Journal of Kerbala University , Vol. 6 No.4 Scientific. 2008

Copyright Watermarking in MPEG 1 Files

Assistant Lecturer Ali Retha Hasoon Al-Moseaway
Computer Science Department, College of Science, Karbala University
Ali_retha@yahoo.com Or aliretha@gmail.com

Abstract:

We are living in a digital world in that extensive information exchange can be performed quickly and easily in digital format over the Internet or different storing devices in a cost-effective way. Digital data, such as music, image, video, text, e-mail, and so forth, are easily copied and transferred without any degradation. Concerns over ownership protection, data protection, and other security issues have therefore arisen. Digital watermarking is a general solution that can be used to identify illegal copying and ownership, authentication, or other applications by inserting information into the digital data in an imperceptible way. In order for a watermark to be useful, it must be perceptually invisible and have robustness against detecting processing and a variety of possible attacks by those who seek to pirate the material. In this paper we embedding watermark (text, icon) in MPEG1 file by using unused space in MPEG1 files to protect file from illegal copy and the same way could be used in secret communication (Steganography). The information embedded is hidden and the effect on the appearance and function of the work is minimized or invisible. It is supposed to survive even after digital-to-analog conversion, compression, or resizing.

الخلاصة :

نحن نعيش في عالم رقمي يتم فيه تبادل المعلومات الرقمية بسرعة وسهولة خلال الانترنيت أو أي وسيلة خزن و بكلفة ضئيلة للغاية. من السهولة استنساخ و نقل البيانات الرقمية مثل الموسيقى ،الصور، الفيديو،النصوص،محتوى البريد الالكتروني و غيرها و بدون أي تشويه . العلامة المائية هي إحدى الحلول لتمييز النسخ غير الشرعية و ذلك بحشر معلومات بصورة غير محسوسة في البيانات الرقمية و لكي تكون العلامة المائية مفيدة يجب أن تكون متينة ضد عمليات الكشف و الهجمات المحتملة لإزالتها. في هذا البحث المقترح ، نحاول تضمين علامة مائية (نص ، أيقونة) في ملف متعدد الوسائط من نوع وذلك باستغلال المساحات غير المستخدمة في ملف الوسائط المتعددة ، و ذلك لحماية الملف الهدف من الاستنساخ غير الشرعي علم بان هذه الطريقة يمكن استخدامها في عمليات التراسل السري بين الأفراد و المؤسسات . المعلومات المضمنة في الملف الهدف مخفية و إن تأثيرها على الملف هو اقل ما يمكن أو معدوم تماماً، و إن المعلومات سوف تنجو من عمليات التحويل من النظام الرقمي إلى التناظري و بالعكس، ضغط البيانات. أو إعادة تشكيل حجم الملف.

1 Introduction:

The best place to begin in this paper is with the bad news: there is no absolute way to prevent people from copying the digital versions of you text, your music, your movie, or your data. If a computer will interpret the data let someone use it, then that same computer can be programmed to grab the data and make a copy of it.

In this decade, the Internet, especially the World Wide Web has been successfully integrated into public and business domains. Recent surveys and public opinion polls have accented the value of the Internet. Traditional television and the Internet converge. In addition, the growth and integration of broadband access points, wireless and mobile technologies and the progress towards one-in-a-box device proves the significance of developing a legitimate marketplace for entertainment and business activities.

Digital networks and libraries, Internet services, and the disposition of non branded digital products within a global accessible network support lead to illegal copying, modification, and redistribution and the loss of high company sales and profits.

Particularly, the music and entertainment industry have struggled with illegal distribution over peer-to-peer and other networks for years. The International Intellectual Property Alliance (IIPA) has estimated that the annual worldwide trade loss due to copyright piracy is up to $10.2 billion excluding Europe and the United States in 2002 (IIPA, 2004). In 2003, the IIPA (special 301 report) estimated $20 to $22 billion in annual losses for the copyright industry. While the copyright industry generates the highest foreign sales for the U.S. economy, the annual loss in 2002 can be estimated up to 15% worldwide by $88 billion foreign sales. [1]

2 Copyright vs. Copyleft:

2.1 Copyright: is a set of exclusive rights granted by governments to regulate the use of a particular expression of an idea or information. At its most general, it is literally "the right to copy" an original creation. In most cases, these rights are of limited duration. The symbol for copyright is ©, and in some jurisdictions may alternately be written (c) [2].

