COMPRESSION INDEPENDENT OBJECT ENCRYPTION FOR ENSURING PRIVACY IN VIDEO SURVEILLANCE

Paula Carrillo[†], *Hari Kalva*[†], *and Spyros Magliveras*[‡]
[†]Dept. of Computer Science and Engineering, [‡]Dept. of Mathematical Sciences
Florida Atlantic University, Boca Raton, FL

ABSTRACT

One of the main concerns of the wide use of video surveillance is the loss of individual privacy. Individuals who are not suspects need not be identified on camera recordings. Mechanisms that protect the identity while ensuring legitimate security needs are necessary. Selectively encrypting objects that reveal identity (e.g., faces or vehicle tags) is necessary to preserve individuals' right to privacy. This paper presents a compression algorithm independent solution that provides privacy in video surveillance applications. The proposed approach is based on the use of permutation based encryption to hide identity revealing features. The permutation based encryption tolerates lossy compression and allows decryption at a later time. The use of permutation based encryption makes the proposed solution independent of the compression algorithms used. The paper presents the performance of the system when using H.264 video encoding.

Index Terms— video surveillance, compression, privacy, identity, encryption

1. INTRODUCTION

With video surveillance becoming an integral part of our security infrastructure, privacy rights are beginning to gain importance. The key concern is the fact that private citizens, who are not suspects, are being recorded and recordings archived through the use of video surveillance systems. Such a record-everything-and-process-later approach has serious privacy implications. The same privacy issues arise when surveillance cameras routinely record highway traffic as vehicle tags are recorded. The solution of removing the identities by blurring/blackening the portions of video is not acceptable to security personnel as they may have legitimate need to review the videos. On the contrary, leaving the videos with identities of people and vehicles public is a breach of privacy.

A solution to the problem is selective encryption of objects that reveal identity (e.g, faces, vehicle tags) in surveillance video. Objects in a video can be encrypted to ensure privacy and still allow decryption for legitimate security needs at anytime in the future. The goals of the video surveillance are still met as selective encryption allows monitoring the activities without knowing the identities of those being monitored. When a suspicious activity needs to be investigated, the identities can be uncovered with proper authorization. The few existing solutions are specific to video and image compression algorithms used and require modification to the video encoders [1,2]. These approaches limit the flexibility of surveillance systems. This paper presents a solution that meets the needs of individuals' privacy and legitimate security needs. The proposed solution is independent of the image and video compression algorithms used. This allows the use of standard video encoders and decoders and also enables smart-cameras that output encrypted video. The proposed solution also survives video transcoding and recoding allowing a normal video distribution chain with multiple video encoding and decoding operations. The innovation in the proposed approach is the use of permutation based encryption that can survive lossy compression.

2. BACKGROUND

The spread of video surveillance has generally been accepted by the public. However, there are also several concerns about loss of privacy and possible abuse of information. Legislation about data protection have been implemented in different countries and as a direct consequence of these regulations and concerns, different kinds of video privacy systems have been proposed and deployed. A privacy preserving video console is proposed in [1]. This approach is a method of rendering face images unusable by face identification software. Computer vision techniques are proposed to select interesting component information of the video and then obscuring that piece of information, or its components, such that face recognition software cannot recognize the faces. This solution requires the use of a special video console and anyone with access to the video sequence will have access to the identities. A transform-domain scrambling technique for MPEG-4 video was proposed in [2]. In this approach a region of interest is detected and then the signs of selected transform coefficients are scrambled. The decoded video will have blocky regions unless a proper key is used for de-

scrambling. This approach is specific to video compression algorithm used. Furthermore, this approach cannot survive operations such as transcoding that may be necessary to distribute video. Another technique proposed in [3] is privacy through an invertible cryptographic obscuration; the authors use DES/AES to encrypt regions of JPEG images during the compression stage, before Huffman encoding. This is similar to the transform coefficient sign scrambling reported in [2]. This method also suffers from the same drawbacks: it is compression algorithm specific and cannot survive transcoding. Another approach to privacy protection reported in [4] is based on detecting skin tones in images and replacing it with other colors, hence making it impossible to determine the race of the individual. This method is compression independent but does not hide the identity completely.

