COST EFFECTIVE WATERMARKING IN THE SET TOP BOX

Joseph Oren
Cinea Inc. a Dolby Company

Abstract

Providing Cable consumers with premium (e.g. HD or early window) content via Video on Demand services is projected to become a key revenue source for system operators. Yet the Hollywood studios insist that before this content will be made available, enhanced content protection technologies must be deployed within the content distribution infrastructure [i]. Specifically, forensic watermarking, defined as the binding of unique traceable information to the video streams, is increasingly mentioned as an essential content protection layer, one that complements existing conditional access and digital rights management solutions. [ii] This paper describes how this new business requirement can be technically and economically fulfilled by watermarking technologies now reaching the market. Our focus will be on watermarking technology implemented in the consumer's equipment, commonly called the Set Top Box (STB).

DISCLAIMER

The author of this paper, Joseph Oren, is employed by Cinea Inc., a Dolby company. Cinea offers commercial products utilizing certain technologies described herein.

INTRODUCTION

DRM and CA technologies have made great strides toward system recognition of the rules agreed upon by content owners and consumers. The available mechanisms to enforce those rules, termed content protection, is, however, limited to encryption during transmission and storage. Once the content is rendered in a consumable form, its digital and analog representations become subject to copying and subsequent unauthorized redistribution (piracy). Figure 1 shows a simplified receiving device, with vulnerabilities identified. A real-world home network may spread these functions over several devices, each with analogous vulnerabilities.

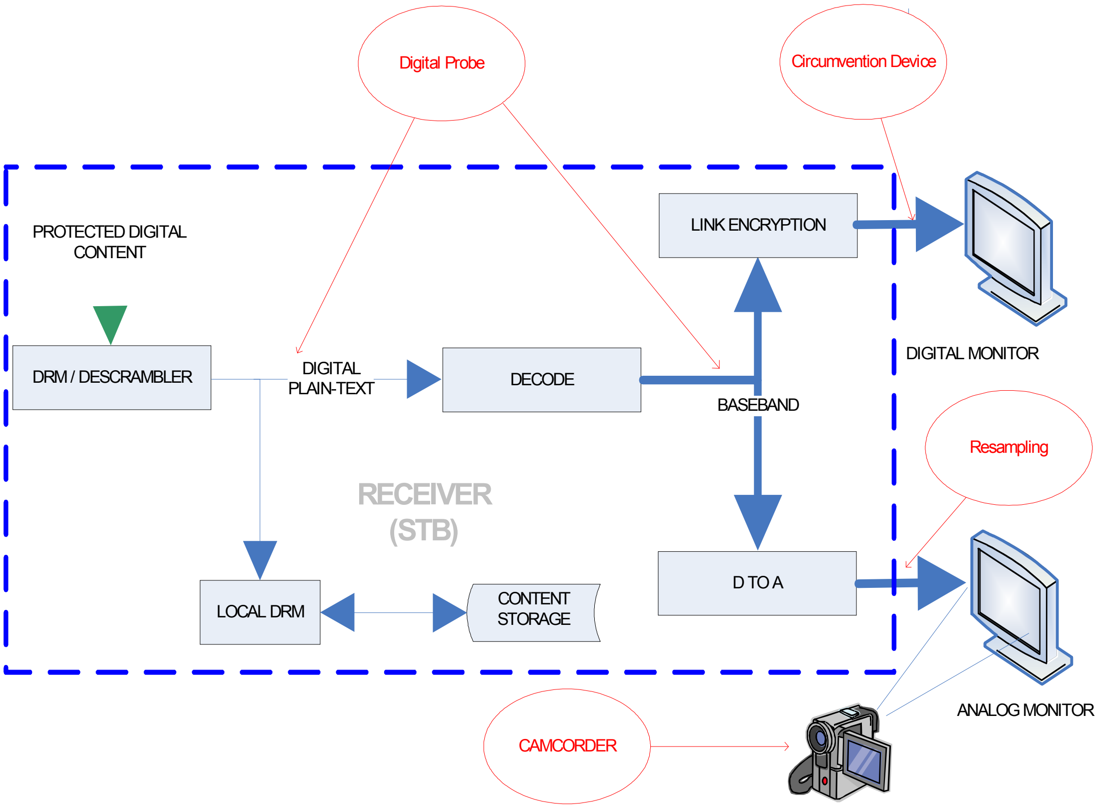

Figure 1 – Receiver with vulnerabilities identified

These vulnerabilities have fed the interest in forensic watermarking technology, whereby each instance of a content item is individuated by information to facilitate the tracing of the content back to its last legitimate holder. Tracing produces valuable evidence in identifying copyright violators. Further, since forensic watermarking is an investigative tool, as opposed to a control tool, it offers the potential to obviate some of the more complex and consumer hostile aspects of strong DRMs. While the DRM acts to constrain the user, actually challenging him to circumvent the technology, forensic watermarking deters piracy by introducing risk of exposure. In the case of large scale re-distribution, forensic watermarking facilitates identification of the point of compromise. With both DRM and watermarking available, a more balanced and appealing approach to content protection becomes possible.

