Encrypt Microarray Image with Your Pictures

Hsi-Chun Wang [1], Chu-Song Chen [2,1], and Chi-Wei Jan [2]

[1] Department of Graphic Arts and Communication, College of Science and Technology, National Taiwan Normal University, Taipei, Taiwan
[2] Institute of Information Science, Academia Sinica, Taipei, Taiwan

Abstract

The objective of this research is to apply the data hiding technique to encrypt micro-array image with the help of contrast properties of a portrait image or company logo. The original micro-array image is captured and digitized from a gene-chip membrane, with 9600 different hybridized and expressed gene fragments, as 9600 features are extracted by image analysis. By adopting an algorithm of "relative-contrast-consistent reshuffle," the features of the micro-array image and an 80x120 grayscale image can result a new image which possesses the features of micro-array's gene expression and a meaningful pattern (e.g., a human portrait or a company logo) as well. The method developed in this research has significant potential in personal identification, security information protection and next generation gene-chip design.
Keyword: digital image processing; data hiding; cDNA microarray

摘要

本研究之目的在於將生物晶片之特徵影像應用另一張數位影像的反差特性來隱藏訊息，所得之結果影像除了仍保有原生物晶片影像之特徵值之外，並能顯現有意義之圖案。本研究之原始數位影像來自於數位化的9600個基因微陣列點顯色影像，將此數位影像再經影像分析後，得到各個基因片段表現顯色點的特徵值，此9600個特徵值與另一80x120灰階影像，經由本研究所發展之「相對反差不變之重整」演算法處理後，可將生物晶片影像之特徵值以有意義之圖案表示。本研究之結果顯示，所發展之方法能成功地顯示不同的人臉或公司行號的識別標誌影像，未來並有極大的發展潛力應用於身分認證或有價證券防偽與圖像化的生物晶片設計。

1. Introduction

In recent years, accompany with rapid innovation of cutting-edge computer technology, multimedia and network, researches on information hiding or watermarking in images, audio and video have drawn large attention of scientists and engineers in related fields [1,2,3,4]. Most of the works have focused on data hiding in electronic data format for media applications. Due to the development of gene-chip technology, digital image processing plays a key role in the quantification of gene expression analysis. Data hiding technique for gene-chip micro-array image therefore has great potential in personal identification and security information preparation, and may effectively help us to strike against unauthorized usage and illegal counterfeiting [2].

Micro-array gene chip system possesses the ability to do the quantification of gene expression and make speedy gene screening. Microarray, which consists of thousands of gene fragments, can be prepared by several methods, such as contact printing, inkjet printing and photolithography on the substrates [5,6,7,8,9,10,14]. After the hybridization with tagged DNA from normal or treated cells, the hybridized micro-array is transferred to different detection procedures for the quantification of gene expression. There are two major detection systems for microarray systems that are distinguished by varied tagged agents: laser-induced fluorescence detection [6,7] and enzyme-linked colorimetry detection [5]. In general, the former has been widely adopted more by research institutes. As for the later, which can be performed even by a flatbed scanner, it has high throughput and can be more cost-effective for saving the delicate laser-induced fluorescence detection equipment. Both methods involve specific bio-chemical reactions and electro-mechanical control modulus. The micro-array image is captured by a photo-sensor, which is either an arrayed CCD or a scanning photomultiplier tube. The digitized image is then sent to a microcomputer for further analysis. [12,13]. Observing a typical micro-array image by colorimetry detection in Figure 1, 9600 gene-fragments are contact printed 1600 times (in 40x40 format) by 6 steel pins which deliver the droplets with different gene sequences on a nylon membrane, and the spots later show varied gradation after the hybridization reaction. The varied gradation is related to the so-called gene expression, and the gradation of those dots (in a matrix of 80x120) look like randomly dispersed. An image analysis algorithm is applied to extract the features of those gene sequences according to the location and local intensity for each dot. The features referred here are the quantification values for the gene expression.
The objective of this research is to apply the data hiding technique to encrypt micro-array image with the help of a picture or a company logo. The resulted image will possess both the features of micro-array's gene expression and a meaningful pattern.

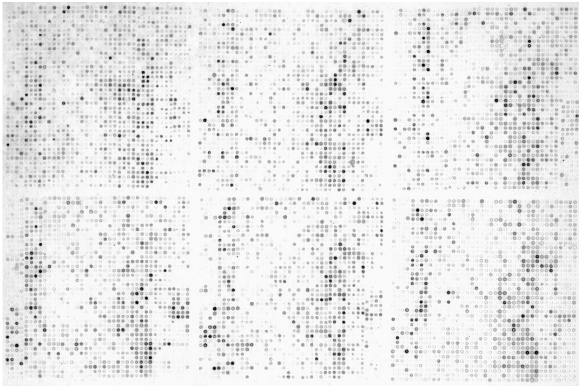

Figure 1. A typical micro-array gene chip image

2. Face Picture-Guided Reshuffle of Micro-array Image

In this section, we will introduce the main principle of our method that can successfully encrypt a microarray image with the aid of a particular picture. Without losing the generality, we assume in this research that the particular picture is taken from a single human face. Our method however allows it to be an arbitrary picture which does not have to be restricted to be a human face. Through our method, the encrypted microarray images not only look like a picture of this human face, but they also inherently hides and preserves complete information about the microarray images. The proposed method is called the *relative-contrast-consistent reshuffle* in this paper, which will be introduced as followed.

