Intelligent Multi-Level Regions-of-Interest (ROI) Document Image Encryption Using an Online Learning Model

Alexander Wong and William Bishop
Department of Electrical and Computer Engineering, University of Waterloo
Waterloo, Ontario, Canada
a28wong@engmail.uwaterloo.ca, wdbishop@uwaterloo.ca

ABSTRACT

Image-based document management systems have become increasingly popular for handling documents that contain both text and graphical elements. When such systems are used to store confidential information, document security is a key concern. Conventional encryption techniques used for images fail to provide the level of flexibility required by such document management systems. Newer image encryption techniques provide improved flexibility at the cost of backwards compatibility and ease of use. This paper presents a novel approach to document image encryption using an online learning model. The proposed system provides backwards-compatible document image encryption in regions of interest (ROI) with support for multiple levels of authority. Furthermore, the proposed system is capable of learning from user feedback to improve the ROI selection used during the semi-automatic document encryption process. Experimental results from the encryption of test documents demonstrate the effectiveness of the proposed system.

KEY WORDS
Document image encryption, regions-of-interest, multi-level, online learning

1. Introduction

Digital document management systems are used by a wide range of organizations, such as government institutions, corporations, and hospitals. A typical requirement of current digital document management systems is the ability to manage scanned documents. Such documents may contain a combination of textual and graphical content. This has lead to the development of image-based document management systems, which store and manage scanned documents in image form. When image-based document management systems are used to store confidential information, document security is a significant consideration. In many systems, document images are simply encrypted using conventional file encryption techniques [1-3]. Encryption schemes designed specifically for images have also been proposed [4-6] to improve encryption speed but have been demonstrated to be less secure than conventional approaches [7-8]. One of the main drawbacks to traditional encryption schemes is the fact that the encrypted document image appears completely illegible. As such, it is not possible to use content-based image retrieval techniques for document image searches.

Recent developments in image security revolve around regions-of-interest (ROI) image security, where only regions containing confidential information are hidden from unauthorized users. One approach to ROI image security is through the use of JPSEC [9], an extension to JPEG2000 that permits conditional access to regions within a JPEG2000 image. Dufaux et al. proposed the use of pseudorandom noise to scramble the desired ROI within the image [10]. The seeds used in the pseudorandom noise generator are then encrypted and embedded into the image. An authorized user can then use the proper encryption key to decrypt the seeds, which can then be used to reveal the hidden ROI in the image. Multiple encryption keys can be used to provide different levels of access to the image. This approach provides multi-level ROI access control by allowing different ROI within an image to be visible only to specific authorized individuals.

For example, let us suppose that User A has a confidential document that contains three ROI, denoted by ROI_1, ROI_2, and ROI_3. User A would like to give User B the lowest permission to view only ROI_1, User C the permission to view both ROI_1 and ROI_2, and give User D the highest permission to view all ROI. However, User A wishes to post the same document publicly so that Users B, C, and D can view it. This can be achieved using multi-level ROI image security by encrypting ROI_1 with Key 1, encrypting ROI_2 with Keys 1 and 2, and finally encrypting ROI_3 with Keys 1, 2, and 3. That way, a user would require all three keys to view all hidden regions of the document.

Another approach to the multi-level ROI access control problem is through the use of a mixed raster content (MRC) model. In the MRC model, an image is divided into multiple layers. Therefore, it is possible to implement multi-level ROI image security by dividing the ROI in the image into individual layers, one for each authority level. The layers are then encrypted using different encryption keys such that only authorized individuals with the appropriate key can view the content within specific ROI.

While the ROI image security approach proposed by Dufaux et al. [10] and the approach based on the MRC model provide improved flexibility when compared to traditional techniques, these approaches limit backwards compatibility and ease of use. Since the first approach makes use of JPSEC, it is only applicable if all document images are stored using JPEG2000 with support for JPSEC. Therefore, this approach cannot be used with popular image formats such as TIFF, BMP, and JPEG. This problem is further compounded by the fact that JPEG2000 is not widely supported by web browsers and that JPSEC has not yet been finalized, making it difficult to implement in web-based digital document management systems. The second approach based on the MRC model requires the support of layers in the image format used to store the document image. With the exception of TIFF, this is not the case with many popular image formats.

Another major problem with both approaches is that, unlike traditional techniques that encrypt the entire image, the proposed ROI image security approaches require the selection of ROI for concealment in each individual document image. This is very time consuming, particularly in situations where a large number of document images need to be encrypted. An approach to ROI image security that reduces the time required by users for ROI selection and maintains backwards compatibility with existing image formats is clearly desirable.