Figure 2 – Forensic watermarking identifies the source of unauthorized distribution

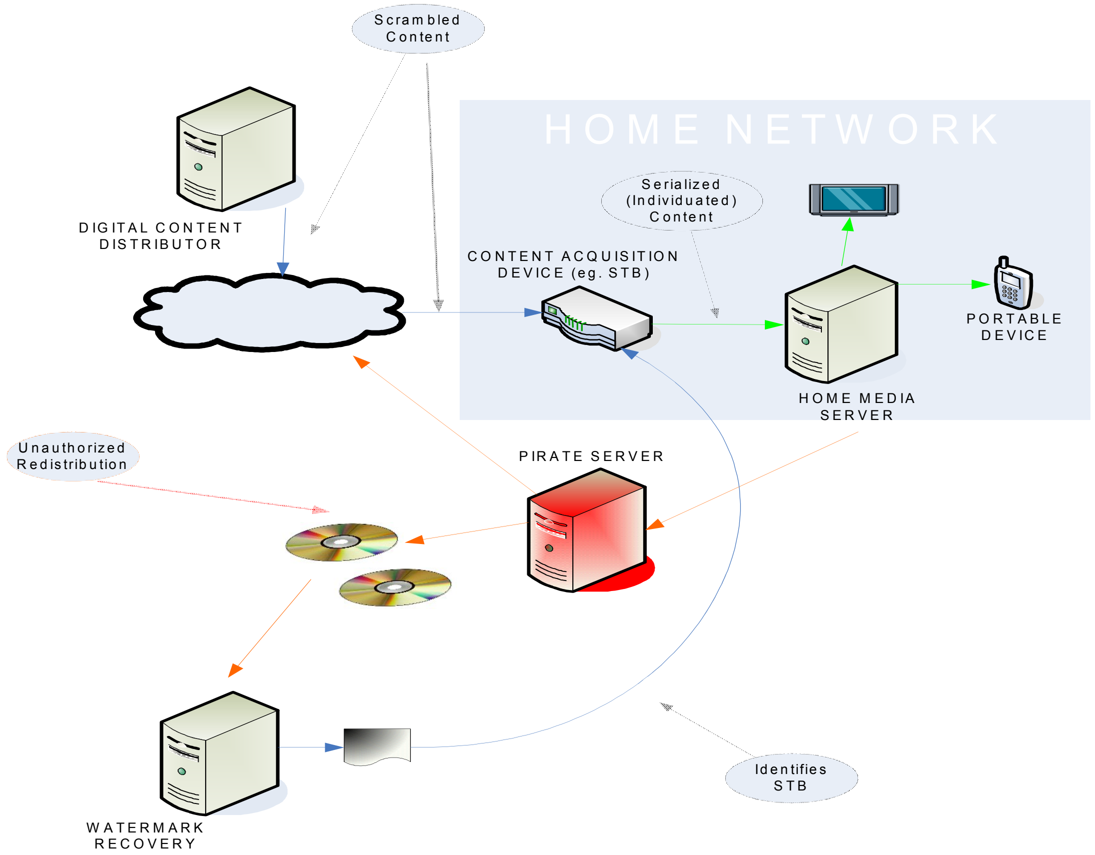
Scrambled Content
HOME NETWORK
Serialized (Individuated) Content
DIGITAL CONTENT DISTRIBUTOR
CONTENT ACQUISITION DEVICE (eg. STB)
PORTABLE DEVICE
HOME MEDIA SERVER
Unauthorized Redistribution
PIRATE SERVER
Identifies STB
WATERMARK RECOVERY

Figure 2 illustrates how forensic watermarking is used to investigate media piracy. In this example, the STB binds an identifying watermark to the content it acquires from the network. The watermark does not interfere with the consumer's enjoyment of the content. But, if the consumer chooses to circumvent the DRM and distribute unauthorized copies, the watermark can be used to determine his identity. Awareness, on the part of the consumer, of the possibility of exposure serves to discourage careless redistribution of the content.

Also called media serialization or content tracing, this paradigm is analogous to the use of serial numbers to track physical machinery or other valuable products. Watermarking technology provides a means of embedding such information into content, through subtle alterations to the content itself.

It is important to differentiate watermarking from simply appending identifying metadata to the content. Identifying metadata can be transparently excised from the content, while erasing a watermark requires specific manipulation of the content itself. Watermarking does not impair the content, but removal of a forensic watermark without impairing content quality is, by design, a very difficult task. We will examine the requirements for an effective forensic watermarking implementation and how those requirements can be addressed within the constraints of the STB.

WATERMARKING CONCEPTS AND TERMINOLOGY

In the broadest sense, forensic watermarking is a steganographic[iii] technique, one that embeds data into an instance of a cover work, in such a way that the data can subsequently be read (recovered) from copies of the watermarked cover work. The cover work may be any communication medium, but we will focus on digital video entertainment content. These principles may, however, be adapted for other media, such as the audio channels.

The process of binding the watermark to a content item is termed *watermark embedding*, and the additional data, in its embedded form, is the watermark itself. The process of reading the watermark from a copy of the content is termed *watermark recovery*.

In a data communications model, the watermark information is data to be communicated, and the cover work is a carrier signal. Indeed, the cover work carrying the watermark is often identified as the *host signal*. In the communications channel, the perceptible features of the content constitute noise that interferes with the watermark's information signal. It is important to recognize that watermark itself is embodied in changes to the cover work features, as opposed to just being ancillary data. Any faithful reproduction of the content (the carrier) will also carry the watermark data. The watermark signal may thus be viewed as modulating a noisy carrier signal, and thus becomes part of the cover work itself.