Consider a pair of digital gray-level images (A, B), where A is a microarray image and B is an arbitrary picture. Without lost of generality, assume that the widths (or heights) of them are normalized to be the same. In detail, both $A(\cdot,\cdot)$: $\boldsymbol{W}\times\boldsymbol{H} \rightarrow \boldsymbol{L}$ and $B(\cdot,\cdot)$: $\boldsymbol{W}\times\boldsymbol{H} \rightarrow \boldsymbol{L}$ are two 2D mappings, where $\boldsymbol{W} = \{w \mid w<W, w\in \boldsymbol{N}^{+}\cup\{0\}$ and $W\in \boldsymbol{N}^{+}\}$, $\boldsymbol{H} = \{h \mid h<H, h\in \boldsymbol{N}^{+}\cup\{0\}$and $H\in \boldsymbol{N}^{+}\}$, and $\boldsymbol{L} = \{l \mid l<L, l\in \boldsymbol{N}^{+}\cup\{0\}\}$. W, H, and L–1 are referred to as the *width*, *height*, and *maximal gray level* of the images, respectively. In particular, a reshuffle of an image is defined to be a permutation of the pixels contained in this image. More specifically, a reshuffle $S(\cdot,\cdot)$: $\boldsymbol{W}\times\boldsymbol{H} \rightarrow \boldsymbol{W}\times\boldsymbol{H}$ is a one-to-one and onto mapping among pixels. We denote A_S to be the image obtained by performing the reshuffle S to image A.

Consider two pixels, (i_1, j_1) and (i_2, j_2), contained in the image A. Let the relative contrast between them, $C(i_1, j_1; i_2, j_2; A)$, be defined as the following mapping:

$$C(i_1, j_1; i_2, j_2; A) = \begin{cases} 1 & \text{if} \quad A(i_1, j_1) > A(i_2, j_2) \\ 0 & \text{if} \quad A(i_1, j_1) = A(i_2, j_2) \\ -1 & \text{if} \quad A(i_1, j_1) < A(i_2, j_2) \end{cases} \quad (1)$$

Some basic properties about the relative contrast are shown here:

Property 1: For all (i_1, j_1) and (i_2, j_2) $\in$ $\boldsymbol{W}\times\boldsymbol{H}$, the following properties hold: (i) $C(i_1, j_1; i_2, j_2; A) = 1$ if and only if $C(i_2, j_2; i_1, j_1; A) = -1$. (ii) $C(i_1, j_1; i_2, j_2; A) = 0$ if and only if $C(i_2, j_2; i_1, j_1; A) = 0$.

The relative contrast defined above characteristics as "brighter than" relation between two pixels of an image. In this paper, we are particularly interested in the transformations allowing us to obtain an image with consistent relative contrasts to another image. Given two images A and B, a reshuffle S is called a relative-contrast-consistent reshuffle of A for B if it satisfies both of the following criteria for all (i_1, j_1) and (i_2, j_2) $\in$ $\boldsymbol{W}\times\boldsymbol{H}$:

$$C(i_1, j_1; i_2, j_2; B) > 0 \text{ implies that } C(i_1, j_1; i_2, j_2; A_s) \geq 0 \quad (2)$$

$$C(i_1, j_1; i_2, j_2; B) < 0 \text{ implies that } C(i_1, j_1; i_2, j_2; A_s) \leq 0 \quad (3)$$

When a reshuffled image of A is consistent with B in terms of relative contrasts, this reshuffled image thus satisfies that the relative contrasts between its arbitrary pairs of pixels are consistent to those in B. In particular, the following property holds:

Property 2: Consider two images A and B, and a reshuffle S that is a relative-contrast-consistent reshuffle of A for B. Given two disjoint regions R^B_1 and R^B_2 in B, [1] assume that all pixels contained in R^B_1 are lighter than the pixels contained in R^B_2. Then, all pixels contained in R^A_2 are not lighter than R^A_1, where R^A_1 and R^A_2 contained in A are the regions with the same pixel indices of R^B_1 and R^B_2, respectively.

An illustration of Property 2 is shown in Figure 2.

[1] A region in an image is referred to as a set of pixels of this image.

