RETHINKING CONVENTIONAL DATA HIDING RESEARCH IN DIGITAL IMAGES: THE CHALLENGES AND THE PROPOSED SOLUTIONS

Simying Ong
Faculty of Computer Science and Information Technology,
Universiti Malaya, Malaysia.

When we take a step back and rethink about the current state and framework of data hiding research, what's the problem and what's left to be done here?

I. The Introduction

Data Hiding is the art and science of concealing the existence of data for various purposes and needs in different application domains. It can be known as steganography, for achieving the purpose of covert communication, or it can also be known as watermark, for claiming the ownership of created digital content, or it can also be called as metadata insertion, for inserting metadata into the digital medium for management and retrieval purposes, just to name a few. Different purpose and need require diverse levels of concealment (security), processing complexity and time, output quality, and hiding capacity. For many years, countless research in data hiding have been done to either focus on certain quality depending on the application domain or trying to strike a trade-off balance between these qualities. Novel techniques or improved techniques are constantly invented to be crowned as the winner of the quality game.

However, when we take a step back and rethink about the current state and framework of data hiding research, what's the problem and what's left to be done here? In this article, the stake and challenges of the current data hiding research are discussed from the technical framework and usability perspectives. Then some preliminary works done in the research group are also shared in this article hoping to get nearer to the mature solution in the future.

II. Must we follow the "Conventional Framework"?

A standard data hiding encoding framework is as shown in Fig. 1, consist of four compulsory components, namely the cover image, the hidden data, the data hiding technique and the output image. The cover image will be modified using the data hiding technique, to accommodate the data and generate an output image with hidden data. In most cases, maintaining the output image quality is utmost important, because the output image will be

stored, displayed, or transferred to the intended user or location for further usage and application.

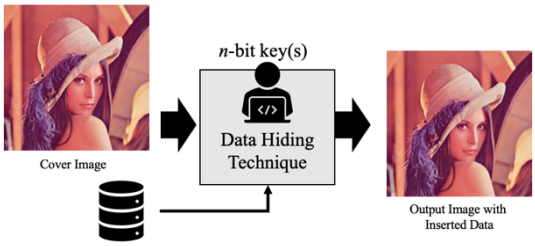

Figure 1: The Conventional Data Hiding Framework.

At the same time, malicious actions can also be performed on the output image to identify the data hiding technique, detect any hidden data, or even crack to obtain the hidden data. These attacks including analyzing its statistical properties such as the signal or histogram distribution against the usual natural image distribution, comparing the output image with its original counterpart, etc. There are many possible attacks centered around the "output image" by trying to find out the trace of modification due to the hiding gesture.

To combat the attacks, there are also researches which proposing data hiding technique that focus on maintaining the natural image statistical distribution or minimizing the modification on the cover image. However, there is no easy way out for this too due to the needs in achieving better results in other qualities and constrained by the cover image. In fact, the cover image acts as a container that creates a boundary to data hiding technique, including the hiding capacity, the distribution, etc. For instance, some data hiding techniques failed to conceal data or only able to achieve very low hiding capacity in certain image type or image components, and it is impossible to hide "outside" of the cover image

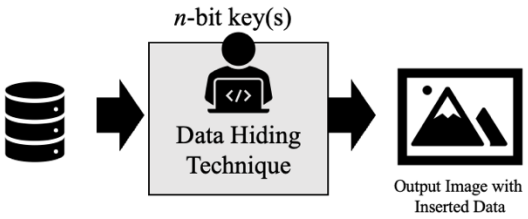

Figure 2: The Coverless Data Hiding Framework.

Hence, is there any other option? Must we use "cover image" and the standard "output image" production method in conventional encoding framework? In year 2007, Otori and Kuriyama [1] proposed a data hiding technique using texture synthesis method by encoding secrets into Local Binary Pattern dots and placed onto a blank image, which later camouflaged as a random texture image. Following this research, several synthesis-based methods have been proposed to generate output image without using "cover image" [2-3]. Here, we started to coin these methods as the "coverless data hiding methods" [4]. For more information about the evolutions and reviews of coverless data hiding methods, kindly refer to [4] for further reading.

