Face Detection and Encryption for Privacy Preserving in Surveillance Video

Suolan Liu(✉), Lizhi Kong, and Hongyuan Wang(✉)

Changzhou University, Changzhou 213164, Jiangsu, China
lan-liu@163.com, hywang@cczu.edu.cn

Abstract. A number of techniques have recently been proposed for privacy preserving in video surveillance. Most of them are irreversible or have interference effect to the observation and recognition of human activities. In this paper, we address these issues by developing an effective method including face detection and encryption. In face detection, skin-color based approach fusing with fuzzy clustering is produced to detect facial candidates coarsely, and then we refine face by using SVM classifier. In face encryption, a reversible hybrid encryption (decryption) scheme based on spatial and value scrambling models is proposed. Simulation results verify the proposed mechanism can effectively detect and obscure faces while leaving the activities comprehensible and has high key sensibility for reducing the probability of attacking.

Keywords: Privacy preserving · Video surveillance · Face detection
Face encryption · Reversible

1 Introduction

Nowadays, video surveillance has become one of the most important auxiliary means in the field of public security monitoring. Video surveillance systems are widely deployed in many public places such as banks, supermarkets, airports, roads and residential areas [1–3]. Everyone is constantly being watched no matter whether you feel like it or not. However, in [4] a report about government surveillance revelations by NAS contractor Edward have raised new concerns about how best to preserve American's privacy in the digital age. What is personal privacy? One approach defines it in property terms as any information which the individual has certain decisional right [5]. Thus one's facial image, actions, location or copyrighted material are personal-partly, because they "belong" to the individual. Among these privacies, facial image is crucial and has highly close relationship with the others, because it can be directly used in face recognition technology to identify the monitored person's identity [2, 3]. In general, privacy preserving measures based on video surveillance can be taken from two aspects [6, 7]. On the one hand, we should enhance law making and law enforcement to regulate videos collection, storage and usage to avoid malicious infringement and disclosure of individual information. On the other hand, it is necessary to take effectively technical measures to protect the data and information, such as using cryptography theories and computer vision algorithms [2, 3, 7]. Cryptography methods mainly focus on encrypting the whole frame images into an unreadable form

J.-H. Lai et al. (Eds.): PRCV 2018, LNCS 11258, pp. 162–172, 2018.
https://doi.org/10.1007/978-3-030-03338-5_14

so that every unauthorized person cannot recover the original video [8]. The traditional encryption algorithms mainly include symmetric cryptographic algorithm and asymmetric cryptographic algorithm. This kind of methods is fit to processing videos for secure transmission over a communication line instead of real-time security monitoring and alarming for some particular activities recognition (e.g. fall and fights, etc.). Therefore encryption of the image as a whole may not be the most fixable method for this application. Recently, privacy preserving method based on computer vision has been a hot topic in the research field. Most of the preserving mechanisms are focused on partly modify the moving targets in the surveillance scenes [8, 9]. Target detection algorithms are used to localize the sensitive regions (e.g. face, eyes) and other methods are applied to obscure or conceal the selected regions, such as video masking, black boxes and replacing techniques. However, these methods are usually irreversible. Objective to recover the original video whenever needed for authorized person, we should apply reversible image processing methods with low computational complexity to meet the requirement of fast and real-time processing and preserving.

In this paper, we address the above-mentioned issues of privacy preserving in surveillance video by fusing image-processing method with encryption and decryption techniques. In particular, the proposed scheme consists of two steps including face detection and scrambling with the purpose of obscuring human face and monitoring his activities without revealing his identity at the same time.

The remaining of this paper is organized as follows. In Sect. 2, we review previous work related to pedestrian face detection, image encryption and decryption algorithms. Section 3 describes the proposed framework. The overview of the scheme is presented. Face detection approach and image encryption based on pixels spatial and value features scrambling models are given. In Sect. 4, simulations and experimental results are reported. Furthermore, we discuss the security of our proposed scheme. Finally, we conclude our work in Sect. 5.