There are numerous watermarking technologies, with varying degrees of suitability for specific applications. The attributes commonly used to characterize a watermarking technology are as follows:

Perceptibility

Watermarks may either be apparent to the viewer, when the content is rendered, or disguised in such a way that the viewer is unlikely to notice the presence of the watermark. Perceptible watermarks are commonly used to proclaim ownership, exemplified by the visible logo appearing in many network broadcasts. In general, watermarks fall along a continuum of perceptibility, according to the needs of the users and the capabilities of the technology. The field of steganography, the technology of hiding messages in content such that the casual observer is unaware of the message's existence, includes imperceptible watermarking.

Readable vs. Detectable Watermarks

A watermark may carry only a single bit of information, that is, it is significant only in its presence or absence. Such watermarks are classified as detectable. A readable watermark, on the other hand, contains a more complex message, typically many bits of information. Mathematically, a readable watermark with N bits of information could be conceptualized as having been chosen from a set of 2^N detectable watermarks. For a message of useful length, the number of marks in such a set becomes unmanageable, so a practical readable watermark implementation must include a means of decomposing the watermark to reconstruct the message from independent parts.

Forensic watermarks must carry a message: a readable watermark, or a series of detectable watermarks is required to identify the particular source of the content instance. If detectable watermarks are used, the message is treated as a series of independent parts, each of which is represented by a single detectable watermark.

Bandwidth

Bandwidth refers to the of data carrying capacity of the watermark, in proportion to the amount of content carrying the watermark. For multi-media content, bandwidth is commonly expressed in terms of message bits per second. In interpreting bandwidth metrics, however, it is important to distinguish between the original message and an encoded message. Forensic watermarking implementers may apply multiple layers of error control coding (ECC) to the message, to compensate for the "noise" in the channel. Such coding can expand the message several fold, thereby reducing the effective bandwidth of the watermarking technique by the same factor. It is also common to embed several copies of the message into the content. For a robust implementation, the bandwidth required is many times that which is implied by the message length alone.

Robustness

Sometimes termed "survivability", robustness is the degree to which the watermark can withstand the various transformations the content may undergo before reaching the recovery process. An effective forensic watermark can tolerate operations such as rescaling, resampling of analog signals, recompression, cropping, rotation, resolution changes, deinterlacing, gamma changes, and temporal averaging, all of which may occur in the course of pirating the content. Additionally, a pirate may undertake targeted attacks to directly suppress the watermark by filtering, noise addition, collusion or other video processing techniques.

Although no watermarking technique is unconditionally robust, an effective technique requires the adversary either to apply an unreasonable amount of effort to suppress the watermark, or to unacceptably impair the content in the process. In signal processing terms, robustness tends to increase with the amplitude of the watermark signal. Paradoxically, if the signal intensity level reaches the threshold of perceptibility, its nature and location become apparent to the attacker, thereby facilitating the attack. Consequently, the watermark intensity must be carefully calibrated to achieve the required level of robustness.

As mentioned previously, error control coding is an important contributor to robustness. Alterations to the content may erase or distort significant portions of the watermark signal. Effective recovery must include mechanisms to compensate for missing or erroneous signal elements. The watermark system communicates over an extremely noisy channel, requiring aggressive error control.

FORENSIC WATERMARKING REQUIREMENTS

The primary requirement of a Forensic Watermarking application is the placement of the watermark embedder at a point in the distribution network where the legitimate recipient of the content instance is known. In uni-cast or download distribution models, the content instance can be watermarked as it is transmitted to the consumer. The consumer thus receives a unique copy of the content, individuated by the watermark that identifies that consumer's identity, account, or purchase transaction. Any reproduction of the content instance can thereby be traced to the consumer when the Forensic Watermark is recovered.

In broadcast or multi-cast systems, however, each consumer receives an identical copy of the content. It is thus only possible to individuate the content instance in the consumer's content receiving device. Consequently, these distribution models require authentication of the receiver, and sufficient security in the receiver to ensure that the watermark is correctly embedded.

Figure 3 – Forensic watermarking in a broadcast distribution model

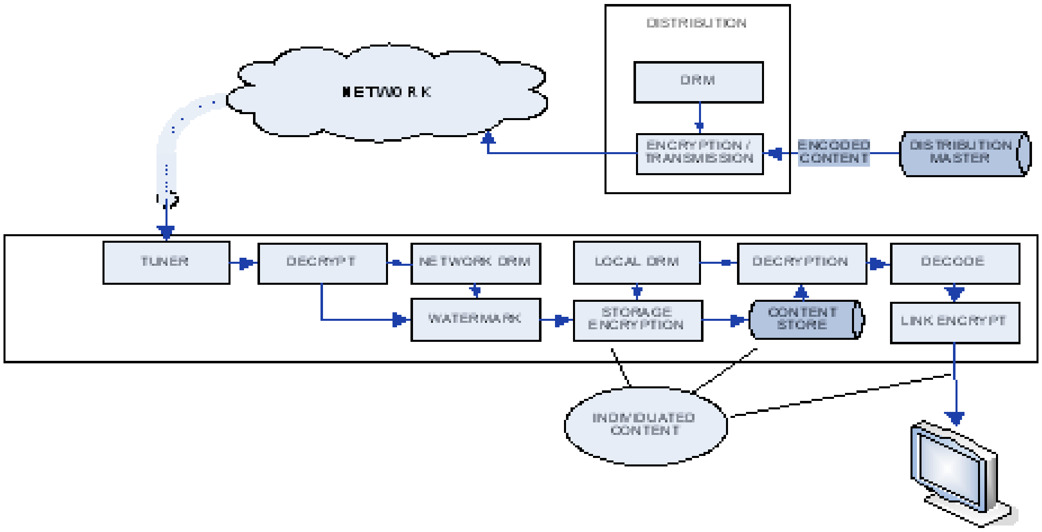

Figure 3 illustrates forensic watermarking in a broadcast distribution model. A watermark is applied following decryption, under the control of the DRM. The content made available to subsequent processes will have been individuated by the forensic watermark.

