Adaptive Perceptual Degradation Based on Video Usage

Alexander Wong
Department of Electrical and Computer Eng.
University of Waterloo
Waterloo, Ontario, Canada, N2L 3G1
a28wong@engmail.uwaterloo.ca

William Bishop
Department of Electrical and Computer Eng.
University of Waterloo
Waterloo, Ontario, Canada, N2L 3G1
wdbishop@uwaterloo.ca

Abstract

The emergence of online digital media sales services has given rise to the issue of online digital rights management (DRM). This paper presents a novel approach to online digital rights management and video content protection through the use of adaptive perceptual degradation based on usage. The proposed system facilitates the sharing of purchased digital videos between users while offering an incentive for users to acquire original video content through legal means. This is accomplished by adaptively degrading the quality of the digital video content based on a number of usage factors such as the number of times viewed and the number of copies made. Experimental results demonstrate the effectiveness of this system in striking a balance between the freedom of file sharing and content protection.

1. Introduction

One of the hottest trends in the last few years is the sale of digital media content such as music and videos over the Internet. One of the most successful cases of this new technological trend is Apple's iTune music store, which sells digital music and video content to consumers that can be played back on their computers and/or portable devices. Recently, plans have been made by several large companies to provide video content such as TV shows and movies for sale on the Internet. These online video sales services include Google Video, MovieLink, and CinemaNow. There are a number of important benefits to digital media sales and distribution services over existing retail venues. First, since no physical media needs to be manufactured, stored, and distributed to individual retailers, the price of the video content can be reduced significantly for consumer purchases. Second, a large library of movies can be made available to consumers, especially older titles that are hard to find in retail. Finally, such services allow users to purchase videos quickly and easily from anywhere in the world.

One of the biggest issues that online video sales services currently face is the problem of digital rights management (DRM) and content protection. It is necessary that such services protect their digital content from piracy and illegal distribution. This is particularly a problem given the widespread abuse of file sharing networks and technologies such as Kazaa and BitTorrent. However, it is also necessary that such services do not impose restrictions on the purchased video content that severely limit the way videos can be watched. For example, Google Video does not currently permit purchased video content from being copied or moved from one computer to another. Instead, the video content must be re-downloaded if a user wishes to view the content on another computer. While this is acceptable for short videos, it is very problematic for full-length movies that require a long time to download. Other services such as MovieLink and CinemaNow restrict the viewing of video content to a fixed number of computers. Such strict DRM restrictions can detract consumers from purchasing video content through such services. Therefore, a different approach is desirable to provide a satisfying user experience while providing a reasonable level of content protection.

This paper introduces a novel approach to online digital rights management and video content protection through the concept of adaptive perceptual degradation based on usage. The proposed system attempts to provide the flexibility of file sharing while still encouraging the legal purchase of video content. This is accomplished by adaptively degrading the quality of digital video content based on a number of usage factors such as the number of times viewed and the number of copies made. The goal of this paper is not to provide an unbreakable DRM system, but to provide a DRM system that provides a high quality user experience that gives users an incentive to purchase video content through legitimate video sales services. Evidence of comparable approaches to digital rights management for video content is not readily apparent in the existing literature on the subject.

In this paper, the theories underlying the proposed system are presented in Section 2. An implementation of the

Proceedings of the Eighth IEEE International Symposium on Multimedia (ISM'06)
0-7695-2746-9/06 $20.00 © 2006 IEEE

proposed system is provided in Section 3. The testing methods and test data are described in Section 4. Experimental results demonstrating the effectiveness of the proposed system are presented in Section 5. Finally, conclusions are drawn and future work is discussed in Section 6.

2. Theory

Prior to outlining the proposed system, it is important to introduce some of the theory behind the key concepts of the architecture. First, the concept of perceptual video degradation is presented as well as a method on how perceptual video degradation can be accomplished in current digital video compression standards. Finally, a number of different usage factors are described in detail alongside methods in which such factors can be used to adaptively degrade video content.

2.1 Perceptual Video Degradation

Traditionally, multimedia content such as images and video have been protected using conventional cryptographic techniques such as block and stream ciphering systems. A number of multimedia encryption algorithms using conventional techniques have been proposed and investigated [2, 3, 18]. In such schemes, the multimedia content is treated as a binary bitstream and all data is encrypted equally. Image-specific cryptographic techniques have also been proposed [15, 1, 14, 17]. Such techniques utilize the characteristics of images to yield better overall computational performance, but have been shown to be less secure than conventional techniques [7, 10, 9, 11].