There are two categories of coverless data hiding methods, namely the constructive/ synthesis-based, and the non-constructive-based. Data hiding methods fall under the coverless range are of those which do not modify the cover image to conceal data. Therefore, for non-constructive-based methods [5-6], input image(s) will still be needed, but only to select suitable images to represent the hidden data. These methods need a large image pool, then using certain calculation technique, such as the average pixel intensity, grayscale gradient co-occurrence matrix, to calculate the represented value of each image. A group of images will be selected in sequence by matching the represented value with the hidden data. The non-constructive-based methods can retain the image quality, but the hiding capacity is very low, and the sequence of the images must be maintained to correctly decode the hidden data.

As for constructive/synthesis-based data hiding methods, no "cover image" or "input image" is required. The "output image" is constructed by representing the data using points [1], patterns [7-8], features [3], or generated images [9]. In 2021, Ng et al. [9] proposed a coverless constructive-based data hiding framework (refer to Fig. 3), which generates art images using ACGAN (Auxiliary Classifier Generative Adversarial Network), mapping each art class to represent specific hidden binary

value, then collage the art images and mask it to generate the final image as shown in Fig. 4 with hidden data.

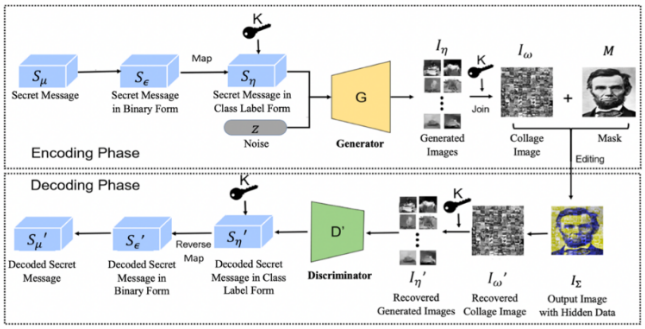

Figure 3: The coverless data hiding method using collage image [9].

Figure 4: Output image using collage image and mask [9].

All these coverless methods are proposed and tested with various experiments to evaluate them with the conventional data hiding methods. At the current stage of work and advancement of technology, it is found that coverless data hiding methods are much more possible and feasible as compared to decade ago. Using coverless approach shows that these methods not only able to avoid the attacks and limitations that applied to the conventional approaches, but it also opens another possible path in achieving various qualities in data hiding.

III The Usability

The usability is an important requirement when research is put into action and use in real life. For data hiding research, one of the biggest challenges now is to fill the real-world gap by encouraging the use of data hiding methods in the actual world. In every day, new data hiding techniques are proposed to improve the technical qualities, but these methods are simply proposed, experimented, and ended up just being archived in the publisher's database.

A preliminary study was performed on a group of university students to understand the possible causes of this gap [10-11]. In this study, it is observed that the user agreed that the importance of data hiding in protecting their digital images but feels cumbersome to perform additional steps in applying data hiding methods on their digital images

Let's imagine the user routine from capturing the image to uploading the final image to social networking site (SNS) in a mobile phone setting. First, user capture the image using camera, the light signal will be captured and processed to the required image format, enhanced to increase the image clarity and features. Then, user apply photo effects such as filter, border, text or adding objects to customize and personalize the image. Lastly, the image is uploaded to SNS and reformat or compressed to be stored in the database.

Considering the user routine, many image processing steps are performed in the process (refer to the italic words in the routine). Therefore, it is certainly possible to integrate data hiding into these image processing step(s) to reduce the hassle of doing additional steps to achieve data hiding. In our research group, several preliminary works have been proposed to test the feasibility of the proposed integration, and these works are broadly divided into two categories, data hiding in image enhancement techniques [13] and data hiding in photo effects [10-12].

In year 2020, we proposed integrating data hiding using Median Filter image enhancement technique [13]. Instead of choosing the median value to be replaced with the targeted pixel, the pixels in the pre-defined sized region are grouped into multiple partitions to carry distinctive binary values. The hidden data will decide which pixel partition that will be used for replacing targeted pixel. In 2020, William et al. [12] proposed to use Sketch, Halftone and Vintage effects, as shown in Fig. 5, to hide data. Later in 2021, Gong [10] and Tan [11] proposed to embed data using Mosaic filter, pencil drawing, pixel extension and collage effects. In addition, these new data hiding methods are put into a mobile application (refer to Fig. 6) as a proof of concept. Currently, these are the on-going works which later will be put into the feasibility and

usability experiments to verify their performances. More detailed results will be reported in the future.

Figure 5: Data hiding using photo effects [12].