Watermark Integrity

To prevent compromise, and thus obtain the maximum benefit from forensic watermarking, the watermark should be embedded immediately following decryption. The physical security envelope in the device must enclose all processing inclusive of the DRM and the watermark embedding to prevent interception prior to watermarking or other circumvention. Any accessible data paths carrying unprotected content invite interception and can be targeted by pirates. Once the forensic watermark is embedded, however, the content becomes traceable. Traceable content is less attractive to the pirate, due to the increased risk of exposure, and, consequently, is somewhat less demanding of physical protection. Watermarking early in the content path also ensures that all outputs of the device are protected.

Embedding Performance

A closely related requirement is performance. The STB platform typically lacks substantial spare processing and memory resources. Watermarking in a receiving device must take place at rendering speed for real-time streaming content. Devices supporting background downloads may require watermarking at network speeds exceeding real-time. More sophisticated home network devices may require simultaneous watermarking of multiple content streams. Thus the watermark embedder must only minimally impact the STB computational resources.

Imperceptibility

As noted above forensic watermarks must be imperceptible. The system objective is to preserve the value of the content, so significant quality impacts are unacceptable. Imperceptibility is particularly important for high definition content, which the consumer expects to be of the highest quality.

Robustness

Robustness is, of course, critical in forensic watermarking. The system is only effective in exposing pirates to the extent that the watermark information can be extracted from the unauthorized copy of the content. Pirated content is often degraded in the capture of the initial copy, as well as trans-coded[iv] for the pirate's distribution channel. The initial capture technique may range from a perfect digital copy, to resampling of analog signals, or even a camcorder directed at a rendered image. Unless the pirate captures a compressed digital signal, recompression - possibly accompanied by cropping, frame rate changes, (de)interlace, and/or resolution changes – will be necessary to re-distribute the content. And finally, the pirate may attack the watermark by injecting noise, dropping frames, filtering, or collusion[v]. It should be made difficult for the pirate to verify that s/he has successfully removed the watermark.

Cost Effectiveness

Economical implementation is of paramount importance in the consumer domain. Security features are of minimal apparent benefit to the consumer, so it is generally not possible to recover a significant cost increment for the material and licensing cost of the watermark embedder in each STB.

Renewability

Another requirement is that of renewability. Content pirates have unfailingly adapted to new media security technology. Watermarking will not be spared. As watermarking technology is deployed, adversaries will build tools to suppress the watermark signal. As such tools are perfected and become widely available, the targeted watermark technology will be rendered ineffective. Renewing the watermark system forces the pirate to analyze a new technique and adapt his countermeasures. Thus the ability to renew watermarking techniques, by varying the watermark signal, is a hallmark of an effective system.

Consistency

Uniform quality is important in an entertainment offering. Similarly, uniform robustness is important in a security system. Both reputation for quality and content security are only as strong as the system's weakest links. Similarly, when an unauthorized content instance is discovered, recovery requires knowledge of the technology used to embed the watermark. If the content has been marked inconsistently, it becomes more difficult to effect recovery, and, if no watermark is detected, very difficult to determine which watermarking technology has failed. An ideal watermark system deployment should, therefore, include a mechanism to ensure that all instances of a given content title are watermarked in a consistent manner.

Flexibility

Analogous to renewability, flexibility describes the ease of adapting watermarking to the needs presented by specific content items. The content universe features broad ranges of exposures to piracy, as well as sensitivity to quality. Content providers are likely to prefer watermark perceptibility-robustness tradeoffs that differ from one content item to another. Ideally, a watermark system should provide a means of control, to conform to the content provider's preferences and policies.

For example, Theatrical content may require very low watermark perceptibility with a corresponding decrease in robustness. Alternatively for the purposes of identifying service theft, a higher degree of watermark perceptibility may be tolerated in order to achieve an increase in robustness.

Bandwidth

Forensic watermarking makes only modest bandwidth demands: DCI requires only 35 message bits in each 5 minute segment of a motion picture (~.117 bit/sec).[vi] As mentioned above, allowance for error control coding increases the raw bandwidth requirement.

FORENSIC WATERMARKING ENGINEERING CONSIDERATIONS

An effective watermarking system for forensic watermarking (or any watermarking application for that matter) must perform three basic functions. First, it must decide where in the content to place the watermarks. Secondly, it must generate the watermark signal used to modify content features such that the signal can be detected and recovered from a copy of the content. Thirdly, it must convey information in the watermark signal. The choice of a method to perform these functions greatly impacts the performance and cost of the system.