The reshuffled image, A_S, looks like image B (because it has the same relative contracts with B) but is essentially a reshuffle of A. We also refer A_S to a *B-guided reshuffle of A*. We consider A_S a useful representation of A. In fact, A_S can be treated as an encrypted image of A, where the reshuffle is served as an encryption process. In the encryption process, the pixels of A are re-organized as a meaningful pattern by using the contents of B. Because A_S is a reshuffle of A, the encrypted image contains the complete information of A. The image A can be reconstructed without any loss from A_S by using the inverse reshuffle of S. Moreover, A_S looks like the image B but inherently hides the complete information of A. It can also be treated as hiding an image (e.g., A) into another one (e.g., A_S) that is a contrast-preserving version of a particular image (e.g. B). An example is shown in Figure 3.

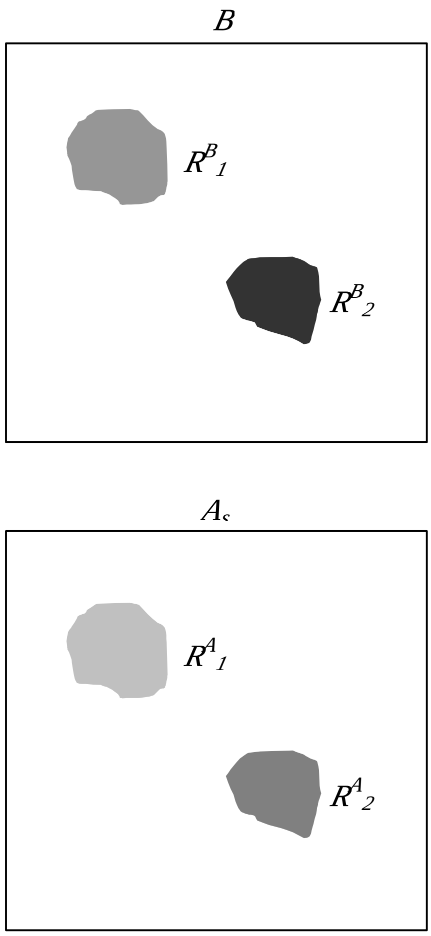

Figure 2. Illustration of Property 2. If the region R^B_1 is lighter than the region R^B_2, then R^A_2 is not lighter than R^A_1 where R^A_1 and R^A_2 are A's regions consist of pixels of the same positions of R^B_1 and R^B_2, respectively.

3. Main Algorithm

In the previous section, we have formulated the concept of relative-contrast-consistent reshuffle. In this section, we will investigate the problem about whether such a reshuffle exists and how to find it when it exists. Fortunately, the existence of such reshuffle is ensured and can be found through a simple algorithm, which is presented below.

In term of the sorting of the pixels of A. The pair of mappings, $i_A(\cdot)$, $j_A(\cdot)$:$\{n \mid n\in \boldsymbol{N}^+$ and $n\leq WH\} \rightarrow \boldsymbol{W}\times\boldsymbol{H}$, is a sorting of A if it satisfies the following property:

$$A(i_A(1), j_A(1)) \leq A(i_A(2), j_A(2)) \leq ...... \leq A(i_A(n), j_A(n)) \leq ...... \leq A(i_A(WH), j_A(WH)), \quad (4)$$

where $(i_A(m), j_A(m)) \neq (i_A(n), j_A(n))$ for all $m \neq n$. The sorting of A may not be unique because there may be some other pixels with equal gray levels in A. Similarly, let $(i_B(\cdot), j_B(\cdot))$ be a sorting of B.

Consider the reshuffle, $S(\cdot,\cdot)$: $\boldsymbol{W}\times\boldsymbol{H} \rightarrow \boldsymbol{W}\times\boldsymbol{H}$ defined as follows:

$$S(i_A(n), j_A(n)) = (i_B(n), j_B(n)) \quad (5)$$

for all $n\in \boldsymbol{N}^+$ and $n\leq WH$, where $(i_A(n), j_A(n))$ and $(i_B(n), j_B(n))$ are two sortings of A and B, respectively. The reshuffle $P(\cdot,\cdot)$ is a relative-contrast-consistent reshuffle of A for B, as shown in the following property:

Property 3: The reshuffle defined in (5) satisfies both the conditions shown in (2) and (3).

Pf. Let (i_1, j_1) and (i_2, j_2) be two arbitrary different pixels. Without lost of generality, we assume that $C(i_1, j_1; i_2, j_2; B) = 1>0$. Since the sorting is a one-to-one and onto mapping, there exists positive integers m and n, $m>n$, such that $(i_B(m), j_B(m))=(i_1, j_1)$ and $(i_B(n), j_B(n))=(i_2, j_2)$, respectively. Then, $C(i_1, j_1; i_2, j_2; A_S) = C(i_B(m), j_B(m); i_B(n), j_B(n); A_S) = C(S(i_A(m), j_A(m)); S(i_A(n), j_A(n)); A_S) = C(i_A(m), j_A(m); i_A(n), j_A(n); A) \geq 0$ because $m>n$ and $(i_A(\cdot), j_A(\cdot))$ is a sorting. Hence, $C(i_1, j_1; i_2, j_2; A) \geq 0$, and thus (2) holds. Similarly, we can prove that (3) holds.