When conventional and image-based cryptographic techniques are applied to encoded video content, the resultant decoded video content appears degraded to the point of being incomprehensible to the viewer. However, there are applications where it is desirable for video content to be visually degraded but remain comprehensible to the viewer. An example of such an application is Video-on-Demand (VOD) services, which may allow the consumer to view a visually degraded preview of the video content prior to purchase. Based on the degraded preview, the consumer may decide to purchase the full video content at a higher quality without visual degradation. To address this issue, a new class of multimedia encryption techniques known as perceptual cryptography was introduced [16, 12, 13, 8]. It is important to understand how perceptual video degradation can be applied to current digital video compression standards such as MPEG [4, 5] and more recently H.264/AVC [6]. Video compression algorithms try to represent video content in as few data bits as possible. Certain data components have a greater effect on the visual quality of the decompressed video output than others. This can be exploited to control the level of visual quality of the output video. Therefore, it is important to describe the different data components that are common to modern block-transform based video compression algorithms as well as their affect on perceptual quality.

Most popular block-transform based video compression techniques are based on the same fundamental concepts. As such, a number of common data components are present in the encoded video data produced by compression standards such as MPEG and H.264/AVC. The key common data components are:

1. DC coefficients
2. AC coefficients
3. Motion vectors

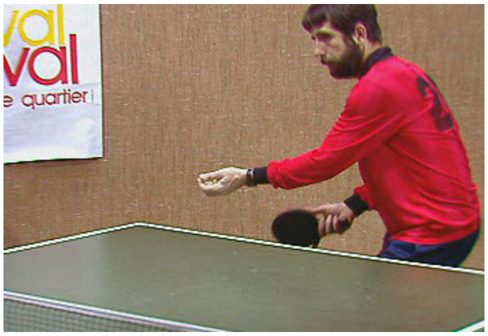
The DC coefficients represent the average energy of a video frame. The AC coefficients represent the details of a video frame. The motion vectors represent the estimated movement of blocks from one video frame to the next and they are used in video motion estimation to reduce the amount of data that needs to be stored. The human vision system is more sensitive to degradation at the lower spatial frequencies than at the higher spatial frequencies. Therefore, DC coefficients have a greater impact on the perceptual quality of video content than the AC coefficients. An example of this is illustrated in Fig. 1. Furthermore, corruption of the AC coefficients at the lower frequencies results in more noticeable perceptual degradation than those at the higher frequencies. This characteristic of the human vision system is very important for controlling the perceptual degradation of video content. For example, if minor perceptual degradation in video content is desired, only the AC coefficients at the higher frequencies need to be scrambled. To increase the level of perceptual degradation in the video content, more AC coefficients at the lower frequencies and/or the DC coefficients need to be scrambled.

For the proposed system, the data components that are potential candidates for degradation are divided into three groups:

1. DC coefficients
2. low frequency AC coefficients
3. high frequency AC coefficients

Breaking the AC coefficients equally into 2 separate groups provides a reasonable level of granularity for controlling the perceptual degradation applied to video content without significantly increasing the complexity of the system. Similar to the approach described in [8], the quality control factors used in the system indicate the probability that the sign bit of a coefficient within a certain group will

Proceedings of the Eighth IEEE International Symposium on Multimedia (ISM'06)
0-7695-2746-9/06 $20.00 © 2006 IEEE



(a) Original Video Content

(b) Scrambled DC Coefficients

(c) Scrambled AC Coefficients

Figure 1. Impact of scrambling DC and AC coefficients on perceptual quality

be scrambled. Only the sign bits are scrambled to reduce the computational overhead of the perceptual degradation process. A total of three control factors are used in the proposed system:

1. p_{dc}: probability that a DC coefficient is scrambled
2. p_{lo}: probability that a low frequency AC coefficient is scrambled
3. p_{hi}: probability that a high frequency AC is scrambled

2.2 Video Usage Factors

As discussed in Section 1, current methods of digital rights management and content protection used in online video sales services are very restrictive in the way a user can view the purchased video content. These DRM restrictions can severely deter consumers from using such services because they do not cater to user needs. Therefore, it is important to analyse how people prefer to use their purchased video content.