Meeting the application requirements discussed in the previous section, some in direct opposition to one another, is a non-trivial undertaking. It is illuminating to examine the major issues individually:

Perceptibility vs. Robustness

Both perceptibility and robustness are directly related to watermark signal strength. As the signal amplitude increases, other factors held constant, the watermarks become both more robust and more perceptible. A desired level of robustness can thus be achieved by increasing the signal strength, at cost of quality. Conversely, decreasing signal strength to the point of watermark imperceptibility can impact robustness.

Informed Embedding

Certain watermarking techniques favorably shift the perceptibility-robustness tradeoff. Watermark placement and composition can be optimized to take advantage of host signal (content) characteristics[vii]. Numerous studies of human perception have determined that sensitivity to a particular sensory input varies according to context (i.e. background). In the watermarking paradigm, the watermark is the sensory input that should be disguised, and the background context is the content itself.

A particular watermark signal will, thusly, be more or less likely to be perceived over various backgrounds. The effectiveness of a given background in disguising a feature is called its “masking” property. Masking is a function of the characteristics of both the background host signal and the disguised feature. It is thus possible to reduce watermark perceptibility by choosing watermark signal characteristics and placements that leverage the masking properties of the host signal.

Exploiting the host signal masking properties accommodates more watermark signal energy at a given degree of perceptibility, thereby improving the perceptibility-robustness tradeoff. Robustness can also be enhanced by choosing watermark characteristics and placements that optimize recoverability. Watermark robustness depends on the watermark and background image characteristics, as well as the recovery technique being used. Recoverability analysis evaluates interference between the background image and the watermark signal. The technique is analogous to “dirty paper coding” where the signal is positioned to sidestep interference. Optimal watermark composition is often a tradeoff between perceptibility and recoverability, as an image area with a high level of masking energy may also interfere with the watermark.

Watermark embedding that conforms to the content background characteristics is called "informed embedding"[viii]. Properly employed, informed embedding significantly and favorably shifts the perceptibility-robustness tradeoff. This advantage comes at significant computational cost, however. Analysis of the masking and recoverability properties of motion video signals requires complex algorithms to be applied to several successive frames. The task is particularly challenging at the data rates required to support high definition content in real time.

Sequencing processors or programmable gate arrays capable of this task can add significant per unit costs. ASICs are an option, but only at high volumes, and are difficult to renew.

Receiver Watermarking Security

For broadcast or multi-cast distribution models, as discussed above, forensic watermarks must be applied at the receiver. An optimal security architecture for forensic watermarking in the receiver applies the watermark immediately following content decryption. Both processes should occur within the device's physical security envelope, so that both encoded and baseband plain-text (deciphered) content is protected from eavesdropping prior to forensic watermarking. A serious complication arises, however, due to the requirement of conventional watermarking techniques for access to the uncompressed (baseband) digital video signal.

The baseband digital video is required for informed embedding analysis, and typically for the composition and embedding of the watermark signal. Consequently, a secure architecture in the receiving device requires that the physical security envelope enclose both the decode and watermark processes, in addition to the DRM and decryption blocks. Both the decoding and masking analysis require complex logic, and thus force a potentially costly expansion of the physical security envelope.

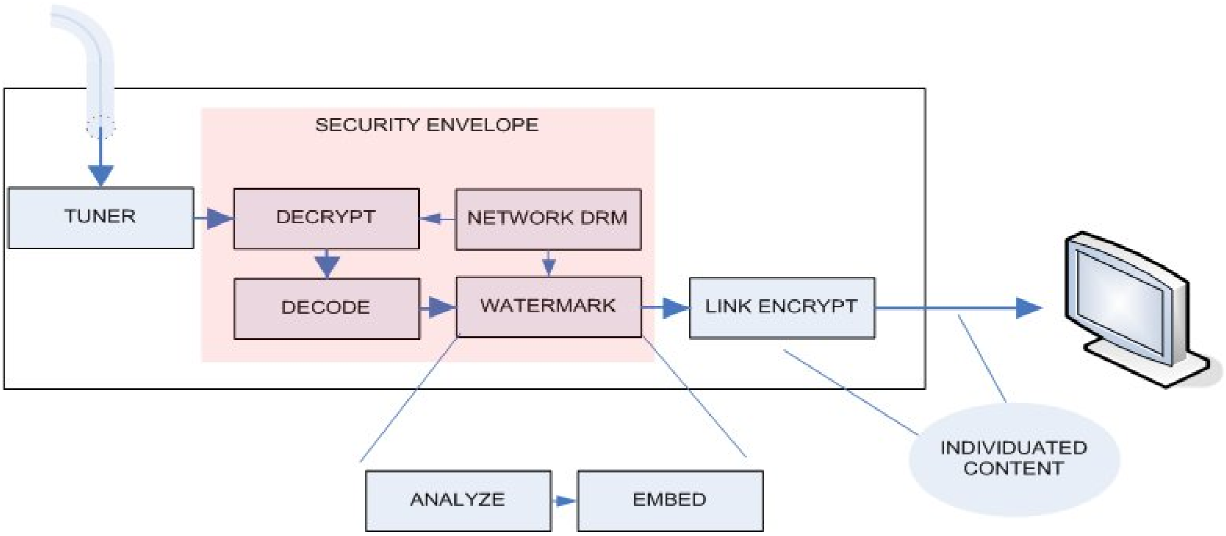

Figure 4 – Forensic watermarking post-decode

Figure 4 diagrams a receiver implementation in which the forensic watermark is applied to the baseband video, subsequent to decode. As shown in the red shaded area, a substantial processing block requires physical security to protect the unmarked data.