2 Related Works

At present, video surveillances are widespreadly set up for the purpose of ensuring security and smart life. From this point of view, one may like surveillance to be carried out with not be willing to reveal any individual information. As the most informative part of human, face is usually used for identification. Therefore, obscuring or concealing face technique becomes an urgent demand for video surveillance with privacy preserving. Face detection is the first step of this application. Many of the current face detection techniques contain two major modules including face localization and verifying by extracting ‘facial’ features. To accurately localize face region, some prior information of human face are required. Skin color and face geometry make explicit use as apparent properties. Human skin color is one of the most robust face features and can be efficiently applied to find the pixels belonged to human skin in a scene. Roughly, physical-based methods and statistical-based methods are two basic kinds of skin color-based face localization approaches. Furthermore, statistical-based approached can be grouped into parametric approaches and non-parametric approaches. In parametric approaches, mean values, covariance matrices, Gaussian or mixtures of

Gaussians are used to build parametric face skin distribution models. For instances, in [10] Pujol et al. developed a fuzzy system to detect facial region by computing and fusing image variances from three color spaces of RGB, HSV and YCbCr. For the considering of error detections, a method of detecting where truly face locates is further proposed to eliminate these similar regions, such as the neck and hands. Experiments showed about 93% correct face detection rate in brief backgrounds and stable light conditions. In RGB space, Zhen et al. [11] built a maximum entropy model called the first order model (FOM) for parameter estimating human face. And then belief propagation algorithm was used to obtain fast selection and exact location for facial skin region. But the output of detection was in a gray scale skin map and the special region was not exactly located and marked. In non-parametric approaches, histogram, Bayesian approach and neural networks are usually developed to distinguish "face" or "non-face". In [12], authors applied the histograms of oriented gradients (HOG) as skin feature extraction clue and a feed-forward neural network was trained to classify the face from candidates. They tested the performance of their proposed scheme in sequences of color images and achieved an accuracy of 91.4%. The recent research of convolutional neural networks (CNNs) as the hottest algorithm in application of videos has proposed different solutions for incorporating the face detection and human recognition. Lu et al. [13] proposed using Clarifai net [14] and VGG-D model [15] to extract features and fuse them before fine-tuning. A binary classification by support vector machine (SVM) was conducted to realize face detection. Experimental results on three public datasets verify its state-of-the-art performance. Although great progress has made in recent years, face detection is still confronted with many challenges and cannot handle the large variations in different poses, occlusion, illumination condition and face in poor-quality video sequences.

As reported in [8], Boult proposed to protect privacy by using and adapting encryption techniques and combining them with intelligent video processing methods. The main contribution showed as cryptographically invertible obscuration only for authorized users in possession of the decryption key. Image encryption methods have been increasingly applied to meet security demands in video surveillance. The traditional encryption algorithms mainly include symmetric cryptographic algorithm and asymmetric cryptographic algorithm. Data encryption standard (DES), Triple data encryption standard (TDEA), Rivest Cipher5 (RC5) and International data encryption algorithm (IDEA) are typically symmetric cryptographic algorithms, while RSA (proposed by Ron Rivest, Adi Shamir and Leonard Adleman in 1977), ELGAMAL, RABIN, Diffe-Hellman and Elliptic curve cryptography (ECC) are asymmetric cryptographic algorithms. Video processing requires meeting its need such as fast and high-level efficiency. Therefore these traditional encryption algorithms may not be the most desirable algorithms for encrypting video frames with large size. By analyzing recent reports and publications, encryption schemes for image application may be grouped into three categories including pixel-position permutation, value permutation and hybrid scrambling methods. Arnold transform, Fibonacci transform and Hilbert transform are position permutation approaches with the disadvantage of not being able to change the original histogram. They only rearrange the positions of the image pixels rather than the pixel values. Once the histogram is revealed, exhaustion method can be used to find the original image. Value permutation-based algorithms such as Virginia encryption [16], chaotic map [17] and gravitational transform [18] aim at changing value by setting some parameters in advance. However, contour of original image can always be found in the encrypted image, which may cause security issues. Hybrid scrambling methods are produced by combing the advantages of the two former methods. In [19], to property compromise between imperceptibility and robustness of logo image encryption, Roy et al. proposed to fuse redundant discrete wavelet transform (RDWT) with Arnold scrambling and furtherly reshape it. Qin et al. [20] presented a novel image hash securely generated scheme by diving the image into several quantizes and scrambling the variances of pixel values. Testing results showed good performances with respect to perceptual robustness and discrimination. In [21], a hybrid encryption scheme based on quaternion hartley transform (QHT) and two-dimensional logistic map are suggested to enhance the security level. Simulation results verified that the novel scheme not only had satisfied security level but also had certain robustness against cropping and noise disturbance.