2.2 Copyleft: is a play on the word copyright and is the practice of using copyright law to remove restrictions on the distribution of copies and modified versions of a work for others and require that the same freedoms be preserved in modified versions.

Most commonly, Copyleft is implemented by a license defining specific copyright terms applied to works such as computer software, documents, music, and art. Whereas copyright law, by default, automatically restricts the right to make and redistribute copies of an author's work, a Copyleft license uses copyright law to ensure that every person who receives a copy of a work has the same rights to study, use, modify, and also redistribute both the work, and derived versions of the work as long as the same license terms apply to all redistributed versions of the work. Thus, in a non-legal sense, Copyleft is the opposite of copyright.

Under a Copyleft form of copyright license, the restrictions imposed are that the work can be copied, modified or used in any subsequent work if, and only if, the author of that subsequent work agrees to grant the same Copyleft rights to the public to freely copy, use and modify the subsequent work. For this reason Copyleft licenses are known as reciprocal licenses.

Authors and developers use Copyleft to allow anyone to use, share and improve the work as a continuing process, disallowing people from sharing derived works with any new restrictions. For many people, Copyleft is a technique which uses copyright as a means of subverting the restrictions traditionally imposed by copyright on the dissemination and development of knowledge. This approach uses Copyleft primarily as a tool in a broadly scaled sniggling operation, whose aim is to permanently reverse such restrictions [2].

3 Watermarking Application:

3.1 Copyright Watermarking: Digital watermarking is described as a possibility to interface and close the gap between copyright and digital distribution. It is based on steganographic techniques and enables useful rights protection mechanisms. Digital watermarks are mostly inserted as a plain-bit sample or a transformed digital signal into the source data using a key-based embedding algorithm and a pseudo noise pattern.

The embedded information is hidden in low-value bits or least significant bits of picture pixels, frequency, or other value domains, and linked inseparably with the source of the data structure.

For the optimal application of watermarking technology, a trade-off has to be made between competing criteria such as robustness, nonperceptibility, nondetectability and security. Most watermarking algorithms are resistant to selected and application-specific attacks. Therefore, even friendly attacks in the form of usual file and data modifications can easily destroy the watermark or falsify it [3].

3.2 Copy Protection and Device Control: Digital watermarks can be used to enable copy control devices. In this combination, the recording device scans the digital data stream for an existing watermark and enables or disables the recording action for a specific movie or stream. Such technology could extend the pay-per-view concept and close the gap between the applied cryptographic approach and its usability. However, the implementation in consumer devices seems to be possible in using the same procedures applied when inserting the Macro Vision and CSS DVD copy mechanisms. By limiting available DVDs to CSS-compliant DVD players, manufacturers had to integrate new encoders that are secured by patent law regulations in their devices to maintain market position [3].

3.3 Broadcast Monitoring:

The production cost of broadcasting material, such as news, shows, and movies, are enormous and can be $100,000 per hour and more. Therefore, it is important for production companies, for example, Warner Bros., Miramax, and Universal Pictures, to secure their intellectual property and not permit illegal rebroadcasting activities. In this case, digital watermarking can enable Digital Watermarking:

Introduction 11 technical frameworks such as TALISMAN, which automatically monitor broadcasting streams at satellite nodes all over the world and identify illegally broadcasted material. Furthermore, TV stations can be monitored and the unlawful use could be tracked and debited individually. In 1997, two Asian broadcasting stations had been identified for intentionally overbooking their advertising time and making customers pay for unplayed broadcasting time.Computer systems can be used for tracking and monitoring advertisement activities on broadcasting channels and for examining advertisement deals. Nielsen Media Research and Competitive Media Reporting offer such computer systems [3].

3.4 Data Authentication: Digital watermarking is often used to prove the authenticity of a specific digital document. The digital watermark contains information that can be used to prove that the content has not been changed. Any such operation on the file destroys or changes the integrated watermark.

If the watermark information can be extracted without errors, the authenticity can be proved. In order to design an effective watermarking algorithm, the watermarking data or procedure can be linked to the content of the digital document. Such watermarks are called fragile watermarks or vapor marks [3].