Another security issue arises when the receiver imports and stores content, as opposed to rendering in real-time. If the watermark process requires access to baseband content, the receiver must either defer watermarking until the content is decoded and rendered; or decode, watermark, and re-encode prior to storage. The former choice weakens security by distancing watermarking from the initial decryption, in both time and space. Figure 5 illustrates this design, including a very large requirement for the physical security envelope.

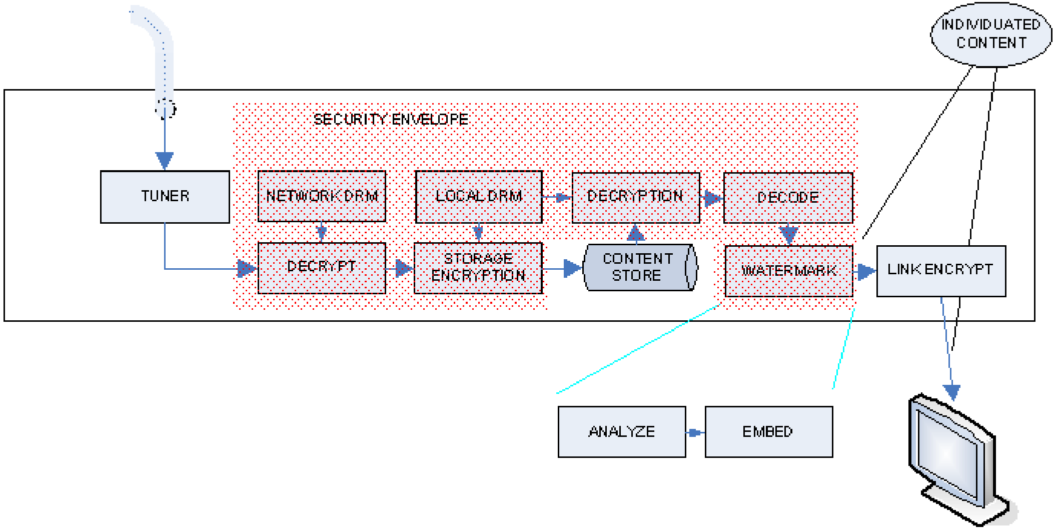

Figure 5 – Late Watermarking Model

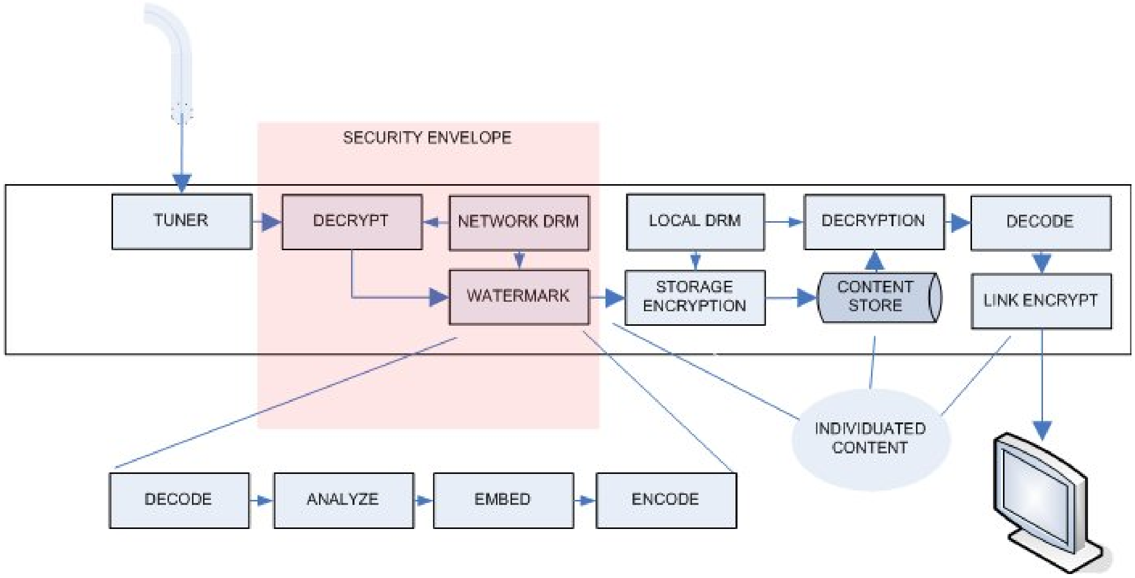

Figure 6 – Watermarking Prior to Storage

The latter choice adds a video encoder to the receiver, at significant component and potential licensing cost, as well as exposure to content degradation due to multiple encodings. This architecture is illustrated in Figure 6. The security envelope appears to have been reduced, but additional decode and encode blocks have appeared as components of the forensic watermarking process.

Integration Issues

A requirement for access to baseband video introduces integration issues for existing equipment designs. It may be difficult to arrange access to the baseband video in a format that is compatible with the watermarking process. Access to several successive frames for sophisticated informed embedding analysis is even more challenging. Additionally, frame latency in the watermarking process could necessitate adjustments to audio synchronization. Large scale integration in media processors can raise formidable barriers to watermark implementation by limiting accessibility to the video frame buffers.