Property 3 gives a way to find a relative-contrast-consistent reshuffle. Since the sorting of a sequence can always be performed, the existence of a relative-contrast-consistent reshuffle is guaranteed and one of them can thereby be found with the sorting technique introduced above. The main algorithm developed in this paper for finding the B-guided reshuffle of A (where A and B are two gray-level images of equal size) is shown in the following:

Algorithm [Finding Relative-Contrast-Consistent Reshuffle by Sorting]:

Step 1: Sort the pixels of A based on an increasing order. That is, compute $(i_A(n), j_A(n))$, $n = 1, 2, ..., WH$.

Step 2: Sort the pixels of B based on an increasing order. That is, compute $(i_B(n), j_B(n))$, $n = 1, 2, ..., WH$.

Step 3: Find a reshuffle of A by mapping the sorted pixels of A and B one by one, respectively. That is, set $S(i_A(n), j_A(n)) = (i_B(n), j_B(n))$ for all $n = 1, 2, ..., WH$.

Step 4: Output S.

Figure 3 gives an example about the application of the relative-contrast-consistent reshuffles to a pair of images. Figure 3(a) and 3(b) show two source images, A and B, respectively. Figure 3(c) and 3(d) show the images, A_S and $B_{S'}$, obtained by applying the reshuffles S to A and S' to B, respectively, where S is the B-guided reshuffle of A and S' is the A-guided reshuffle of B.

Figure 4 illustrates an interesting relationship between "relative-contrast-consistent reshuffle" and "histogram equalization"[16,17]. The purpose of histogram equalization is to provide a transformation between gray levels, which can transform an image into another image whose histogram is a uniform distribution. In fact, if B is an image whose gray levels are uniformly distributed from o to $L-1$ (or equivalently, the histogram of B is a uniform distribution), the image obtained by applying an A-guided reshuffle of B can be treated as a histogram- equalized image of A because this image preserves the relative contrasts of A and its histogram is uniformly distributed. Figure 4(a) is a 256x256 grayscale image with 256 gray levels. Figure 4(b) is a 256x256 ramp image with grayscale uniformly distributed from 0 to 255. Figure 4(c) is the resulted image from 4(a)-guided reshuffle of 4(b). By observing both 4(a) and 4(c), the details in dark regions can be distinguished in Figure 4(c). Histogram equalization, therefore, is a special case for our "relative-contrast-consistent reshuffle" while the histogram of the guiding image is evenly distributed.

4. Experimental Results

In this section, we present the experimental results about encrypting an image with the help of another image by using the relative-contrast-consistent reshuffle introduced in Sections 2 and 3. In particular, we focus on the encryption of a microarray image based on a particular pattern (e.g., a face picture). Figure 5(a) and 5(c) show a micro-array image constructed by enzyme-linked colorimetry detection [5] and its feature image, respectively. Figure 5(b) shows a face picture that is served as a meaningful pattern to encrypt this microarray image. After finding the Figure 5(b)-guided reshuffle of Figure 5(c) by using the algorithm described in Section 3, a reshuffled image of Figure 5(c) can be obtained as shown in Figure 5(d). It is observed that the reshuffled image, Figure 5(d), is a meaningful pattern similar to Figure 5(b) but is in content a microarray image. Given the reshuffle associated with this encryption, the microarray feature image can be completely recovered from Figure 5(d).

Figure 6 also shows some other examples that encrypt microarray images based on face pictures. According to the above results, the reshuffled images are clear face patterns that can be used for identifying the identities of person. Moreover, they are also the microarray images that contain gene-expression information of the persons. Hence, the method is able to integrate both types of person identity information, face patterns and gene expression, into a single image. Figure 7 shows some other examples in which the meaningful patterns are not faces but logotypes. Figure 7(a), 7(c), and 7(e) are some particular logotypes while Figure 7(b), 7(d), and 7(f) show their corresponding reshuffled images, respectively. These examples show that a microarray image may not have to be manufactured in its original way where image patterns consisting of scattered dots are usually constructed. On the other hand, the reshuffle can be performed in the manufacturing process and thus the produced microarray images look just like the logotypes. Since any microarray images produced in this way are in content logotypes from the perspective of appearance, the gene-expression information is encrypted and thus personal data can be protected. It can also be treated by hiding an original gene-expression image in the logotype of the company that manufactures the gene chip.

5. Conclusions, Discussion, and Future Work

In this research, the data hiding and encryption for micro-array image has successfully been implemented while the encrypted image possesses both features of gene expression and a meaningful pattern. Compared to the original human face image and the encrypted image, it looks like a faded image lack of rich gradations because of the generic attribute of the bio-chemical image development process. However, the meaningful pattern is very clear to human vision. We believe the results of this research would have great potential in producing next-generation personal identification and preparing secured information.