3.5 Further Applications: Though the main application of digital watermarking is to secure the intellectual property, it can also be used in the medical field. In using digital watermarks as container for information about patients and their diagnosis, medical images, for example, X-ray or nuclear magnetic resonance (NMR) tomography could be automatically associated with the patient. Furthermore, digital watermarking could be used to save context or meta-information in source documents. In using special watermarking agents, generic search machines are able to retrieve such information and can offer time-based media documents as a result [3].

4 MPEG STANDARDS:

To date, MPEG has developed five major sets of technical standards as listed below. Work on most of these is still going on, including maintenance on the earlier MPEG standards. Since its start in May 1988, MPEG has developed the following sets of technical standards2:

• ISO/IEC 11172 (MPEG-1), entitled ‘Coding of Moving Pictures and Associated Audio at up to about 1.5 Mbps’

• ISO/IEC 13818 (MPEG-2), entitled ‘Generic Coding of Moving Pictures and Associated Audio’

• ISO/IEC 14496 (MPEG-4), entitled 'Coding of Audio–Visual Objects'
• ISO/IEC 15938 (MPEG-7), entitled 'Multimedia Content Description Interface'
• ISO/IEC 21000 (MPEG-21), entitled 'Multimedia Framework'

MPEG follows the principle that a standard should specify as little as possible while guaranteeing interoperability. This increases the lifetime of the standard by allowing new technology to be introduced and encouraging competition. For example, MPEG coding standards only specify bit stream syntax and semantics as well as the decoding process; nothing is said about how encoders are supposed to create compliant bit streams.

This leaves developers much freedom to optimize their choice of encoding tools, for example, motion estimation, rate control, psychoacoustic model, to maximize the performance for their target applications and content. It also allows the standard's performance to increase over time, following the developers' increasing knowledge on how to best exploit the coding tools. Lastly, it enables the various players to compete in quality and price, while providing mutual interoperability [4].

5 THE MPEG-1 STANDARD:

The MPEG-1 standard represents the first generation of the MPEG family; it was set in the period from 1988 to 1991. At the end of the eighties, with ITU-T recommendation H.261 on video coding for communication purposes almost finalized, it became clear that the same coding technology could provide a digital alternative to the widely spread analogue video cassette player. MPEG-1's goal was to provide a complete audio-visual digital coding solution for digital storage media such as CD, DAT and optical drives. Since CD was the major target, the standard was optimized for 1.5 Mbps bit rate range, but the standard works at lower and higher bitrates as well. To compete with analogue videotape recorders, the MPEG-1 solution had to provide the special access modes typical of these devices such as fast forward, fast reverse and random access, and the video quality had to be at least comparable to VHS quality.

This highlighted that coding efficiency is not the only requirement of a coding scheme; there are other important functionalities, depending on the targeted area of use (e.g., random access, low delay and the object-based interactive functionalities built into the MPEG-4 standard).

MPEG usually develops audio and video coding standards in parallel, together with the multiplexing and synchronization specifications. Although designed to be used together, the individual specifications can also be used independently and with other tools, for example, a different video format can be used together with the MPEG-1 Systems and Audio solutions. For this reason, the MPEG standards are organized in Parts, each one defining a major piece of technology. MPEG-1 has five Parts:

• Part 1 – Systems: Addresses the combination of one or more data streams (MPEG-1 Video and Audio) with timing information to form a single stream, optimized for digital storage or transmission.

• Part 2 – Video: Specifies a coding format (video stream and the corresponding decoding process) for video sequences at bitrates around 1.5 Mbps. The target operational environment was storage media at a continuous transfer rate of about 1.5 Mbps, but the coding format is generic and can be used more widely. It supports interactivity functionalities or special access modes such as fast forward, fast reverse and random access into the coded bit stream. The coding solution is a typical hybrid coding scheme based on block-based Discrete Cosine Transform (DCT) applied to a single picture or to a prediction error obtained after temporal prediction (on the basis of one or two pictures) with motion compensation. DCT is followed by quantization, zigzag scan and variable length coding.

MPEG-1 Video only supports progressive formats and the number of lines is flexible.