The main contribution of this paper is a novel system for ROI image security for document images using an online learning model. The proposed system provides multi-level ROI access control to document images. The system is backwards compatible with existing image formats. Furthermore, the system provides semi-automatic document image encryption using an online learning model based on user feedback. In this paper, the proposed system is described and explained in detail in Section 2. Experimental results are presented in Section 3 to demonstrate the effectiveness of the proposed algorithm. Conclusions are drawn in Section 4.

2. Proposed System

The proposed document image encryption system provides an effective ROI image security solution for document images that addresses the issues faced by other approaches when used for the purpose of image-based document management systems. First, the proposed system addresses the issue of backwards compatibility by providing a scheme for ROI image encryption that is compatible with existing image formats. This allows existing content-based image retrieval schemes to be used with little or no modification required. Furthermore, it allows public content from encrypted document images to be viewable by the general public using a standard web browser or image viewer without the need for proprietary software. Finally, the proposed system addresses the issue of ease of use by providing a semi-automatic ROI selection scheme using an online learning model. This level of automation significantly reduces the manual labor required by the user.

2.1. Overview

The proposed document image encryption system is comprised of two main components:

i) ROI selection, and
ii) ROI image encryption.

The general process of the proposed system is illustrated in Figure 1. It can be seen from Figure 1 that the proposed system has a feedback loop configuration. In the proposed document image encryption system, the user inputs a document (or a batch of documents) into the system for encryption in image form. This can be accomplished by inputting a document image or using another method such as scanning a document and converting it into an image. The user may optionally provide the document type to reduce the processing time of the system. In the ROI selection component, the proposed system attempts to automatically select the appropriate ROI for the document image to be encrypted as well as determine its document type (if not given by the user) using an online learning system. The selected ROI and document type are shown to the user. The user is then given an opportunity to adjust and modify the selected ROI as well as the document type for the final document encryption process. The user feedback is then sent back to the online learning system, which then trains itself based on the feedback data. Finally, user feedback is passed to the data encryption component for ROI document image encryption.

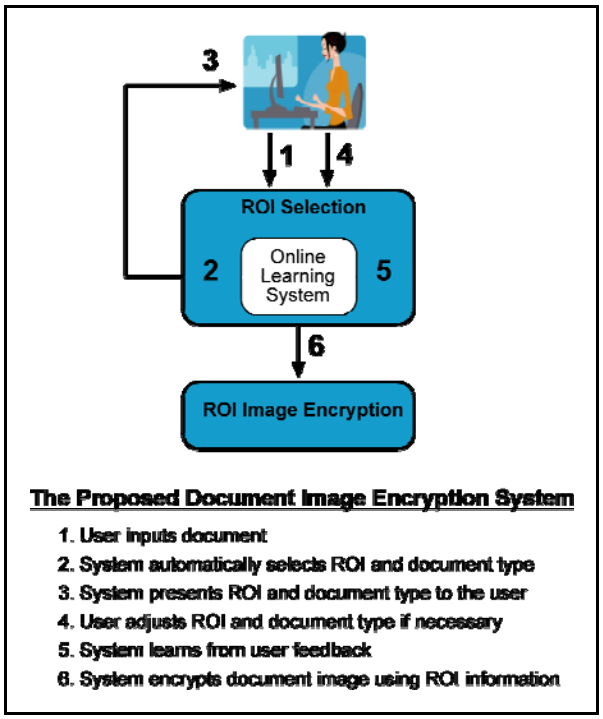


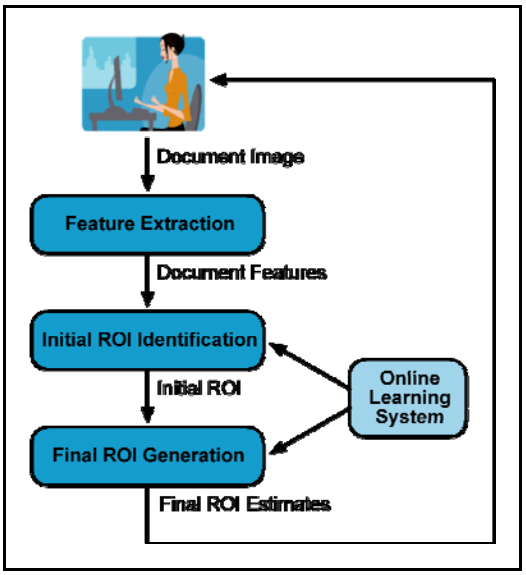
Figure 1. The proposed document image encryption system