Compared with the methods that appear in the literature, our method is compression algorithm independent, survives transcoding and recoding to other formats, and uses a provably secure permutation based encryption. This enables flexible surveillance systems that can select compression algorithms that suite their needs.

3. SYSTEM ARCHITECTURE

Figure 1 show a video surveillance system based on the proposed approach. The object selection module performs the object detection (e.g, face or vehicle tag). Object detection is not the focus of this paper. Based on the selected objects, the regions of a video frame are encrypted using a secret key and then encoded using a standard video encoder. The encoded video can be stored or delivered live to a remote location. The video with the encrypted objects can be decoded and played on any standard decoder but the objects will remain encrypted. With a proper authorized key, the encrypted objects can be decrypted and displayed.

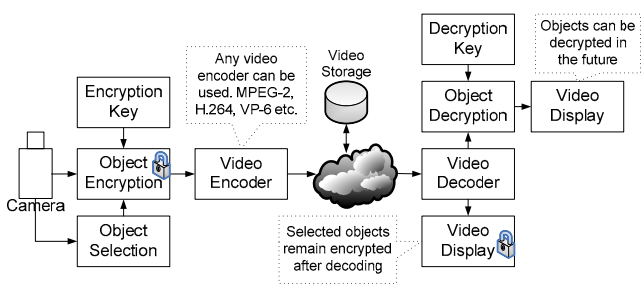

Fig. 1. System Architecture

The objects in the video are encrypted on a block basis. The 16x16 blocks that cover the selected objects are determined and then a block based encryption is applied. For each frame sequence of 16x16 blocks to be encrypted, a pseudorandom sequence of row-column permutation pairs (α_t,β_t) are applied to the cleartext sequence of blocks to yield the encrypted block sequence. Here, t is the frame number and the permutation sequences α_t and β_t are mutually independent, and are generated by a key-dependent random permutation generation algorithm based on *logarithmic signatures* [6,7,8]. Each key choice yields a sequence of permutation pairs (α_t,β_t) of periodicity $(16!)^2 \cong 4.4\text{x}10^{26}$. A much more secure process can also be used according to which the 256 pixels of the tth frame are permuted by the permutations of degree 256 in a random sequence σ_t again using the techniques of [6,7,8]. Now, the periodicity of σ_t can be made to be $256! \cong 10^{507}$. The encryption key can be generated dynamically based on the frame number and block number. The encryption key can be varied on a per-block or per-frame basis if desired. We have used permutation based video encryption in our prior work.

The encrypted macro blocks are then encoded using the standard encoding process. Since encryption rearranges the pixels in a block, the correlation is decreased and the compression rate for the block decrease. The application of the proposed encryption leads to increase in bitrate and the amount of increase depends on the content and the number of blocks encrypted. The video can be decoded on any standard decoder for the compression format used and the video remains encrypted after decoding and a key is necessary for decrypting the video.

(a) (b)

Fig.2. a) Original frame b) Frame with encrypted faces

4. EXPERIMENTAL SETUP AND EVALUATION

The system is evaluated with H.264, MPEG-2, MPEG-4, and H.263 video encoders using the Crew sequences at 352x288 resolution. Face detection for the experiments was done manually and face regions are input to the system. Figure 2 shows a frame of the original video and Figure 3 shows the video with encrypted regions that is input video encoders.

When video with encrypted objects is encoded, the encrypted blocks also suffer distortion due to lossy coding. The loss of correlation in the encrypted blocks leads to larger non-zero coefficients and quantization of these large coefficients increases the distortion in these blocks.

Figure 3a shows video with some visible distortion, but with largely acceptable video quality for faces. As QP increases, the quality of decrypted faces deteriorates. Figure 3b shows that the face quality is not acceptable when QP increases to 35.

(a) (b)

Fig.3. a) Decoded and decrypted H.264 video with Quantization Parameter (QP) of 26 b) Decoded and decrypted H.264 video with QP of 35.