• Part 3 – Audio: Specifies a coding format (audio stream and the corresponding decoding process) for monophonic (32–192 kbps) and stereophonic (128–384 kbps) sound.
This standard specifies three hierarchical coding layers – I, II and III – which are associated to increasing complexity, delay and efficiency. Layer III is more commonly known as MP3. These coding solutions are designed for generic audio, and exploit the perceptual limitations of the human auditory system, targeting the removal of perceptually irrelevant parts of the signal. They do not rely on a model of the signal source, like voice coders do.
• Part 4 – Conformance Testing: Specifies tests to check if bit streams (content) and decoders are correct according to the specifications in Parts 1, 2 and 3.
• Part 5 – Software Simulation: Consists of software implementing the tools specified in
Parts 1, 2 and 3; this is a Technical Report that has only informative value. In later MPEG standards, the so-called Reference Software became a normative Part of the standard [5].
MPEG-1 is still a very popular format for Internet video streaming and downloading.
It did not become widely used in the area of digital storage media, partly because the DVD (using MPEG-2) soon followed it; still, the MPEG-1–based Video CD (VCD) is selling to the millions in China. It is also well known that the success of MP3, needless to explain, set the stage for the ongoing revolution in digital music distribution.

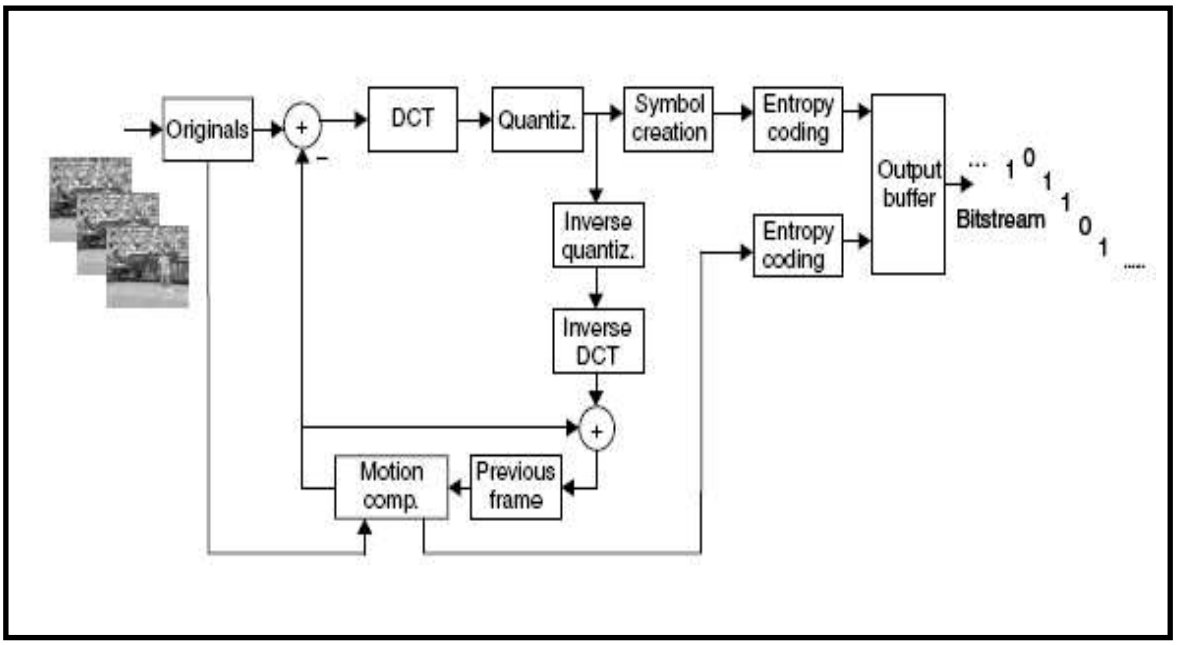

Figure (1) Simplified MPEG-1 Video encoder architecture [6]

6 The Watermark Embedding Methods:
In this research, there are two embedding methods used to insert the watermark in MPEG4 file.

Algorithm (1)
Embedding Watermark (Text) in padding space (unused space) of MPEG 1 File:
Input : MPEG 1 File, Watermark (Text)
Output : MPEG 1 File with Watermark
Step1: Load MPEG 1 File and Analyze it.