2.2. ROI Selection Component

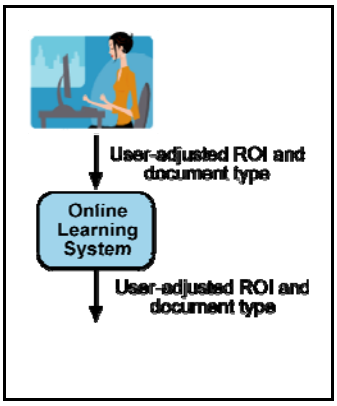
The ROI selection component used in the proposed system is based on previous research by Wong et al. [11]. The proposed ROI selection component extends upon the basic concepts presented in their paper by making use of an online learning model that is based on user feedback. Prior to describing the proposed ROI selection component, it is important to first provide an overview of the online learning model. In the online learning approach to machine learning, the learning system is adjusted dynamically to account for new data as it arrives into the system. This model of machine learning has a number of advantages. First, it allows the system to quickly evolve to handle new types of data. Second, it allows the system to improve its robustness to variations in existing data as more samples of that data type is fed into the system. As such, the system has the potential to improve incrementally with each subsequent sample. It is these characteristics that make an online learning model well suited to the system.

Given just the initial input document image (and optionally the document type), the proposed ROI selection component attempts to determine the appropriate ROI for the document image using an online learning system that is based on user feedback. The general architectural flow of the proposed semi-automatic ROI selection system is illustrated in Figure 2. The ROI selection system is divided into two phases:

a) ROI selection phase, and a
b) Feedback training phase.



(a) ROI Selection Phase



(b) Feedback Training Phase

Figure 2. Architectural flow of proposed ROI selection component

2.2.1 Online Learning System

The knowledge core of the ROI selection component is the online learning system, which is used in both the ROI selection phase and the feedback training phase. The online learning system is structured as a hierarchy tree. This structure is chosen for a number of reasons. First, it permits fast knowledge traversal and comparisons to be performed, particularly for situations where the general document type is known. Finally, it permits new knowledge to be easily

added to the online learning system. Each non-leaf node in the online learning system represents a particular document type, while each leaf node represents the knowledge retained from a sample document instance that was processed by the system.

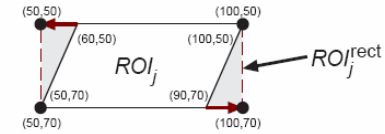
The knowledge stored within a leaf node can be seen as a concise and compact representation of a particular document image sample that can be used to identify ROI in similar documents. For the purpose of the proposed system, each leaf node contains the two sets of information extracted from a particular document image sample. The first set of information is the set of ROI and the corresponding authority levels selected for the regions. The second set of information stored in the leaf node is the coordinates and the corresponding local point descriptors for a set of interest points that represent the document image at a fixed image resolution as defined for the particular system. A fixed image resolution is used so that document comparisons between similar documents can be performed at the same scale. The set of interest points is sorted in the order of feature significance. For the proposed system, the set of interest points was extracted using a modified Harris corner detector described in [12] and the local point descriptor used for a particular interest point was an 11×11 patch of pixel values centered about the interest point.

2.2.2 ROI Selection Phase

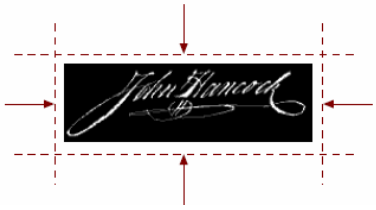
In the ROI selection phase, the document image inputted by the user is first scaled to same fixed image resolution used by the online learning system. The input document image is then processed using a feature extraction algorithm to obtain a set of interest points for the document. The feature extraction algorithm used is the same as that described in Section 2.2.1.

At the initial ROI identification stage, a document similarity comparison process is performed using the input document image and a set of searchable nodes from the online learning system. The set of searchable nodes is determined based on the optional document type specified by the user. If no document type is specified, all nodes within the online learning system are marked as searchable. If a document type is specified, then all leaf nodes under the specified document type are marked as searchable. Since the document type is typically known, the search space can be reduced. This leads to a reduction in the computational cost for the document similarity comparison process. The similarity comparison process used in the proposed system is the same as that proposed in [11], where a similarity score is calculated between the input document image and each searchable leaf node by matching local point descriptors using a normalized correlation similarity metric. The leaf node with the highest similarity score is selected as the closest match and the information contained in the leaf node is used at the final ROI generation stage to determine a final set of ROI for the input document image. Furthermore, if the user does not specify the document type, it is set as the document type of the matched leaf node.