Renewability and Flexibility

As mentioned above, both ease of renewal and flexibility are desirable features in a watermarking system. It is important for the operator to be as nimble as the pirates, so the overhead of a change to the watermarking technique must not constrain his ability to respond to new challenges.

Consistency

The need for consistency in forensic watermarking was discussed above. Consistency is most difficult to achieve when complex algorithms are deployed, particularly in renewal scenarios.

REPLACEMENT WATERMARKING IN THE COMPRESSED DOMAIN

So far, we have discussed forensic watermarking assuming a monolithic implementation – one in which the entire watermarking process takes place in a single device or component, apart from any other device or process. In such architectures, the watermarking process requires access to (at least partially) decoded content. As discussed above, the more sophisticated informed embedding techniques require full access to baseband video. This requirement has caused some system designers to question the feasibility of watermarking in a practical consumer device.

It is clear, on the other hand, that several of the engineering issues could be resolved if the watermark embedder were capable of operating directly on the encoded content. The much lower data rate of encoded content translates to smaller frame buffers and a lesser processing resource requirement. Encoded content is available immediately following decryption, where watermarking could be more securely integrated with the DRM. In use cases where the content is stored or downloaded subsequent to watermarking, the costly decode-encode steps could be avoided. Thus there is ample motivation to develop a technique for watermark embedding in the compressed domain.

Encoded content is, however, highly complex to modify directly. MPEG, the most common video codec family in the consumer domain, contains numerous interdependencies such as inter and intra frame references, differential coding, and entropy coding. Such dependencies make it impossible to interpret a single frame, much less part of a frame, or to make a localized modification. The only obvious approach appears to be the cumbersome decode-watermark-reencode paradigm.

An enhancement to the watermark system architecture can, however, circumvent these problems and actually permit watermark embedding to take place in the compressed domain. Assume that an upstream watermark processing step operates on a single content master instance. This process performs the analysis required for informed embedding, and generates compressed, watermarked video fragments, such that each fragment can be inserted into the encoded content stream, at a specific location, in place of pre-existing video data. Now, package these watermarked fragments with the content, and a downstream watermark embedder need only choose watermarks from this watermark "metadata" to substitute for parts of the encoded content.