Step2: Load the watermarking String (text) and encrypt the watermark .
Step3: Locate MPEG 1 Padding Header.
Step4: Calculate the padding size.
Step5: If (the watermark size larger than padding space and Not End Of File (EOF)) then go to Step 3 else if (EOF) go Step 9 .
Step6: Put the Flags in beginning and ending of encrypted watermark.
Step7: Insert the encrypted Watermark in the padding space and replacing with the same blocks size of padding.
Step8: End.
Step9: Print Windows Message "The watermark is larger than padding space of this file " and end.

Algorithm (2)

Embedding Watermark (icon, small image) in padding space (unused space) of MPEG 1 File:
Input : MPEG 1 File, Watermark (icon, image).
Output: MPEG1 file with watermark (hidden) image.
Step1: Load MPEG 1 File and Analyze it.
Step2: Load the Watermark file (icon, small image).
Step3: Convert icon file (watermark file) to array of bits and calculate its size.
Step4: Locate MPEG 1 Padding Header.
Step5: Calculate the padding size.
Step6: If(the watermark size larger than padding space and Not End Of File (EOF)) then go to Step 4 else if (EOF) go Step 10 .
Step7: Put the Flags in beginning and ending of watermark.
Step8: Insert the Watermark (array of bytes) in the padding space and replacing with the same blocks size of padding.
Step9: End.
Step10: Print Windows Message "The watermark is larger than padding space of this file " and end.

7 The Watermark extracting Methods:

By inverse the steps in algorithms (1,2) we can extract the invisible embedded watermark from the MPEG1 file.

Algorithm (3)

Extracting Watermark (Text) form padding space of MPEG1 File:
Input : MPEG 1 File with Watermark
Output: MPEG1 File, Watermark (Text)
Step 1: Load MPEG1 File and Analyze it.
Step 2: Locate MPEG1 Padding Header and search it to locate Flags.
Step 3: If locate beginning Flags then read encrypted watermark (Text) until Ending Flags
Else If Not (EOF) go Step 2
Else go Step 6.
Step 4: Decrypt watermark and print it in window message.
Step 5: End.
Step 6: Print Windows Message "This File is not watermarked" and end.

Algorithm (4)

Extracting Watermark (Icon, Small Image) form padding space of MPEG1 File:
Input: MPEG1 File with watermark (hidden) image.

Output: MPEG1 File, Watermark (Icon, Small Image).
Input : MPEG1 File with watermark (hidden) image.
Output: MPEG1 File, Watermark (Icon, Small Image).
Step 1: Load MPEG1 File and Analyze it.
Step 2: Locate MPEG1 Padding Header and search it to locate Flags.
Step3: If locate beginning Flags then read watermark (icon, small image) until Ending Flags
Else If Not (EOF) go Step 2
Else go Step 9.
Step 4: Read the first three bytes of watermark (signature of file), if it (ico) go to step 5 else, if it (bmp) go to step 6.
Step 5: convert watermark (array of bytes) to binary file and then save this file in "c:\temp\master\watermark.ico".
Step 6: convert watermark (array of bytes) to binary file and then save this file in "c:\temp\master\watermark.bmp"
Step 7: Display watermark (icon, small image) in window message.
Step 8: End.
Step 9: Print Windows Message "This File is not watermarked" and end.