Our experiments show that the quality of decrypted regions is good for videos encoded with QP of up to 28. The same set of experiments were repeated and the upper bounds for QP to ensure acceptable quality for decrypted regions is QP of 6 for H.263 and MPEG-4 and 3 Mbps for MPEG-2. The experiments show that the video has to be recorded with good quality in order to preserve the quality of decrypted regions. When lower bitrate surveillance is necessary, the encoder can enforce an upper bound on the QP used for the encrypted blocks.

The proposed approach increases the bitrate of the encrypted and then encoded video compared to the video encoded without encryption. This increase in bitrate is the tradeoff for the desired encryption features. Figure 4 shows the plot of encrypted video bitrate vs. standard video bitrates for the crew CIF sequence. The bitrate increase because of encryption is around 23% on average. The total bitrate can be decreased if the encrypted objects are treated as regions of interest (ROI) and coded using a fixed QP that results in a good reconstructed object while increasing the QP for the sequence. This ROI based approach is described in the next section.

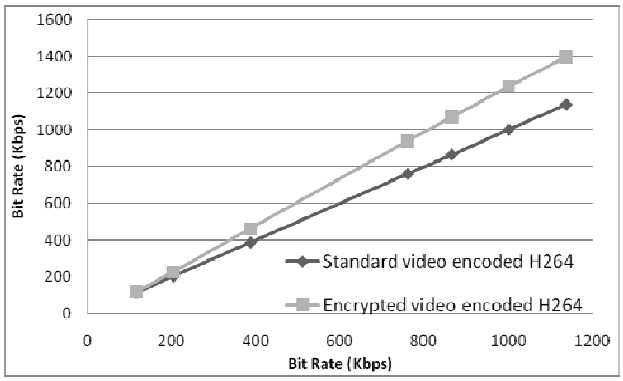

Fig. 4. Bitrate increase because of encryption

5.1. Fixed region of interest QP for low bitrate surveillance

Limiting the distortion of the encrypted objects, referred to as Region of Interest (ROI) here, allows surveillance systems to record video at lower bitrates. The relatively higher quality of ROI maintains the quality of decrypted objects at an acceptable level. The upper bound on QP, however, increases the bitrate. This increase in bitrate is, as in the previous case, the *cost* of providing privacy in video surveillance systems.

Fig. 5. Decoded and decrypted H.264 video with QP of 40

Figure 5 shows the quality of video at low bitrates, encoded with QP of 40. The high QP value distorts the encrypted blocks and the decrypted area are essentially lost. Figure 6 shows the video with encoded with QP of 40 but with ROI QP set to 26. With minimal quality maintained for the ROI, the faces can be clearly seen in the decoded video. As in the previous case the bitrate of the video increases.

Figure 7 shows the increase in bitrate when the QP is fixed for the ROI. The increase depends on the ROI QP. As expected the bitrate increases are higher for lower QP and lower for higher QP.

Fig. 6. Decoded and decrypted H.264 video with QP of 40 and with ROI QP fixed to 26

With a QP of 26, the encrypted video takes 23% more bits as explained in section 4. However, if the ROI QP is fixed at 26 and if a higher QP is used for the video sequences, the video bitrate can be reduced without affecting the quality of encrypted objects. Table 1 shows the reduction in overall bitrate compared to the encrypted video coded with QP of 26. The encrypted video at QP 26 was taken as the base bitrate for percentage comparison.

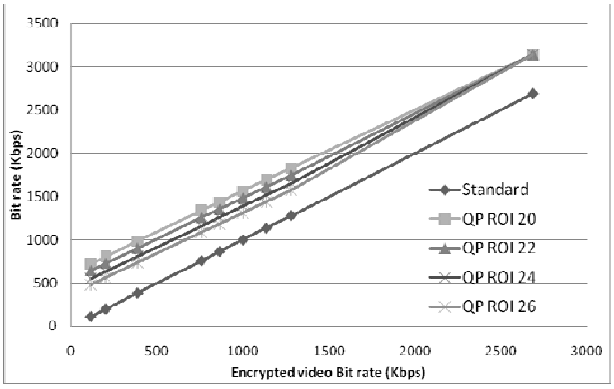

Fig. 7. Bitrate increase because of constant QP for the ROI

Table 1. Comparison between a video encoded with a QP of 26 and videos with QP > 26 and a fixed QP ROI of 26.