First, people typically like to be able to watch purchased video content at a number of different locations. For example, an individual may purchase a movie using their home desktop computer, but would like to watch the movie on their laptop computer. Therefore, he or she would like the ability to move the video content from one location to the next without having to re-download it. Second, people like to share multimedia content with others. This is evident by the popularity of peer-to-peer file sharing networks and technologies such as Kazaa and BitTorrent, as well as the popular practice of loaning purchased DVDs to friends and family. Therefore, people would like the ability to share their purchased video content with other people. While it is not feasible from a business perspective to give a user complete freedom to share their purchased video content with everyone else, it is also important not to restrict this activity completely. As such, a balance between content protection and freedom of usage is needed.

One approach to addressing these issues is through the concept of controlling perceptual quality of purchased video content based on usage factors. The inspiration for this concept is the VHS tape system. Due to the technological limitations of VHS tape systems, the perceptual quality of VHS video content is implicitly altered by two main video usage factors. The first main usage factor is the number of times a video has been viewed. With frequent use, the perceptual quality of a VHS video tape degrades naturally. The second main usage factor is the number of times that a video has been copied. Each copy of a VHS video tape is slightly worse than the source VHS video tape in terms of visual quality. Taking a cue from the VHS tape system, digital video content purchased through online sales services

Proceedings of the Eighth IEEE International Symposium on Multimedia (ISM'06)
0-7695-2746-9/06 $20.00 © 2006 IEEE

can similarly be altered based on these two video usage factors and perceptual degradation techniques.

In the proposed system, the perceptual quality of the purchased video content is controlled by two video usage factors:

1. The number of times the video content has been viewed
2. The number of times the video content has been copied

When a user views the purchased video content, the usage is recorded by the video player application. When a user wishes to share a copy of the purchased video content with another user, the usage information is used to determine the level of perceptual degradation suffered by the copied version of the video content. Furthermore, each copy of the video content is worse than the previous copy of the video content in terms of perceptual quality. Finally, the source video content is perceptually degraded based on the number of copies shared by the user.

The proposed system can be illustrated using the following example, as shown in Fig. 2. Alice purchases a movie from an online video sales service and watches it several times. Alice decides to share a copy of the movie with Bob and Margaret. The copies of the movie are created by making a copy of Alice's copy of the movie and performing perceptual degradation based on the number of times Alice has watched the movie. As such, the copies of the movie are inferior to Alice's original copy of the movie. If Alice wishes to then make a copy of the movie for Jim, her own copy of the movie will be degraded to act as an incentive for her not to make further copies. If Margaret makes a copy of her copy of the movie for Jill, the copy of the movie that Jill receives is further degraded and therefore inferior when compared to Margaret's copy of the movie. Since the quality of Jill's copy is noticeably degraded relative to Alice's copy, Jill has an incentive to purchase an original copy of the movie from the online sales service.

There are a number of advantages to this proposed system. First, it permits a user to freely share his or her purchased content with others. Second, it gives a user incentive not to share too many copies of his or her purchased content. Third, it allows other users to view the purchased content at a degraded level and thus provide incentive to purchase a copy of the video at a higher quality. Therefore, this system achieves the goal of providing a high quality user experience that strikes a balance between the freedom of media sharing and content protection.

3 Implementation

Based on the theory presented in Section 2, the following system was implemented. The basic implementation