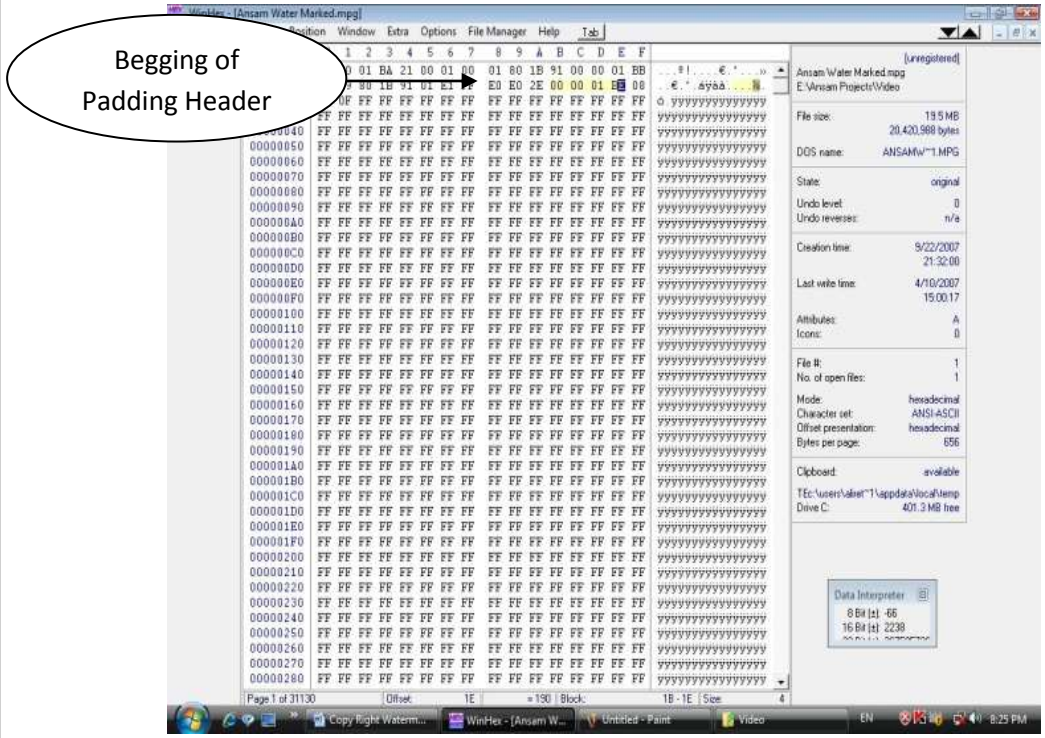

Figure (2) Begging of Padding by Hex Editor

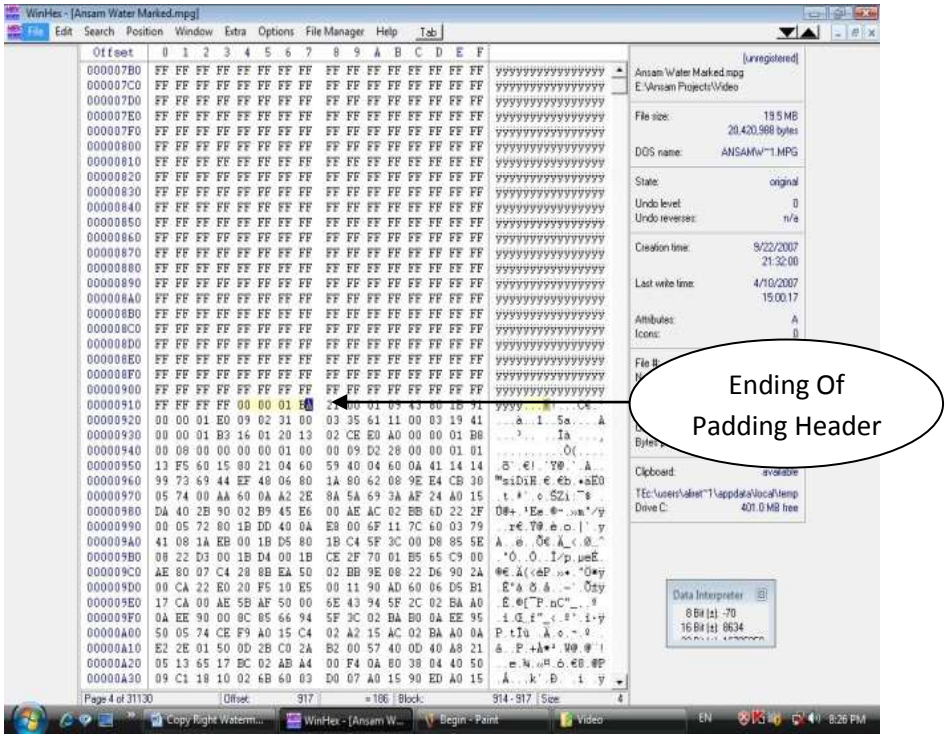

Figure (3) Ending of Padding by Hex Editor [7].