3 The Proposed Method

In this section, we describe the proposed privacy preserving method based on two steps: face detection and face encryption. The framework is shown in Fig. 1. In our scheme, first, we develop cascaded classifiers to extract face from coarse-to-fine. Then, a hybrid encryption approach based on spatial and value scrambling models are used to change and rearrange pixels in facial region. The following subsections will discuss the procedure detailedly.

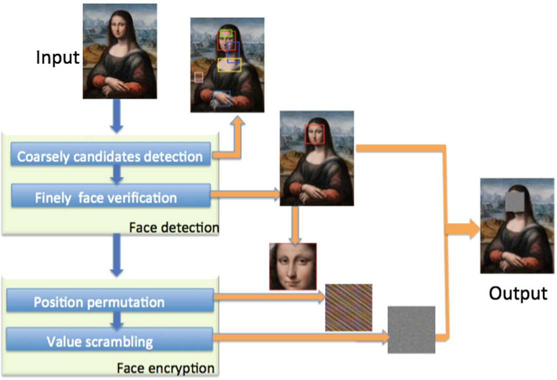


Fig. 1. Framework of our proposed scheme

3.1 Face Detection

Since most of the monitoring devices provide RGB video streams, approaches developed in this paper are based on RGB applications. Skin color model can be used to coarsely search facial candidates. Obviously, RGB has the negative property of each coordinate (red, green and blue) is subject to luminance effects from light changes, which may cause misclassification of skin and non-skin regions. Reported researches show that skin color models work effectively only on the chrominance subspaces such as Cb-Cr [22, 23] and Hue-Saturation (H-S) [24]. Inspired by the work in [25, 26], in our approach skin candidates are produced using fuzzy c-means clustering (FCM) based on pixel local properties termed as LFCM in Cb-Cr subspace. In [27], the standard FCM is used to localize skin-like regions. However, they only consider pixel value instead of other useful information, such as the relationships between pixels, which play important roles in discriminating the category of a pixel. With this in mind, we improve FCM by considering attributions from 8-neighbor pixels of a point. Therefore, the conditional probability of a pixel x_i categorized into the *jth* class can be expressed as:

$$f(j_i|\eta_i) = \frac{e^{\beta\delta_i(j)}}{\sum_{i=1}^{c} e^{\beta\delta_i(j)}}, i = 1, 2\cdots, N \tag{1}$$

where j_i means that the pixel x_i is classified into *jth* class. η_i is the class label from 8 neighbors. $\delta_i(j)$ is the statistical number of 8-neighbor pixels belonged to *jth* class. β is the weight factor, $\beta \geq 0$. We set $\beta = 0.5$ in all of our tests in Sect. 4. The following criterion can be used to discriminate the pixel's category:

$$j^* = arg \max u_{ij}, i = 1, \cdots, N; j = 1, \cdots, c \tag{2}$$

where u_{ij} is the fuzzy membership value and can be calculated by the following formula:

$$u_{ij} = u_{ij}^{'} \times f(j_i|\eta_i) \tag{3}$$

$u_{ij}^{'}$ can be obtained from the standard FCM.

To refine facial region from several candidates, we conduct finely classification by SVM. To reduce the influences from illumination and different sizes, we do preprocessing including light compensation [27] and resizing every candidate to 64 * 64. Define the block size as 16 * 16 composed by cells sized 8 * 8 with moving step 8 * 8. Next, nine gradient orientation bins are selected to produce HoG features and concatenate them as final feature vector to train SVM model by using polynomial kernel function [28].

3.2 Face Encryption

Once face region is properly detected, the next step is scrambling it for security and privacy protection. Note that the encrypted face should be able to be recovered as needed [29]. Motived by this requirement, a reversible hybrid encryption (decryption) scheme is proposed in this section, which uses Arnold transform in spatial position permutation [30] combining with gravitational transforms termed as GTs in value permutation [18] to encrypt and decrypt human facial region. In our numerical setting, to facilitate Arnold transform, facial region is located in a bounding box sized N*N. The facial region image is expressed as $f(x_i, y_i)$. In mathematics, the hybrid encryption operation is described as follows:

$$F(x_o, y_o) = G\{\mathrm{A}(\mathrm{f}(\mathrm{x}_i, \mathrm{y}_i))\}(\mathrm{x}_o, \mathrm{y}_o) \tag{4}$$

where $F(x_o, y_o)$ represents the output. The symbol “A” means Arnold transform (ART), “G” denotes GTs.