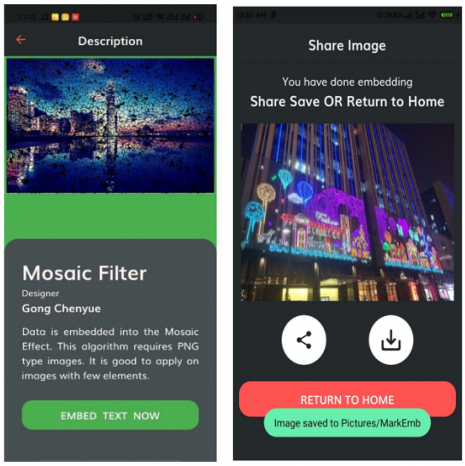

Figure 6: Mobile App prototype for Data Hiding using Photo Effects [10-11].

IV The Conclusion

In this article, two challenges in current data hiding research, viz., the limitation of using "cover image" in conventional framework, and the usability gap within data hiding research and real-world, are briefly explained. For the first challenge, some existing coverless data hiding methods are brought forward as possible solutions, because the coverless approach can avoid leaving modification traces on the cover image. As for the second challenge, data hiding techniques are proposed to integrate the hiding capability with the image processing step in user routine, including image enhancement and photo filters, to eliminate the extra steps required in using data hiding in real life scenario.

IV Acknowledgement

Special thanks to all the student researchers, including Ms. Ng Koi Yee (PhD student), Mr. William Kan Wei Yeong (Master student – graduated year 2022), Mr. Gong Chenyue and Mr. Tan Qing Lin (Final Year Project students – graduated year 2021) for investigating into these challenges, proposing potential solutions, and providing valuable insights/findings to these issues.

References

[1] H. Otori and S. Kuriyama, "Data-embeddable texture synthesis," in International Symposium on Smart Graph- ics. Springer, 2007, pp. 146–157.
[2] K.-C. Wu and C.-M. Wang, "Steganography using reversible texture synthesis," IEEE Transactions on Image Processing, vol. 24, no. 1, pp. 130–139, 2014.
[3] S. Li and X. Zhang, "Toward construction-based data hiding: from secrets to fingerprint images," IEEE Trans- actions on Image Processing, vol. 28, no. 3, pp. 1482– 1497, 2019.
[4] K. Y. Ng, S. Ong and K. Wong, "Delving into the Methods of Coverless Image Steganography," 2019 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), 2019, pp. 1763-1772.
[5] L. Zou, J. Sun, M. Gao, W. Wan, and B. B. Gupta, "A novel coverless information hiding method based on the average pixel value of the sub-images," Multimedia Tools and Applications, pp. 1–16, 2018.
[6] J. Wu, Y. Liu, Z. Dai, Z. Kang, S. Rahbar, and Y. Jia, "A coverless information hiding algorithm based on grayscale gradient co-occurrence matrix," IETE Techni- cal Review, vol. 35, no. sup1, pp. 23–33, 2018.
[7] W. K. Lee, S. Ong, K. Wong and K. Tanaka, "A Novel Coverless Information Hiding Technique Using Pattern Image Synthesis," 2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), 2018, pp. 1122-1127.
[8] Z. Qian, H. Zhou, W. Zhang, and X. Zhang, "Robust steganography using texture synthesis," in Advances in Intelligent Information Hiding and Multimedia Signal Processing. Springer, 2017, pp. 25–33.
[9] K. Y. Ng, S. Ong, Y. P. Loh and C. S. Chan, "Relabel, Scramble, Synthesize:A Novel Coverless Steganography Approach via Collage Image," 2021 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), 2021, pp. 1877-1882.
[10]Chenyue Gong, 2021, "Aesthetic Photo Effects Implementation via Mobile Application in Enabling Information Hiding (Mosaic Filter and Pencil Drawing)" [Universiti Malaya Final Year Project Report – to be published].
[11] Qing Lin Tan, 2021, "Aesthetic Photo Effects Implementation via Mobile Application in Enabling Information Hiding (Pixel Extension and Collage Effects)" [Universiti Malaya Final Year Project Report – to be published].
[12] W. K. W. Yeong, S. Ong and K. Wong, "Data Embedding Method Using Photo Effects with Resistance to Compression," 2020 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), 2020, pp. 1361-1368.
[13] S. Ong and K. Wong, "Information Hiding In Image Enhancement," 2020 IEEE International Conference on Image Processing (ICIP), 2020, pp. 1261-1265.