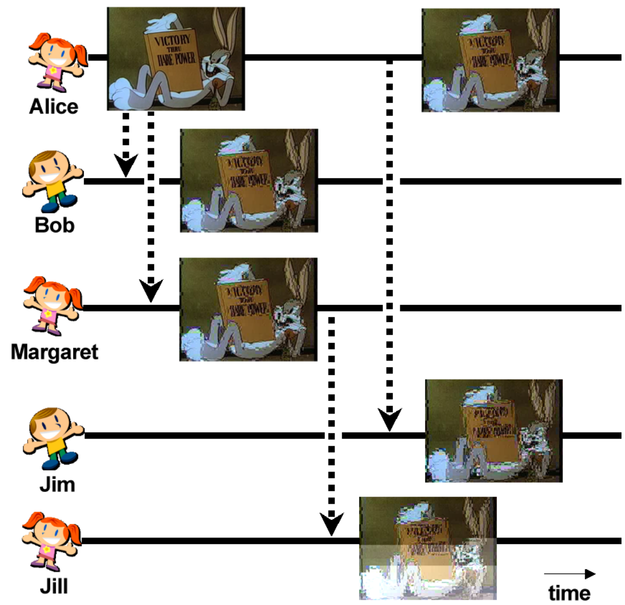


Figure 2. Example of usage-based adaptive perceptual degradation

of the system consists of both client and server software. When User A completes a purchase transaction with the video sales service, the server generates a copy of the video, referred to as Copy A, and sends it to the user. Copy A is an exact duplicate of the original video with no degradation. The checksum of Copy A is stored in User A's profile on the server for validation purposes. Each time User A views Copy A through the client, the checksum is sent back to the server to validate that the user has permission to view Copy A. Furthermore, data is sent back to the server to record the number of times that Copy A has been viewed (V_A) and the number of times that Copy A has been copied and shared with another user (C_A). When User A wishes to share a copy of Copy A with User B, a request is sent to the server. The server then calculates a set of control factors used for making the copy based on V_A and C_A:

$$p_{\mathrm{dc},B} = p_{\mathrm{dc,base}} + p_{\mathrm{dc,inc}} \times \max(\lfloor \frac{V_A}{D_{\mathrm{dc,copy}}} \rfloor, \lfloor \frac{C_A}{D_{\mathrm{dc,copy}}} \rfloor) \quad (1)$$

$$p_{\mathrm{lo},B} = p_{\mathrm{lo,base}} + p_{\mathrm{lo,inc}} \times \max(\lfloor \frac{V_A}{D_{\mathrm{lo,copy}}} \rfloor, \lfloor \frac{C_A}{D_{\mathrm{lo,copy}}} \rfloor) \quad (2)$$

$$p_{\mathrm{hi},B} = p_{\mathrm{hi,base}} + p_{\mathrm{hi,inc}} \times \max(\lfloor \frac{V_A}{D_{\mathrm{hi,copy}}} \rfloor, \lfloor \frac{C_A}{D_{\mathrm{hi,copy}}} \rfloor) \quad (3)$$

where $p_{\mathrm{dc,base}}$, $p_{\mathrm{lo,base}}$, and $p_{\mathrm{hi,base}}$ are the base probabilities for each group, $p_{\mathrm{dc,inc}}$, $p_{\mathrm{low,inc}}$, and $p_{\mathrm{hi,inc}}$ are the probability increment sizes for each group, and $D_{\mathrm{dc,copy}}$, $D_{\mathrm{lo,copy}}$, and $D_{\mathrm{hi,copy}}$ are the number of views before a copy of Copy A is degraded to the next level for each of

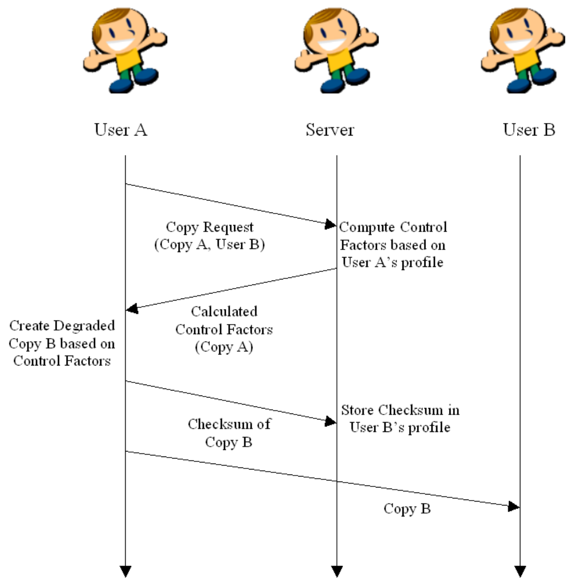


Figure 3. Example overview flow of the copying process

the groups. When Copy A is viewed repeatedly, the coefficients for the copies are altered accordingly to degrade the copies. Based on C_A, the control factors used for degrading Copy A are calculated as follows:

$$p_{\mathrm{dc},A} = p_{\mathrm{dc,inc}} \times \left\lfloor \frac{C_A}{D_{\mathrm{dc,source}}} \right\rfloor \quad (4)$$

$$p_{\mathrm{lo},A} = p_{\mathrm{lo,inc}} \times \left\lfloor \frac{C_A}{D_{\mathrm{lo,source}}} \right\rfloor \quad (5)$$