Furthermore, the facial image is imported to Arnold transform function [18], which is defined as:

$$\mathrm{A}_N : \begin{bmatrix} x_i' \\ y_j' \end{bmatrix} = mod\left(\begin{bmatrix} 1,1 \\ 1,2 \end{bmatrix} \begin{bmatrix} x_i \\ y_j \end{bmatrix}, N \right) \tag{5}$$

where $\left(x_i, \mathrm{y}_j\right)$ and $\left(x_i', y_j'\right)$ are the coordinates before and after position permutation A_N.

The GTs can be given as:

$$G : \left[\gamma \frac{m_r \times m_{x_i' y_j'}}{(x_r - x_i')^2 + \left(y_r - y_j'\right)^2 + k^2} \right] mod256 \oplus \mathrm{V}\left(x_i', y_j'\right) \tag{6}$$

γ is gravitational coefficient and assigned a large positive number in experiments. $m_r = 1$ is the quality of unit particle which location is (x_r, y_r). k is an adjusting parameter to ensure $(x_r - x_i')^2 + \left(y_r - y_j'\right)^2 + k^2 > 0$. $m_{x_i' y_j'}$ is the quality of the pixel point $\left(x_i', y_j'\right)$ with pixel value $\mathrm{V}\left(x_i', y_j'\right)$. Note that V can be a three-tuple corresponding to components of color images.

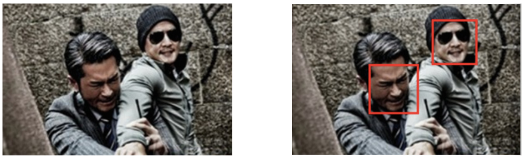
4 Numerical Simulations and Discussion

The main idea of our work is to develop a reversible method for human face obscuring while having no interference to recognizing and monitoring their activities. To verify the performance of the proposed scheme, we do experiments by choosing several video clips with life scenarios. The operations in the processes of face detection, encryption and decryption will be conducted in Matlab running on a laptop.

4.1 Test One

In this test, the original testing image shown in Fig. 2(a) contains two faces with variations in illumination, position, orientation and accessories. As displayed in Fig. 2 (b), our approach can effectively detect faces with a certain range of skin color changes. Even though the left-side person is lowing the head, his facial region is properly localized. For the right-side person, accessories such as sunglasses greatly increase the difficulties of face refining, which may result in partial detection of human face.

However, our algorithm can successfully suppress this kind of influence and detect the whole face region. Obviously, the effectiveness of this part will greatly facilitate the next step of encryption.



(a) original image (b) facial localization

Fig. 2. An example of face detection

In Table 1, we list the encrypted results by setting different parameters. For the sake of conducting fair comparisons, in GTs we set the unit particle's position as mean values for each facial position and assign the adjusting parameter k = 100. The first list displays the closeup of the detected faces; in the second list the ART results with different numbers of iterations are presented. We show the GTs results based on ART position permutation in the third list. The final encrypted results are displayed on the original images in the last list. From Table 1, one may find that with the changes of iterations from 3 to 80, the position scrambling effects show better from vision. Note that, once the number of iterations increases to a certain extent, it becomes a decryption operation. On the other hand, with the increase of gravitational coefficient the permutation of pixel values show more uniform and indistinguishable.

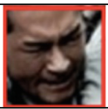
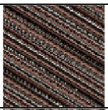
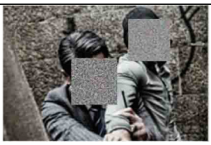
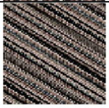
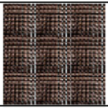
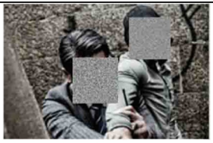
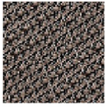
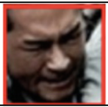
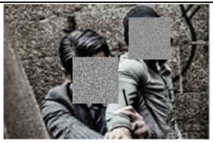
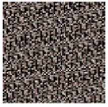
4.2 Test Two

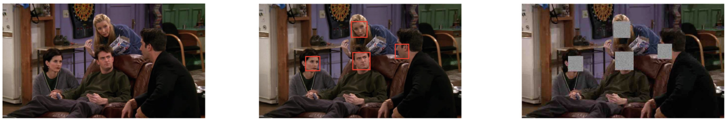
A frame image contains multiple faces from different views coupled with cluttering background is utilized to test the robustness and security of our proposed scheme.

