical deployment.

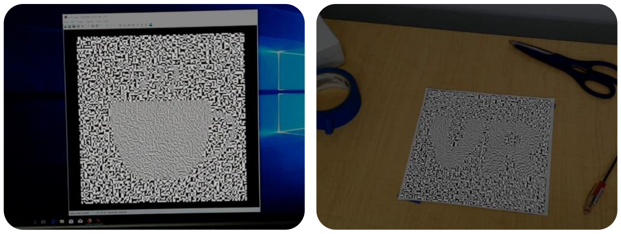
3.1 Comparison with Classical Visual Cryptography

We generate visual cryptography images at the resolution of 1024×1024 pixels using a custom C++ program using the proposed T^2VC algorithms and the classical visual cryptography algorithm under four conditions: exact match, one-row misalignment, one-column misalignment, and one-row, one-column misalignment (please refer to the supplementary material and [2] for more detail). We summarize the following insights:

1. The classical visual cryptography algorithm does not work with even a single row or column of misalignment, making it extremely challenging to interpret the image with the visual tracking being even slightly off.
2. T^2VC^* can deal with one row or one column misalignment (2 pixels) while preserving as good a contrast as the original visual cryptography algorithm. However, the contrast drops with misalignment.
3. T^2VC provides better contrast than T^2VC^* when misalignment occurs and even works with two pixels misaligned both horizontally and vertically. After increasing the size and scale of the Gaussian kernel, we can still see the secret message even with two rows (four pixels) of misalignment.

3.2 Deployment

We implement our system in Unity and deploy it on *Magic Leap One*. As shown in Fig. 2, the user can still observe the decrypted information even when the visual tracking module of Magic Leap One misaligns the two shares. In the supplementary material, we further suggest smoothly varying the brightness level of the overlaid image.



(a) revealing a teacup icon with one share displayed on a desktop monitor

(b) revealing two characters *VR* with one share printed on a piece of paper

Figure 2: Results of seeing through *Magic Leap One to* align the other share (a) in a desktop monitor, and (b) on a piece of paper.

4 Conclusion

In this paper, we have adapted visual cryptography for the current-generation AR HMDs. Our system T^2VC uses a novel visual cryptography algorithm which is tolerant to users' head jitter and slight misalignment of the two shares of encrypted visual information when visual tracking is enabled. We achieve this by modeling the misalignment through a 2D Gaussian distribution of the visual cryptography's random patterns. This allows us to trade off precise alignment with perceived contrast. As one of the first steps towards practical visual cryptography for VR and AR, we believe that our algorithm provides a versatile, commodity, off-the-shelf solution for embedding encrypted augmented reality information in the real-world displays and virtual environments [3], thereby protecting confidential data while facilitating an easy-to-use visual decryption.

Acknowledgement

This work has been supported in part by the NSF Grants 1823321, 1564212, 1429404, and the State of Maryland's MPower initiative.

References

[1] S. J. Andrabi, M. K. Reiter, and C. Sturton. Usability of Augmented Reality for Revealing Secret Messages to Users But Not Their Devices. In *11th Symposium on Usable Privacy and Security*, pp. 89–102, 2015.

[2] R. Du. *Fusing Multimedia Data Into Dynamic Virtual Environments*. PhD thesis, University of Maryland, College Park, Nov. 2018.

[3] R. Du, D. Li, and A. Varshney. Geollery: a Mixed Reality Social Media Platform. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI)*, CHI, p. 13. ACM, May. 2019. doi: 10.1145/3290605.3300915

[4] M. Naor and A. Shamir. Visual Cryptography. In *Workshop on the Theory and Application of of Cryptographic Techniques*, pp. 1–12. Springer, 1994. doi: 10.1109/ICRITO.2016.7784984