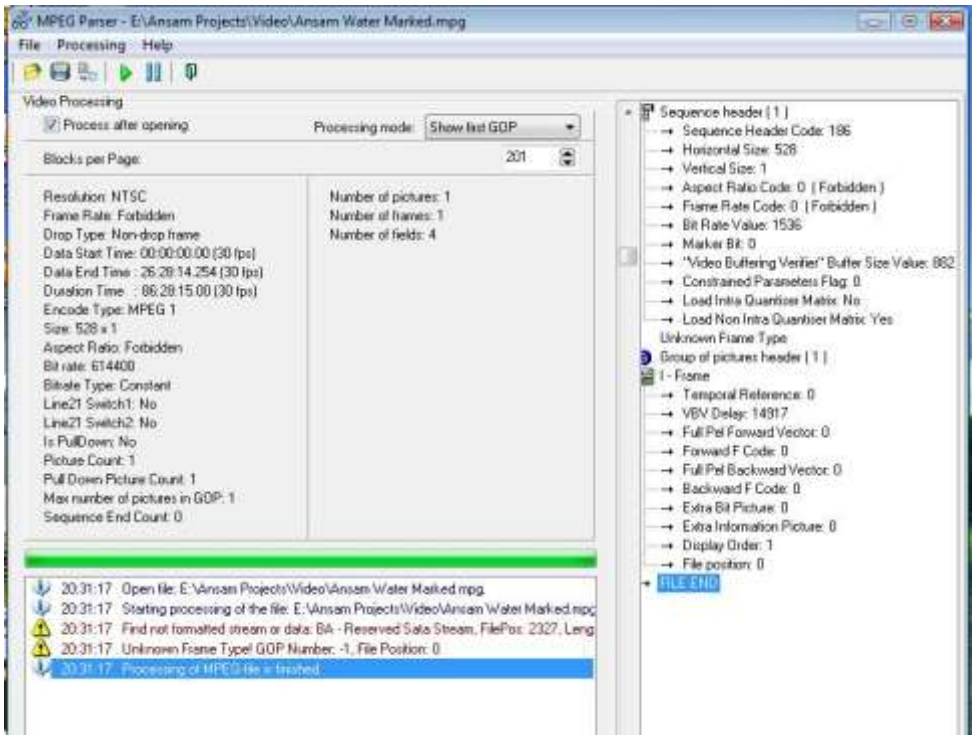

Figure (4) Analyzing MPEG1 File by using MPEG Parser 1.2 [8].

8 Conclusions:

1. The embedding by using unused spaces (padding spaces) will not increase the file size and this will increase the undetectability.
2. The proposed method of embedding technique can be used in steganography .
3. The proposed method can improve by imbedding cascade padding signature in padding space and this will confuse the possible attacker.
4. The proposed method is robustness against digital to analog conversion, compression, or resizing.
5. Because of the proposed method use the unused spaces then the resolution of MPEG1 file will not defect by embedding.

9 References:

[1] Juergen Seitz **," Digital Watermarking For Digital Media"** University of Cooperative Education Heidenheim, Germany, 2005.
[2] URL: www.wikipedia.com , 2008.
[3] Tino Jahnke , **" Digital Watermarking: An introduction"** , University of Cooperative Education Heidenheim, Germany, 2005.
[4] Jong-Nam Kim, Pukyong University, Republic of Korea, Byung-Ha Ahn ,Gwangju Institute of Science and Technology, Republic of Korea, **" MPEG Standards and Watermarking Technologies"**, 2007.
[5] Ling Guan , Sun-Yuan Kung ,Jan Larsen , **"MULTIMEDIA IMAGE and VIDEO PROCESSING "**, CRC Press , Washington, D.C. , USA, 2001.
[6] Ian S Burnett, University of Wollongong, Australia ,Fernando Pereira, Instituto Superior T´ecnico, Portugal ,Rik Van de Walle, Ghent University, Belgium Rob Koenen, MPEG Industry Forum, USA , **"The MPEG-21 Book ",** John Wiley & Sons Ltd, England ,2006.
[7] **Hex Editor 3.0**, By HHD, 2004.
[8] **MPEG Parser**, version 1.2.1, DVD Tools for professionals, Logic Software, 2004.