For the purpose of strengthening the security, in this test we set encrypt key as KEY4 as following: the number of iterations is 150 and three different sets of parameters for GTs corresponding to 3 channels [18] are applied. For red, $m_{x'_i y'_j} = 85 \times x_i'^2 + y_j'^3 + 230$, $\gamma = 9 \times 10^{14}$; for green, $m_{x'_i y'_j} = 60 \times x_i'^2 + y_j'^3 + 175$, $\gamma = 11.8 \times 10^{15}$; for blue, $m_{x'_i y'_j} = 115 \times x_i'^2 + y_j'^3 + 70$, $\gamma = 10.5 \times 10^{13}$. The cipher-image is displayed in Fig. 3(c). As can be seen that even though the image shows small scaled faces and one of the actors in his profile, our method still achieves good detection rates and localizes the core areas of all faces.

To verify the key sensibility of the proposed method, we select the face from "Monica" shown in Fig. 4(a) and try to recover the encrypted image in Fig. 4(b) by using different decryption keys. Firstly, we decrypt it by using KEY5 of incorrect iterations as 90 for inverse ART operation, but no change to other parameters. The decrypted result displays in Fig. 4(c). Furthermore, we utilize only incorrect keys for inverse GTs operation with $m_{x'_i y'_j} = 50 \times x_i'^2 + y_j'^3 + 60$, $\gamma = 9 \times 10^{13}$ for all 3 channels

Table 1. Scrambling processings and encrypted results

Closeup of face	ART	GTs following ART	Encrypted results
			 KEY1: For ART, the number of iterations is 3; For GTs, $m_{x_i'y_j'} = 50 \times x_i'^2 + y_j'^3$+80, $\gamma = 10^{13}$
			
			 KEY2: For ART, the number of iterations is 50; For GTs, $m_{x_i'y_j'} = 70 \times x_i'^2 + y_j'^3$ +100, $\gamma = 10^{14}$
			
			 KEY3: For ART, the number of iterations is 80; For GTs, $m_{x_i'y_j'} = 90 \times x_i'^2 + y_j'^3$ +150, $\gamma = 10^{15}$
			



(a) original image (b) face detection result (c) encrypted result

Fig. 3. An example of multiple faces detection and encryption

as KEY6 and show the result in Fig. 4(d). Figure 4(e) is decrypted image with correct keys. Concludely, Figs. 4(c) and (d) indicate that the cipher-image can withstand some potential attacks. Experimental results show the high key sensibility in our scrambling scheme.

4.3 Discussion

Correlation coefficient between plain-image and cipher-image can be used to quantify the performance of an encryption algorithm. The lower correlation coefficient indicates that the encryption algorithm can better hide the feature information of the plain-image.

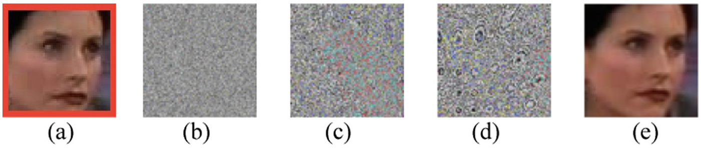


Fig. 4. Test results of key sensibility on a face

In this way, it will become more difficult to be attacked. Here we analyze the performance of the proposed algorithm by calculating the correlation coefficients of red, green and blue color channels respectively. The correlation coefficient between two-dimensional image matrix A and B can be defined as:

$$C_{AB} = \frac{\left|\sum_{i=1}^{N}\sum_{j=1}^{N}\left(A_{i,j} - mean(A)\right)\left(B_{i,j} - mean(B)\right)\right|}{\sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N}\left(A_{i,j} - mean(A)\right)^2 \times \sum_{i=1}^{N}\sum_{j=1}^{N}\left(B_{i,j} - mean(B)\right)^2}} \tag{5}$$

Where $mean(x)$ is the mean value.