At the final ROI identification stage, a number of different techniques proposed in [11] are used to fine-tune and fit the ROI information contained in the matched leaf node to the input document image. First, the ROI coordinates from the matched leaf node are mapped onto the input document image using an image registration algorithm. A bounding box algorithm is then used on each transformed ROI on the input document image to readjust the ROI into a rectangular shape. Finally, a ROI boundary adjustment algorithm is used to reduce the amount of unnecessary white space that is covered within each ROI. This final adjustment step produces a final set of ROI estimates, which are then sent back along with the document type to be shown to the user. The bounding box and ROI boundary adjustment algorithms are illustrated in Figure 3.



(a) Bounding box algorithm



(b) ROI boundary adjustment algorithm

Figure 3. Algorithms used in final ROI generation stage [11]

2.2.3 Feedback Training Phase

The key advancement made in the proposed ROI selection component over the algorithm in [11] is realized in the feedback training phase. The algorithm in [11] relies solely on offline training using a fixed set of training data created by a human expert. Therefore, a system based on the algorithm in [11] is only as good as the set of offline training data used to train the system. Furthermore, such a system does not improve with subsequent uses, as it does not learn dynamically from new input document image samples. To address this issue, an online learning system based on user feedback is utilized. In the feedback training phase, the user is shown the set of ROI estimates generated at the ROI selection phase as well as the determined

document type. The user can then adjust the generated ROI estimates to refine the ROI selection and adjust the assigned document type. This feedback is then submitted by the user to the system and used to train the online learning system.

Before the user feedback is used to train the online learning system, it is first determined whether the system would noticeably benefit from the knowledge provided by the user feedback. This is determined in the proposed system using a set of rules. The user feedback is used to train the system if either of these conditions holds true:

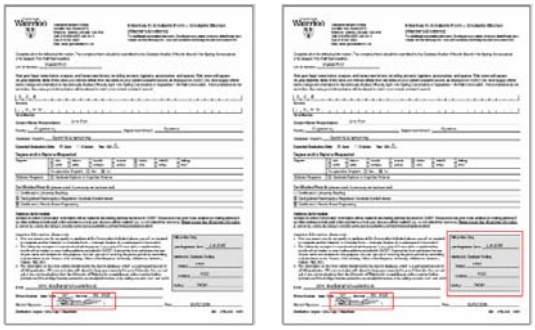
1. The document type determined by the system in the ROI selection phase does not match the document type specified in the user feedback.
2. The intersecting area between the generated ROI estimates and the ROI specified in the user feedback is less than 95% of the total area of the generated ROI estimates.

These conditions ensure that the online learning system only trains on user feedback when the proposed system will benefit from the training to improve accuracy. This helps to effectively reduce the size of the online learning system by storing only what is necessary for producing accurate ROI estimates.

If the user feedback is deemed useful for training the online learning system, a leaf node is created using the refined ROI information provided in the user feedback as well as a set of interest points and their corresponding local point descriptors extracted from the input document image. If the document type specified in the user feedback exists in the online learning system, the leaf node is then added to that document type. If no such document type exists, a new document node is created and the leaf node is added to it.

2.3. ROI Image Encryption Component

In this component, the final ROI information generated by the ROI selection component is used to determine the information within the document image that needs to be encrypted. The proposed ROI image encryption scheme is performed using a series of passes, with each pass performing encryption at a particular level of authority. Assuming g authority levels, the encryption passes are performed sequentially, starting from the highest authority level (denoted as authority level 1) down to the lowest authority level (denoted as authority level g). Therefore, at pass i, all ROI that correspond to authority levels i and below (i.e., i-1, i-2, …, 1) are encrypted at the bit-stream level with cipher key K_i using a data encryption scheme such as RC4, SEAL [13], and AES. An example of the multi-pass ROI encryption scheme is illustrated in Figure 4. For example, if there were a total of g authority levels, a user must possess all g cipher keys and use them sequentially to decrypt and view all concealed content. This multi-pass encryption scheme is used in the proposed system to preserve backwards compatibility with a wide range of image formats. Unlike the JPSEC and MRC approaches described in Section 1, this approach supports many popular image formats such as JPEG, BMP, and PNG that do not provide native support for multiple layers and/or ROI. Therefore, by using a single image layer for all document content, backwards compatibility is preserved by the proposed scheme.