The method we proposed in this research is very useful for micro-array images from both laser-induced fluorescence detection and enzyme-linked colorimetry detection.

Although the background colors for the former and the later are dark and bright, respectively, the reconstructed hidden image is synthesized according to the features of gene expression, which has offset the background effect in the image analysis procedure. Recently, some researchers proposed quantum-dot-tagged micro-beads for gene expression studies and the precision can be up to single-bead level [15]. It might be helpful to increase the sensitivity for micro-array system in the near future.

Another thing should be noted here is the signals discussed above in this research only deal with the grayscale signal. The algorithm developed in this paper indeed can be extended to color image by dividing the color signal into 3 channels, R,G and B, and be applied to each of them. In order to improve the quality of the resultant appearance, color space other than RGB is appropriate. This is scheduled as the future work of this research.

At last, in this research, we manipulate the micro-array image and resultant hidden image only in electronic data format. Can the resultant hidden image be a "physical image"? That is, the resultant hidden image can be printed on a substrate (like paper) or even a "graphic" micro-array can be fabricated by altering the sequence of the 9600 gene fragments contact printed on the substrate. This is especially valuable for enzyme-linked colorimetry detection micro-array, because without image capture and digitization, the hybridized micro-array physical image (on a nylon membrane) itself is visible to human vision. This image can be incorporated with ID cards for personal identification or other security documents. A manufacturer's logo can also be hidden in it for customers and researchers' quick recognition to select the correct analysis tool. A tentatively synthesized "graphic" micro-array is shown in Figure 8, and the physical one is under fabrication, which is as listed as the future works.

Acknowledgments

This research is partially supported by the National Science Council of the Republic of China under the grant of NSC89-2320-B-128-002-M08.

References

[1] W. Bender *et al.*, Techniques for data hiding, IBM System Journal, vol 35, Nos 3&4, pp 313-336, 1996.

[2] D. Gruhl, W. Bender, Information Hiding to Foil the Casual Counterfeiter, proceedings (Lecture Notes in Computer Science; 1525), 2nd International Workshop on information hiding, pp 1-15, Portland, Oregon, April 1998.

[3] H.C. Wang, Data Hiding Techniques for Printed Binary Images, IEEE International Conference on Information Technology: Coding and Computing (ITCC-2001) -- Special Session on Multimedia Security and Watermarking Application, pp 55-59, Las Vegas, Nevada, April 2001.

[4] D.L. Hecht, Embedded data Glyph technology for hardcopy digital documents, SPIE 2171, pp. 341-352, 1995

[5] Chen, J.W. *et al.*, "Profiling expression patterns & isolating differentially expressed genes by cDNA microarray system with colorimetry detection", *Genomics*, 1998, 51(3), 313-324.

[6] Ermolaeva, O. *et al.*, Data management and analysis for gene expression arrays, Nature Genetics, vol 20, September 1998.

[7] Wen, L., *et al.*, Large-scale temporal gene expression mapping of central nervous system development, Proc. Natl. Acad. Sci. USA. Neurobiology, 95:334-339, January 1998.

[8] Iyer, V.R., *et al.*, The transcriptional program in the response of human fibroblasts to serum, Science, 283:83-87, January1, 1999.

[9] Eisen, M.B., *et al*, Cluster analysis and display of genome-wide expression patterns, Proc. Natl. Acad. Sci. USA, Genetics, 95:14863-14868, December 1998.

[10] Marshall, A. and Hodgson, J, DNA chips: an array of possibilities, Nature Biotechnology, 16:27-31, 1998.

[11] Ramsay, G., DNA chips: State-of-the-art, Nature Biotechnology, 16:40-44, January 1998.

[12] http://www.nhgri.nih.gov/DIR/LCG/15K/HTML/img_analysis.html

[13] http://cmgm.stanford.edu/pbrown/mguide/index.html

[14] S.K. Moore, Making chips to probe genes, IEEE Spectrum, pp 54-60, March 2001.

[15] M. Han, *et al.*, Quantum-dot-tagged microbeads for multiplexed optical coding of bio-molecules, Nature Biotechnology, Vol 19, pp 631-635, July 2001.

[16] W.K. Pratt, Digital image processing, 2nd Ed, John Wiley & Sons, Inc., 1991.

[17] B. Jahne, Practical handbook on image processing for scientific applications, CRC Press, New York, 1997.