$$p_{\mathrm{hi},A} = p_{\mathrm{hi,inc}} \times \left\lfloor \frac{C_A}{D_{\mathrm{hi,source}}} \right\rfloor \quad (6)$$

where $D_{\mathrm{dc,source}}$, $D_{\mathrm{lo,source}}$, and $D_{\mathrm{hi,source}}$ are the number of copies shared before the source copy is degraded to the next level for each of the groups. Therefore, the more copies shared, the greater the degradation. The server sends the control factors to the client of User A and a copy of Copy A is created by scrambling the sign bits of DC and AC coefficients in Copy A based on the control factors. The copy of Copy A is referred to as Copy B. Copy A is also degraded based on the control factors to deter users from excessively copying their video content. The new checksum for Copy A is sent back to the server to replace the previous checksum. The checksum for Copy B is also sent back to the server and is added to User B's profile. Copy B can then be shared with User B. An overview of this copying process is illustrated in Fig. 3. Subsequently, User B can use Copy B to create a further degraded Copy C for User C using a similar process.

For testing purposes, the parameters of the system were configured as follows:

$$\begin{aligned}
p_{\mathrm{dc,base}} &= 0\\
p_{\mathrm{lo,base}} &= 0\\
p_{\mathrm{hi,base}} &= 0.3\\
p_{\mathrm{dc,inc}} &= 0.01\\
p_{\mathrm{lo,inc}} &= 0.05\\
p_{\mathrm{hi,inc}} &= 0.1\\
p_{\mathrm{dc,copy}} &= 3\\
p_{\mathrm{lo,copy}} &= 1\\
p_{\mathrm{hi,copy}} &= 1\\
p_{\mathrm{dc,source}} &= 4\\
p_{\mathrm{lo,source}} &= 2\\
p_{\mathrm{hi,source}} &= 2
\end{aligned}$$

In a practical situation, the parameters can be configured based on the needs of the video sales service provider. The parameters used for testing are simply examples chosen to demonstrate the effectiveness of the proposed system.

4 Testing Methods

To demonstrate the video scrambling capabilities of the proposed system, the system was tested using segments from four different MPEG compressed video clips. A brief description of each video segment is provided below.

Animation 1: A segment from a traditional animated cartoon.

Animation 2: A segment from a 3D animated cartoon.

Live 1: A segment from a US Election debate.

Live 2: A segment from the 'Tennis' video.

The following scenarios were tested for each segment:

1. First user copy of the original video content
2. Second user copy of the original video content
3. Third user copy of the original video content
4. Sixth user copy of the original video content

To judge the visual quality of the scrambled video content quantitatively, the average PSNR was measured for each test video clip.

5 Experimental Results

The average PSNR for each scenario is shown in Table 1. It can be observed that the degradation on the video content

Table 1. Average PSNR of video content for test scenarios

Scenario	PSNR (dB)			
	Animation 1	**Animation 2**	**Live 1**	**Live 2**
First User Copy	28.8987	35.9591	25.5528	30.9752
Second User Copy	25.0294	30.4780	22.3275	26.8409
Third User Copy	20.2709	19.4823	13.6470	18.0005
Sixth User Copy	11.9269	17.0881	12.0365	13.1700

increases progressively as the number of copies made increases.

Fig. 4, Fig. 5, and Fig. 6 show decoded video frames from the degraded test videos under the each of the four different scenarios. It can be observed that the perceptual quality is still good for the first copy with only minor degradation of the details. The resultant perceptual quality of the third copy of the test videos is very poor and suitable only for previewing the original content. For many users, the degradation of video quality may serve as an incentive to purchase the original video content with higher quality through an online video sales service. The degradation of video quality also serves as a deterrent to subsequent duplication of the video content. This demonstrates the effectiveness of the proposed system in providing a mechanism for sharing video content while discouraging the excessive duplication of video content.

6 Conclusions and Future Research

This paper proposes a novel system for digital video content protection using the concept of adaptive perceptual degradation based on usage. The proposed system is highly flexible and adapts the level of degradation applied to shared video content based on a number of important usage factors such as the number of times viewed and the number of copies made. It is believed that the proposed system can be implemented into existing online video sales services to improve the experience of customers while maintaining a reasonable level of content protection. Future work may include the integration of the proposed system into digital set-up boxes.

Acknowledgements

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

References

[1] R. Chen, J. Lai, and W. Lu. Image encryption using progressive cellular automata substitution and scan. *Proceedings of IEEE ISCAS*, pages 1690–1693, 2005.