(a) First pass (b) Second pass

(c) Third pass

Figure 4. Example of multi-pass encryption scheme (ROI encrypted at each pass is indicated by a red rectangle)

A conventional encryption scheme is used in the proposed system for a number of reasons. First, unlike the simple pseudorandom noise approach used in [10], conventional encryption schemes have been proven to be cryptographically secure. Using a combination of strong encryption schemes and strong cipher keys causes the encrypted image content to appear as random noise. This is important for situations where a high level of security is required to ensure privacy, such as the storage of medical records in hospitals. The second reason is that the ROI that needs to be encrypted in a document image is typically relatively small compared to the full document image size. Hence, the actual computational cost of encryption is relatively low. This fact is particularly important for large document images, where bit-stream level encryption becomes computationally expensive. Therefore, it is

reasonable to perform conventional encryption techniques in such a situation to improve data security without adding a great deal of computational overhead. To ensure the uniqueness of the key stream if a stream cipher is used, a unique identifier is combined with the cipher keys. By doing this, different document images are encrypted using different key streams even if the same cipher keys are used, since the unique identifier of each document image is different.

Besides the greater level of flexibility in terms of access control, there are several technical advantages to the use of the proposed ROI image encryption scheme for document image encryption. First, the proposed algorithm encrypts only regions of interest in an image. It should be noted that as image content is encrypted, it becomes more difficult to compress the content. The proposed approach effectively improves the compression rate when compared to a fully encrypted document image because only portions of the image content are encrypted. Finally, since only parts of the image are encrypted, there is enough visible content available in the document image such that conventional content-based image retrieval systems can be used for document image searches.

Once the document image has been encrypted, it is necessary to store the ROI and authority level information. In image formats such as JPEG and TIFF, this information can be stored as meta-data in the image without compatibility issues since both formats support header data. However, popular image formats such as BMP and GIF do not provide native support for custom meta-data storage. To maintain backwards compatibility with these formats, the proposed system uses lossless image watermarking techniques to conceal the ROI and authority level information in the actual document image content. As such, no modifications are needed to the actual image format and therefore the encrypted document image can be viewed using standard image viewers without compatibility issues. In the case of lossy image compression formats such as JPEG, the lossless image watermarking techniques are applied after the lossy compression process.

3. Experimental Results

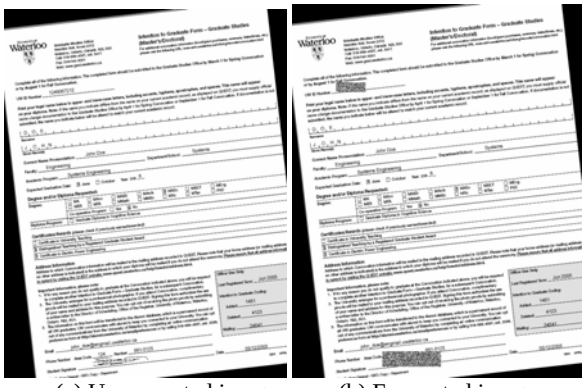
The proposed system was implemented for testing purposes using the RC4 stream cipher with 256-bit cipher keys. Stronger ciphers may be used depending on the requirements of the actual application. The document images are stored in the TIFF image format using the lossless PackBits compression algorithm. The test set consists of 5 documents of different document types from the University of Waterloo Registrar's Office, with two ROI selected for each document. To evaluate the compression rate achievable when typical documents are encrypted using the proposed ROI image encryption system, the test documents are encrypted using both the proposed ROI image encryption system and a conventional image encryption scheme that encrypts the entire document image. A summary of the results is shown in Table 1. It is observed from the experimental results that a high level of compression performance is achieved using the proposed algorithm. This is due to the fact that only necessary portions of the image are encrypted so the randomness in data distribution caused by the encryption process is minimal. Since data compression schemes rely heavily on finding patterns in the data, minimizing randomness in data distribution allows for higher compression rates.