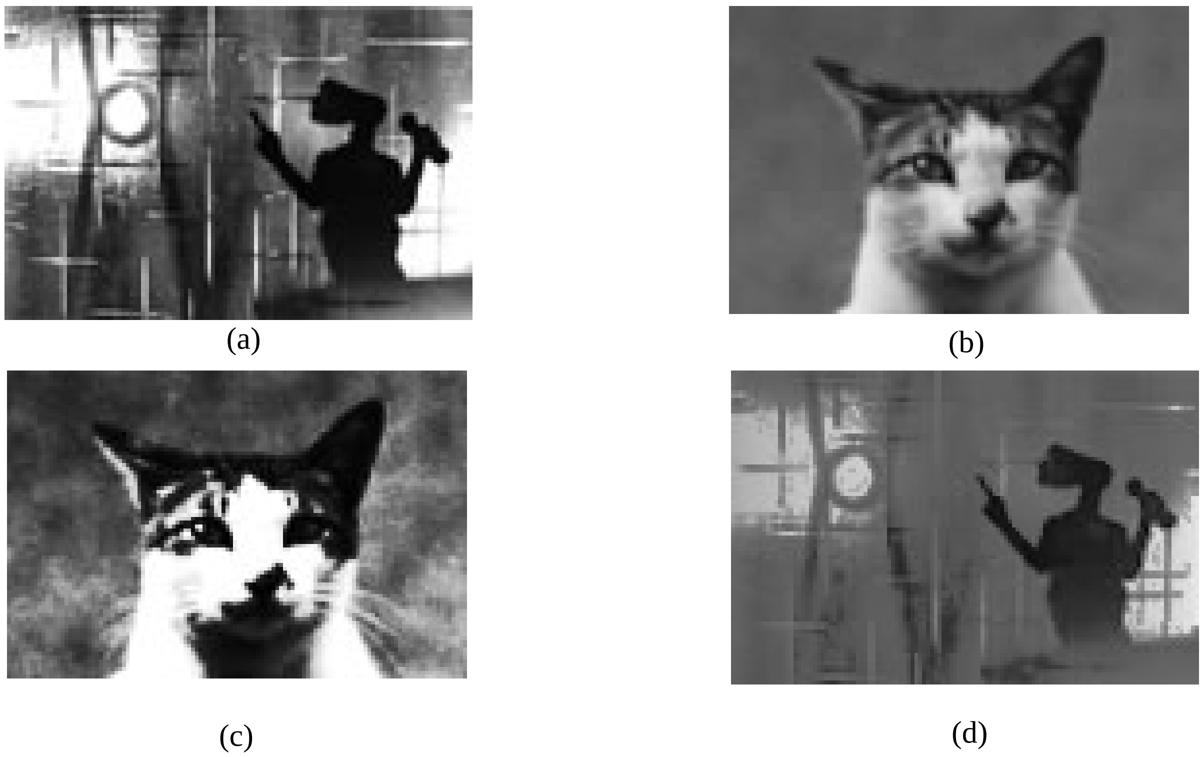

Figure 3. (a)(b) Two images A and B. (c) The reshuffled image of A, obtained by applying the B-guided reshuffle of A. (d) The reshuffled image of B, obtained by applying the A-guided reshuffle of B.

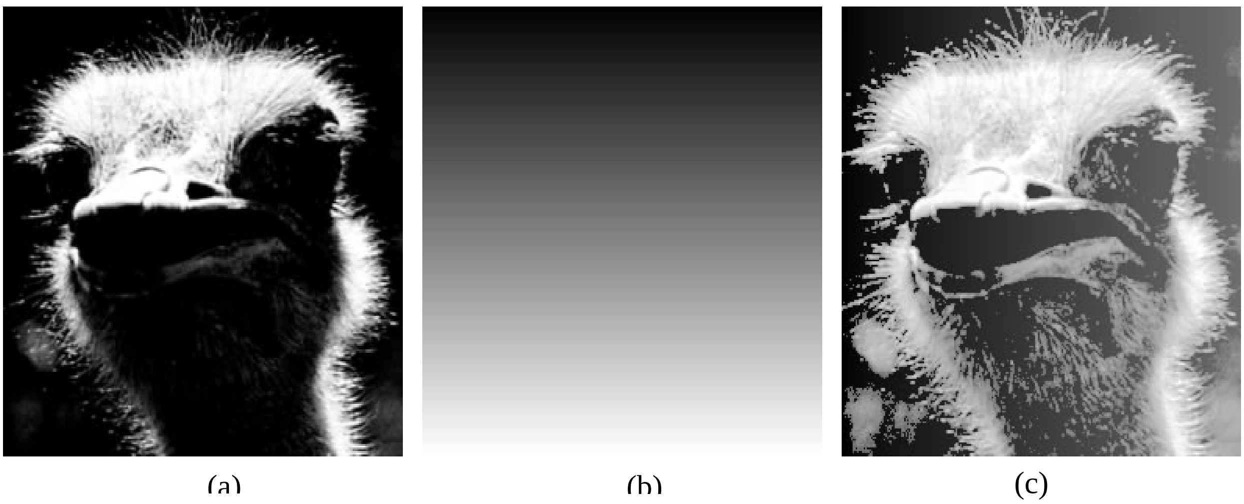

Figure 4.(a) image A. (b) a ramp image B. (c) The reshuffle image of A, obtained by applying the B-guided reshuffle of A.