General video QP	Bitrate (Kbps) ROI QP = 26	% bitrate reduction compared to video coded with QP of 26
26	1571	0%
27	1438	-8%
28	1311	-17%
29	1184	-25%
30	1085	-31%
35	734	-53%
40	561	-64%
45	478	-70%

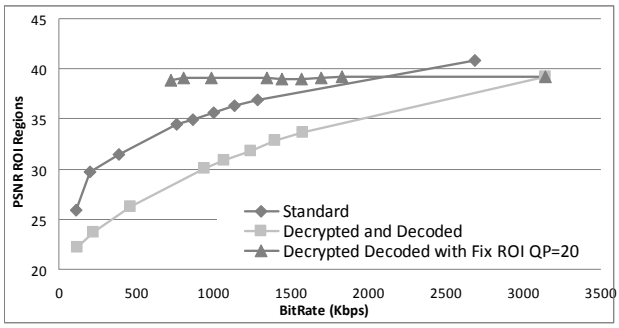

Fig. 8. Bitrate vs ROI PSNR

Figure 8 shows the PSNR of the ROI (faces) with standard encoding and the encoding of encrypted regions proposed in this paper. The figure shows that the minimum expected ROI PSNR when the ROI QP is fixed at 20 is approximately 39dB. Another important observation is that the ROI PSNR when the ROI QP is fixed is better or equal to that of non-ROI decryption and decoding process. However, the key metrics for evaluating this system are the ability to encrypt selected objects, ability to support multiple video compression algorithms, and the resulting increase in bitrate. Based on these metrics we conclude that the proposed system meets all the requirements of surveillance systems that can protect individual privacy rights. The tradeoff of bitrate for ensuring a minimum quality for the encrypted objects is a reasonable tradeoff and is within bounds of practical systems.

5. CONCLUSION

This paper presents a system for encrypting selected objects in videos. The system can be used for ensuring privacy in video surveillance by hiding the identity of the objects in the video. The selected objects can be decrypted in the future with the right decryption keys. The proposed system is independent of the compression algorithms used. The system was tested using H.264, MPEG-2, MPEG-4, and H.263 to verify compression algorithm independence. Providing encryption increases the video bitrate and experiments show that the increase is around 23%. This bitrate can be reduced by keeping the ROI QP constant and increasing the frame QP. The increase in bitrates depends on the type of video and the size of encrypted regions. The increase in bitrate is a reasonable cost to pay for protecting individual privacy.

6. REFERENCES

[1] A. Senior, *et. al.*, "Enabling video privacy through computer vision," IEEE Security & Privacy Magazine, vol.3, no.3, May-June 2005, pp. 50-57.

[2] F. Dufaux and T. Ebrahimi, "Scrambling for Video Surveillance with Privacy," 2006 Conference on Computer Vision and Pattern Recognition Workshop, pp. 160-160, 17-22 June 2006.

[3] T.E. Boult, "PICO: Privacy through Invertible Cryptographic Obscuration," Proceedings of the Computer Vision for Interactive and Intelligent Environment, 2005 , Nov. 2005, pp. 27-38.

[4] A. M. Berger, "Privacy mode for acquisition cameras and camcorders," US. Patent 6,067,399 May 23, 2000.

[5] D. Socek et. al., "New Approaches to Encryption and Steganography for Digital Videos", Multimedia Systems Journal, Vol. 13, No.3, Sept. 2007, pp. 191-204.

[6] S.S. Magliveras and N.D. Memon, "Random permutations from logarithmic signatures", Computing in the 90's First Great Lakes Comp. Sc. Conf., Lecture Notes in Computer Science, Springer-Verlag 507 1989, pp 91-97.

[7] S.S. Magliveras and N.D. Memon, "The algebraic properties of cryptosystem PGM", J. of Cryptology, 5, 1992, pp 167-183.

[8] S.S. Magliveras, Tran van Trung and D.R. Stinson, "New approaches to designing public key cryptosystems using one-way functions and trap-doors in finite groups". J. of Cryptology, 15, 2002, pp 285-297.