Table 1
Compression Rate of ROI Encrypted Documents

Document Type	Proposed algorithm* (%)	Conventional approach* (%)
Course override	30.96	98.65
Cross registration	44.27	97.53
Special guest seating	17.38	98.58
Intent to graduate – undergraduate	22.16	98.58
Intent to graduate - graduate	27.35	98.40
Overall average	28.42	98.35

* Percentage relative to uncompressed original image size

To test the accuracy of the ROI estimates produced by the proposed system, ten test document images were generated from the "Intent to graduate – graduate" document as it was shown to have the lowest generated ROI coverage accuracy for the algorithm proposed in [11]. These test document images were distorted using random rotations of +/- 10^o and random x and y translations of +/- 5% of the respective document dimensions. The proposed system was able to achieve an average coverage accuracy of 95.38% after 6 test document images were processed using the proposed system. An example test trial is shown in Figure 5.



(a) Unencrypted image (b) Encrypted image

Figure 5. Example test trial

4. Conclusion

In this paper, we have introduced a semi-automatic system for multi-level ROI document image encryption using an online learning system. The proposed system produces encrypted document images that are backwards compatible with existing image formats. Experimental results show that the proposed system produces accurate ROI estimates that help reduce the amount of manual labor required by the user. Furthermore, it is demonstrated that ROI encrypted document images produced by the proposed system can be compressed at a higher rate than if a conventional encryption algorithm is used. It is our belief that this system can be successfully integrated into image-based document management systems to provide a greater level of flexibility.

Acknowledgements

This research has been sponsored in part by Epson Canada and the Natural Sciences and Engineering Research Council of Canada.

References

[1] P. Dang and P. Chau, "Image Encryption for Secure Internet Multimedia Applications," in *IEEE Transactions on Consumer Electronics*, vol. 46, no. 3, pp. 295-403, August 2000.

[2] Q. Hou and Y. Wang, "Security traffic transmission based on EZW and AES," in *Proceedings of IEEE Intelligent Transportation Systems*, vol. 1, pp. 86-89, 2003.

[3] I. Ziedan, M.M. Fouad, and D.H. Salem, "Application of data encryption standard to bitmap and JPEG images," in *Proceedings of the Twentieth National Radio Science Conference*, C16, pp. 1-8, March 2003.

[4] R. Chen, J.L. Lai, and W.K. Lu, "Image Encryption Using Progressive Cellular Automata Substitution and SCAN," in *Proceedings of the International Symposium on Circuits and Systems*, p. 1690-1693, May 2005.

[5] M. Salleh, S. Ibrahim, and I.F. Isnin, "Enhanced chaotic image encryption algorithm based on Baker's map," in *Proceedings of the International Symposium on Circuits and Systems*, vol. 2, p. 508-511, May 2003.

[6] H. Zhang, W.X. Feng, L.Z. Hui, L.D. Hai, and L.Y. Chou, "A New Image Encryption Algorithm Based on Chaos System," in *Proceedings of the IEEE International Conference on Robotics, Intelligent Systems and Signal Processing*, vol. 2, p. 778-782, 2003.

[7] C. Li, S. Li, D. Zhang, and G. Chen, "Cryptanalysis of a chaotic neural network based multimedia encryption scheme," in *Advances in Multimedia Information Processing-PCM 2004 Proceedings, Part III*, Volume 3333 of Lecture Notes in Computer Science, Springer-Verlag, pp. 418-425, 2004.

[8] S. Li and X. Zheng, "Cryptanalysis of a Chaotic Image Encryption Method," in *Proceedings of the International Symposium on Circuits and Systems*, vol. 2, pp. 708-711, 2002.

[9] F. Dufaux, S. Wee, J. Apostolopoulos, and T. Ebrahimi, "JPSEC for secure imaging in JPEG 2000", in *Proceedings of SPIE: Applications of Digital Image Processing XXVII*, vol. 5558, pp. 319-330, 2004.

[10] F. Dufaux, and T. Ebrahimi, "Video surveillance using JPEG 2000", in *Proceedings of SPIE: Applications of Digital Image Processing XXVII*, vol. 5558, pp. 268-275, 2004.

[11] A. Wong and W. Bishop, "Expert Knowledge Based Automatic Regions-of-Interest (ROI) Selection in Scanned Documents for Digital Image Encryption," in *Proceedings of the Third Canadian Conference on Computer and Robot Vision*, pp. 51, 2006.

[12] J. Noble, Descriptions of Image Surfaces, D.Phil. Thesis, Robotics Research Group, Department of Engineering Science, Oxford University, 1996.

[13] P. Rogaway and D. Coppersmith, "A software optimized encryption algorithm," in *Journal of Cryptology*, 11(4):273-287, 1998.