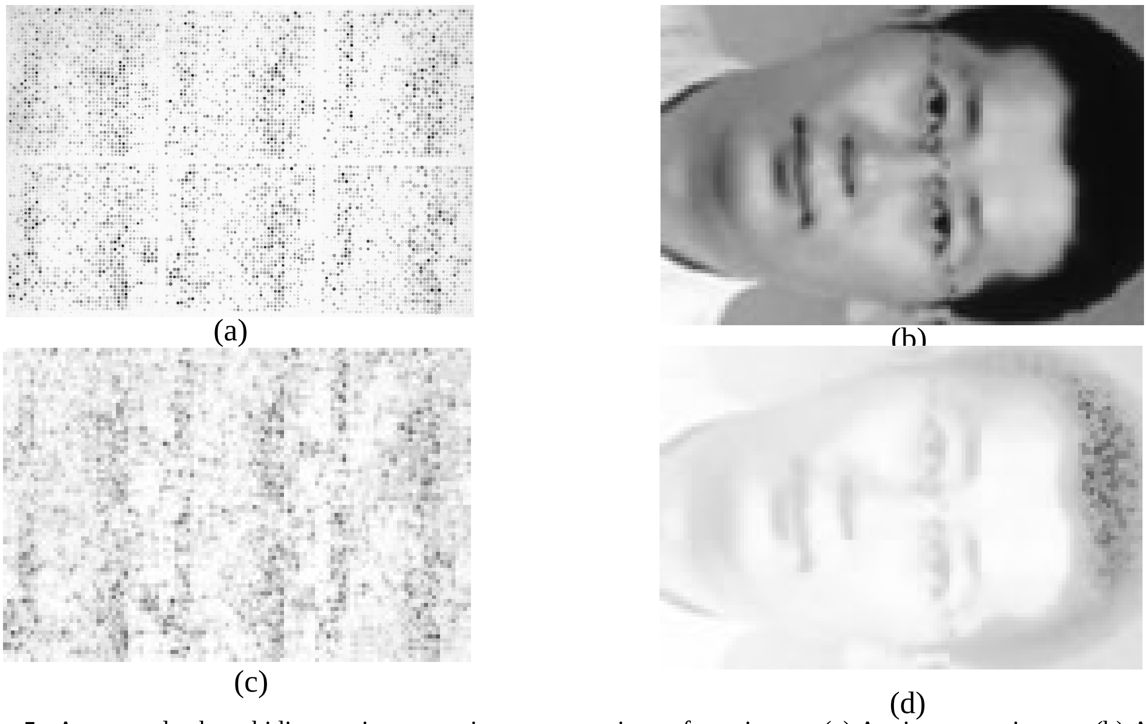

Figure 5. An example about hiding a micro-array image pattern into a face picture. (a) A micro-array image. (b) A face image. (c) The feature image extracted from (a). (d) The reshuffle image obtained by using the relative-contrast-consistent reshuffle to the micro-array image.