[2] P. Dang and P. Chau. Image encryption for secure internet multimedia applications. *IEEE Transactions on Consumer Electronics*, 46(3):295–403, 2000.

[3] Q. Hou and Y. Wang. Security traffic transmission based on ezw and aes. *Proceedings of IEEE Intelligent Transportation Systems*, 1:86–89, 2003.

[4] ISO/IEC. Coding of Moving Pictures and Associated Audio for Digital Storage Media up to About 1.5 Mbits/s - Part 2: Video, November 1993.

[5] ISO/IEC. Generic Coding of Moving Pictures and Associated Audio Information - Part 2: Video, November 1994.

[6] Joint Video Team of ITU-T and I. JTC 1. Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification, May 2003.

[7] C. Li, S. Li, D. Zhang, and G. Chen. Cryptanalysis of a chaotic neural network based multimedia encryption scheme. *Advances in Multimedia Information Processing-PCM 2004 Proceedings*, 3:418–425, 2004.

[8] S. Li, G. Chen, A. Cheung, and B. Bhargava. On the design of perceptual mpeg-video encryption algorithms, 2005.

[9] S. Li, C. Li, G. Chen, and X. Mou. Cryptanalysis of the rces/rses image encryption scheme, 2004.

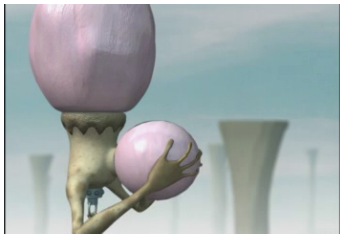
[10] S. Li and X. Zheng. Cryptanalysis of a chaotic image encryption method. *Proceedings of IEEE International Symposium on Circuits and Systems*, pages 708–711, 2002.

[11] S. Li and X. Zheng. On the security of an image encryption method. *Proceedings of International Conference on Image Processing*, pages 925–928, 2002.

[12] S. Lian, J. Sun, and Z. Wang. Perceptual cryptography on jpeg2000 compressed images or videos. *Proceedings of The Fourth International Conference on Computer and Information Technology*, pages 78–83, 2004.

[13] S. Lian, D. Ye, J. Sun, and Z. Wang. Perceptual mpeg-4 video encryption and its usage in video-on demand systems. *IEEE International Symposium on Consumer Electronics*, pages 83–86, 2004.

[14] M. Salleh, S. Ibrahim, and I. Isnin. Enhanced chaotic image encryption algorithm based on baker's map. *Proceedings of IEEE ISCAS*, 2:508–511, 2003.



(a) Original

(b) First User Copy

(c) Second User Copy

(d) Third User Copy

(e) Sixth User Copy

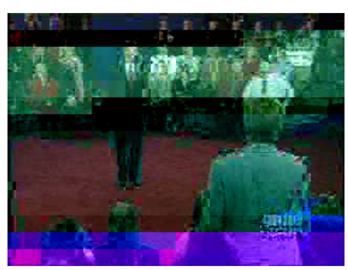
Figure 4. Degradation results for ’Animation 2’ video

(a) Original

(b) First User Copy

(c) Second User Copy

(d) Third User Copy



(e) Sixth User Copy

Figure 5. Degradation results for ’Live 1’ video

(a) Original

(b) First User Copy

(c) Second User Copy

(d) Third User Copy

(e) Sixth User Copy

Figure 6. Degradation results for 'Live 2' video

[15] Y. Seo, D. Kim, J. Yoo, D. Suijit, and A. Agrawal. Wavelet domain image encryption by subband selection and data bit selection. *Proceedings of World Wide Congress (3G Wireless)*, 2003.

[16] A. Tormbia and F. Mora. Perceptual cryptography of jpeg compressed images on the jflf bit-stream domain. *Proceedings of the IEEE Inlcmutiunal Symposium on Consumer Electronics*, pages 58–59, 2003.

[17] H. Zhang, W. Feng, L. Hui, L. Hai, and L. Chou. A new image encryption algorithm based on chaos system. *Proceedings of Robotics, Intelligent Systems and Signal Processing*, C16:1–8, 2003.

[18] I. Ziedan, M. Fouad, and D. Salem. Application of data encryption standard to bitmap and jpeg images. *Proceedings of National Radio Science Conference*, C16:1–8, 2003.