Table 2 displays the correlation coefficient by using different keys. For Test One, we calculate the correlation coefficient between plain-image (Fig. 2(a)) and cipher-images (displayed in the forth list of Table 1). For Test Two, we calculate the correlation coefficient between cipher-image (Fig. 4(b)) and decrypted images (Figs. 4(c) and (d)).

From Table 2 one may find that for Test One most correlation coefficients are low as approximately zero. It indicates that the relevance between plain-image and cipher-image is very weak. From the aspect of encryption sensitivity, it means that the algorithm presented in this paper has superior sensitivity. Conversely, for Test Two while KEY5 is used to decrypt the cipher-image, correlation coefficient varies from 0.0243 to 0.0618, which shows high relevance. The reason for this phenomenon is the incomplete decryption of spatial position. However, we can see that once KEY6 is applied to decrypt, the average correlation coefficient is dramatically reduced from 0.0400 to 0.0061. As expected, localize and encrypt multiple faces in a picture is more challenging, but our proposed scheme is able to perform quite well with satisfied anti-attack property.

Table 2. Correlation coefficient between the red (r), green (g) and blue (b) color channels

Correlation coefficient		C_{rr}	C_{rg}	C_{rb}	C_{gr}	C_{gg}	C_{gb}	C_{br}	C_{bg}	C_{bb}	**Average**
Test one	KEY1	0.0007	0.0021	0.0015	0.0009	0.0018	0.0011	0.0014	0.0006	0.0023	**0.0014**
	KEY2	0.0013	0.0007	0.0003	0.0024	0.0007	0.0015	0.0008	0.0002	0.0009	**0.0010**
	KEY3	0.0004	0.0018	0.0021	0.0026	0.0015	0.0003	0.0020	0.0014	0.0031	**0.0017**
Test two	KEY5	0.0317	0.0243	0.0430	0.0357	0.0532	0.0618	0.0351	0.0426	0.0322	**0.0400**
	KEY6	0.0079	0.0050	0.0071	0.0068	0.0082	0.0064	0.0047	0.0038	0.0046	**0.0061**

5 Conclusion

We have proposed a practical privacy preserving technique for the application of video surveillance. Faces corresponding to privacy sensitive information are detected and encrypted. We aim to conceal faces while not interfere the observation and recognition of human activities using in intelligent monitoring and alarm systems. Our method is reversible for revealing faces whenever needed to the authorized person. Simulation results demonstrate that the proposed scheme can successively detect and obscure faces while leaving the activities comprehensible. Finally, the performance evaluation with key sensibility shows that the developed mechanism can withstand some potential attacks.

Acknowledgment. This work is supported by National Natural Science Foundations of China (No. 61572085, 61502058), Jiangsu Joint Research Project of Industry, Education and Research (No. BY2016029-15) and Changzhou Science and Technology Support Program (Social Development) Project (No. CE20155044).

References

1. Otto, C., Wang, D., Jain, A.: Clustering millions of faces by identity. IEEE Trans. Pattern Anal. Mach. Intell. **2**(40), 1–14 (2018)
2. Torre, M., Granger, E., Gorodnichy, D.: Adaptive skew-sensitive ensembles for face recognition in video surveillance. Pattern Recognit. **11**(48), 3385–3406 (2015)
3. Radtke, P., Granger, E., Sabourin, R.: Skew-sensitive boolean combination for adaptive ensembles: an application to face recognition in video surveillance. Inf. Fusion **15**(20), 31–48 (2014)
4. Maddern, M., Rainie, L.: Americans' attitudes about privacy, security and surveillance. http://www.pewinternet.org/2015/05/20/americans-attitudes-about-privacy-security-and-surveillance/
5. Haggerty, K., Ericson, R.: Varieties of personal information as influences on attitudes toward surveillance. http://web.mit.edu/gtmarx/www/vancouver.html
6. Bonetto, M., Korshunov, P., Ramponi, G.: Privacy in mini-drone based video surveillance. In: Workshop on De-Identification for Privacy Protection in Multimedia, vol. 4, pp. 2464–2469 (2015)
7. Dufaux, F., Ebrahimi, T.: Scrambling for privacy protection in video surveillance systems. IEEE Trans. Circuits Syst. Video Technol. **8**(18), 1168–1174 (2008)
8. Boult, T.: PICO: privacy through invertible cryptographic obscuration. In: Proceedings of the Computer Vision for Interactive and Intelligent Environment, pp. 27–38, October, 2005
9. Carrillo, P., Kalva, H., Magliveras, S.: Compression independent reversible encryption for privacy in video surveillance. J. Inf. Secur. **1**, 1–13 (2009)
10. Pujol, F., Pujol, M.: Face detection based on skin color segmentation using fuzzy entropy. Entropy **26**(10), 1–22 (2017)
11. Zhen, H., Daoudi, M., Jedynak, B.: Blocking adult images based on statistical skin detection. Electron. Lett. Comput. Vis. Image Anal. **2**(4), 1–14 (2004)