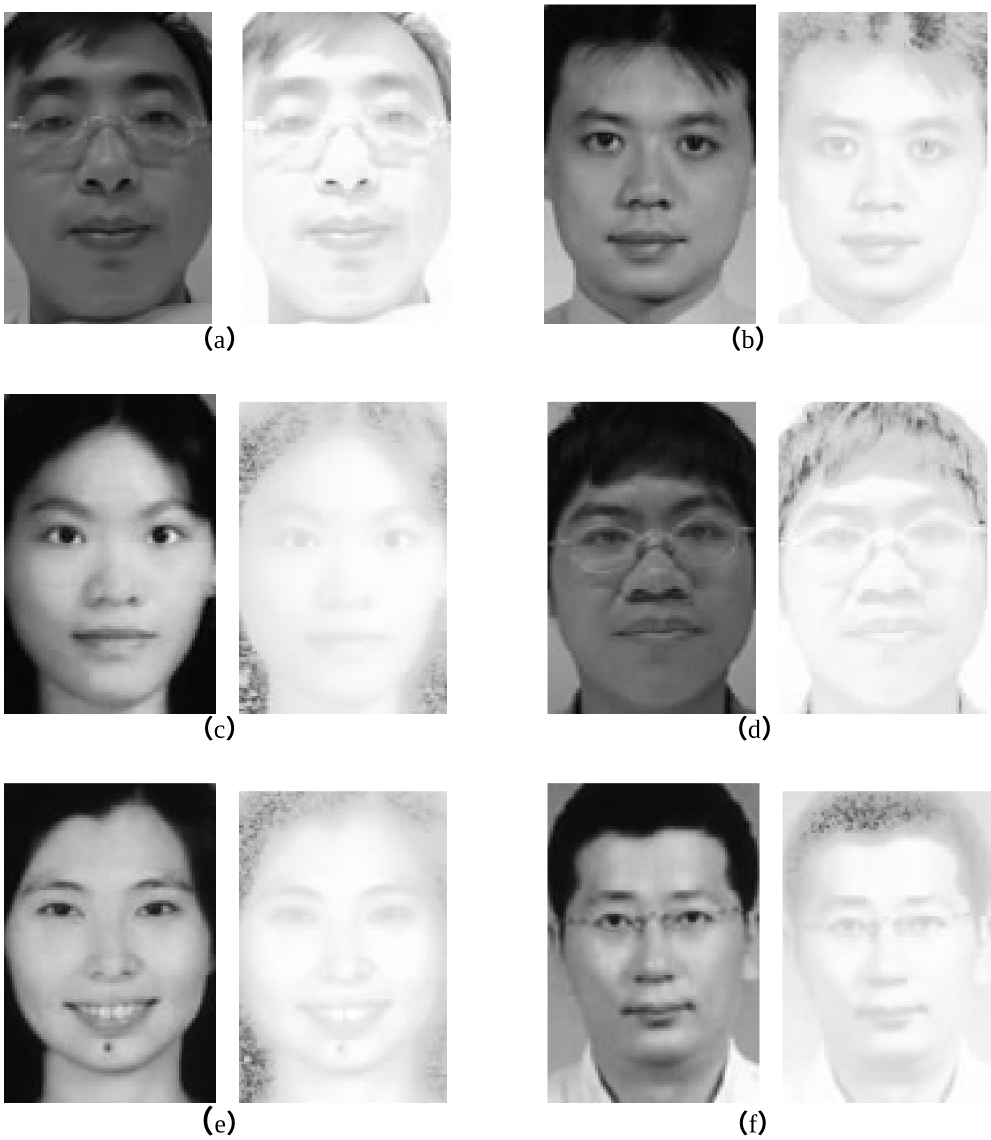

Figure 6. Some other examples of the face images and their reshuffled images obtained by using the associated relative-contrast-consistent reshuffles to micro-array images.

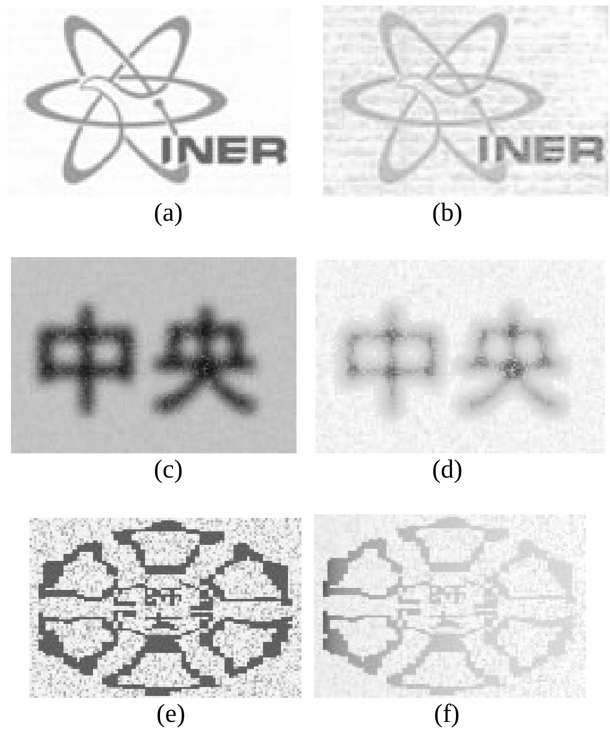

Figure 7. (a) (c) (e) logo images. (b) (d) (f) hiding micro-array images in these logotype.

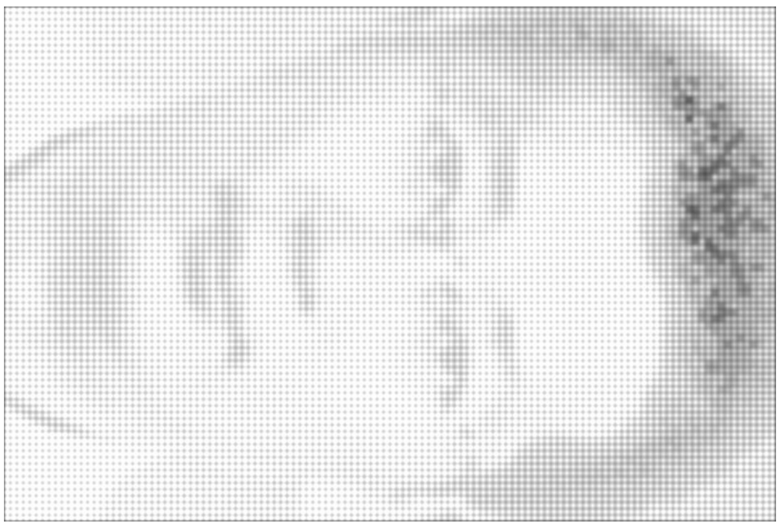

Figure 8. A composed 'graphic' micro-array gene chip image