12. Aulestia, P.S., Talahua, J.S., Andaluz, V.H., Benalcázar, M.E.: Real-time face detection using artificial neural networks. In: Lintas, Alessandra, Rovetta, S., Verschure, P.F.M.J., Villa, A.E.P. (eds.) ICANN 2017. LNCS, vol. 10614, pp. 590–599. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-68612-7_67
13. Lu, X., Duan, X.: Feature extraction and fusion using deep convolutional neural networks for face detection. Math. Probl. Eng. **3**(2), 1–9 (2017)
14. Zeiler, M.D., Fergus, R.: Visualizing and understanding convolutional networks. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds.) ECCV 2014. LNCS, vol. 8689, pp. 818–833. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-10590-1_53
15. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. https://arxiv.org/abs/1409.1556
16. He, M., Qiang, S.: Novel image scrambling algorithm based on changing pixel values. Appl. Res. Comput. **12**(29), 4635–4638 (2012)
17. Belazi, A., Hermassi, H., Rhouma, R., Belghith, S.: Algebraic analysis of a RGB image encryption algorithm based on DNA encoding and chaotic map. Nonlinear Dyn. **4**(76), 1989–2004 (2014)
18. Liu, S., Yue, C., Wang, H.: An improved hybrid encryption scheme for RGB images. Int. J. Adv. Sci. Technol. **4**(95), 37–44 (2016)
19. Roy, S., Pal, A.: A robust blind hybrid image watermarking scheme in RDWT-DCT domain using Arnold scrambling. Multimed. Tools Appl. **2**(76), 1–40 (2017)
20. Qin, C., Sun, M., Chang, C.: Perceptual hashing for color image based of hybrid extracting of structural features. Signal Process. **142**, 194–205 (2017)
21. Li, J.: Hybrid color and grayscale images encryption scheme based on quaternion hartley transform and logistic map in gyrator domain. J. Opt. Soc. Korea **3**(20), 42–54 (2016)
22. Gundimada, S., Tao, L., Asari, V.: Face detection technique based on intensity and skin color distribution. In: International Conference on Image Processing, pp. 1413–1416, November 2004
23. Qing, L., Min, L.: Face detection using skin color and location relation. Comput. Eng. Des. **13**, 3396–3398 (2008)
24. Sabottka, K., Pitas, I.: Segmentation and tracking of faces in color images. In: International Conference on Automatic Face & Gesture Recognition, Vermont, pp. 236–241 (1996)
25. Anwar, N., Rahman, A.: RGB-H-CbCr skin colour model for human face detection. http://pesona.mmu.edu.my/~johnsee/research/papers/files/rgbhcbcr_m2usic06.pdf
26. Lu, J., Yuan, X., Yahagi, T.: A method of face recognition based on fuzzy c-means clustering and associated sub-NNs. IEEE Trans. Neural Netw. **1**(18), 150–160 (2007)
27. Hsu, R., Mottaleb, M.: Face detection in color image. IEEE Trans. Pattern Anal. Mach. Intell. **5**(24), 696–706 (2012)
28. Patilkulkarni, S., Lakshmi, H.: Vanishing moments of a wavelet system and feature set in face detection problem for color images. J. Comput. Appl. **16**(66), 36–42 (2013)
29. Liu, Z., Li, Q.: Image encryption based on random scrambling of the amplitude and phase in the frequency domain. Opt. Eng. **8**(48), 1–6 (2009)
30. Li, C., Lin, D., Lu, J.: Cryptanalyzing an image-scrambling encryption algorithm of pixel bits. IEEE Trans. Multimed. **3**(24), 64–71 (2017)