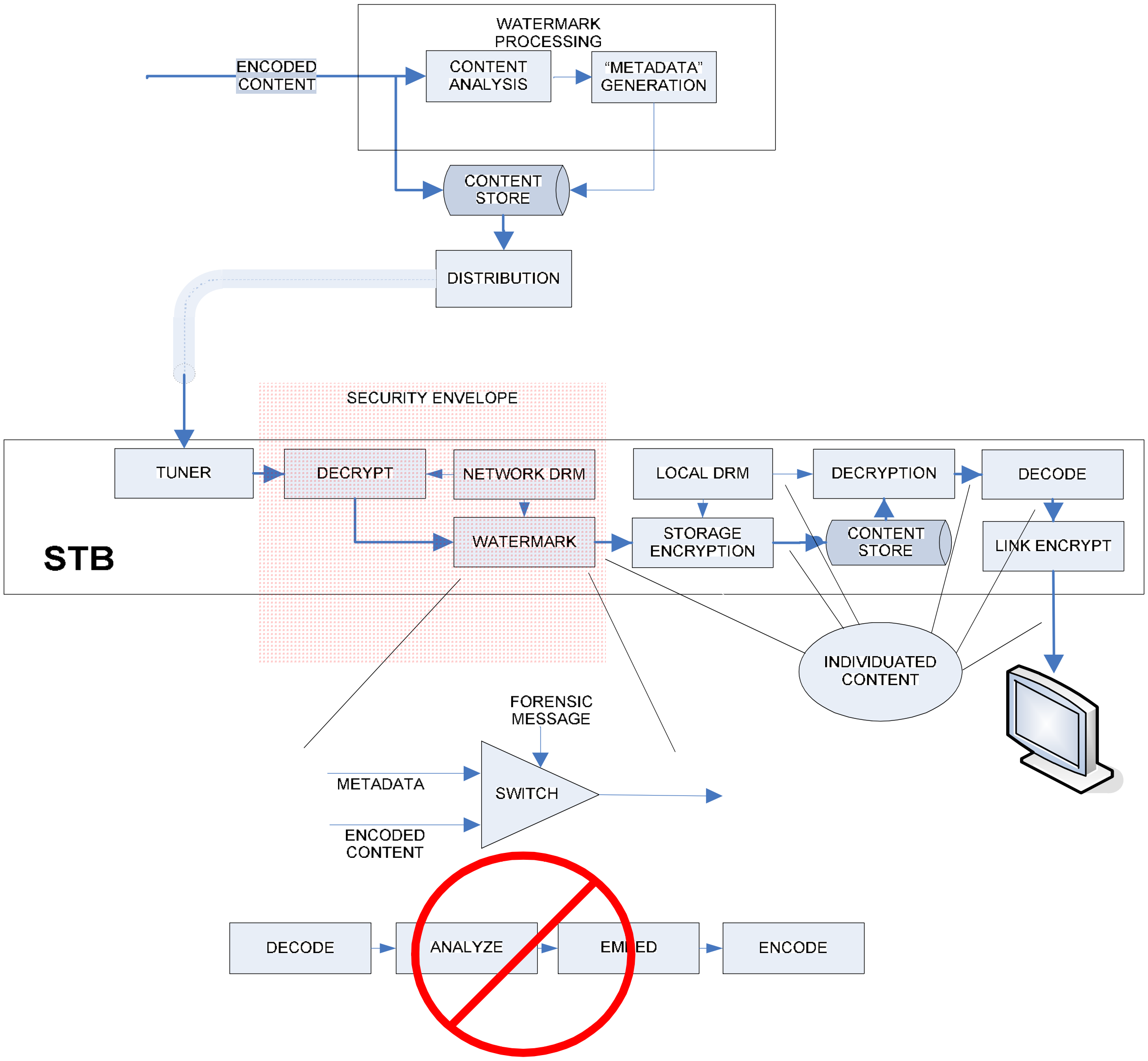

Figure 7 – Replacement Watermarking in the Set Top Box

This architecture, diagrammed in Figure 7, is termed *replacement watermarking*. The embedding process has become a simple switch, capable of selectively replacing segments of the content stream with the watermarked video fragments from the metadata. A substantial efficiency gain is realized by performing complex video analysis and watermark composition only once, where sufficient processing resources can be easily applied, and preserving the results for reuse. Most importantly, decode and encode processing in the watermarking block is completely obviated. Fielded implementations of the replacement embedder require little or no additional processing power than what is commonly available in existing STB designs. The perceptibility-robustness-cost dilemma is thereby alleviated, and the possibility exists that deployed STBs can be upgraded with the light weight embedder.

The STB receiver can utilize replacement watermarking to watermark encoded content immediately following decryption. The very light-weight embedder can be easily integrated with the DRM and decryption processes, and reside within the same physical security perimeter. Whether the resulting content is decoded and rendered immediately, or is stored for later viewing, it will have been individuated by the forensic watermark. The marked content can subsequently be moved throughout a home network, with no requirement for a watermark embedder in each playback device.

Since the embedder is essentially a simple switch, it operates on streams of content data. Frame buffers are not required. Very modest memory demands facilitate integration within the security perimeter inside a consumer electronics device. Integration of the replacement embedder into a STB is straightforward, with reduced impact on time-to-market.

The replacement architecture greatly facilitates renewal. Since all of the watermark placement and composition decisions are made in the upstream watermark processor, these attributes of the watermark system can be altered without affecting the operation of the downstream watermark embedders. Such changes are reflected in the watermark metadata created by the upstream watermark processor, and become effective immediately when the content is processed for replacement watermarking. Significantly, there is no need to update software in widely deployed watermark embedders.

The same techniques used to achieve renewal can be used to increase system flexibility. Watermarking placement and composition are controllable upstream, at the point of watermark metadata creation.

The replacement architecture ensures consistency across the fielded devices, alleviating concerns about heterogeneous fielded watermarking technologies and versions applying watermarks of differing perceptibility and robustness, and requiring diverse recovery techniques. Effectively, the control of watermarking is centralized, and less subject to variations in fielded devices.

REPLACEMENT WATERMARKING ENGINEERING CONSIDERATIONS

Several factors must be considered in the implementation of replacement watermarking. An obvious issue is how to incorporate the watermark metadata into the content package, such that the metadata is available to the watermark embedder. There are several requirements affecting this mechanism.

Foremost, the bandwidth must be available to deliver the metadata to the embedder, properly synchronized with the content. Prior to each frame being processed, any of the watermarked fragments in the metadata affecting that frame must be available to the embedder. Watermarked metadata volume in existing implementations is minimal, but is subject to a number of factors, principally the density of the marks (e.g. marks per second) and the size of the marks.

To secure the watermark embedding process, particularly when it takes place in the potentially hostile environment of the consumer's receiver, it is necessary to ensure that the watermark metadata is bound to the content such that it cannot be identified and removed, prior to decryption. If an adversary were able to isolate the metadata stream, it would be a simple matter to delete or corrupt the metadata and thus suppress forensic watermarking.

Two techniques for conveying the metadata stream have been explored for commercial implementation. One is to simply utilize the codec "user data" features to carry the metadata within the compressed frame structure to which it refers. This approach has the advantage of transparency in distribution, since the metadata simply appears to be part of the encoded content, secured by the same encryption wrapper.

A second approach is to utilize a side channel. When a side channel is employed, it is necessary to secure the side channel to prevent tampering with the data that might disrupt the watermarking process.

On the upstream process side, the metadata must be created in such a way that a valid encoded content stream results when the watermarked fragments are embedded. The techniques for doing this are largely dependent on the codec in use.

SUMMARY

Forensic watermarking involves mass production of individuated content instances. The process is very challenging to securely implement in a large distribution system, using autonomous watermark embedders. By centralizing computationally intensive tasks, distributed watermark embedding can be accomplished through simple operations in the compressed domain. The perceptibility and robustness advantages of informed embedding can be realized with minimal cost impact at scale, along with improved security, renewability, consistency, and flexibility.

[i] Fred Dawson: Studios Eye 1st-Run Service As IPTV Security Advances ScreenPlays January 2008

[ii] Mark Kirstein: MultiMedia Intelligence Identifies Digital Watermarking & Fingerprinting As Key New Opportunity, http://multimediaintelligence.com/index.php?option=com_content&task=view&id=49&Itemid=1

[iii] Steganography is defined as the technology of hiding a secret message inside of a larger cover work, such that the existence and content of the secret message are hidden.

[iv] Media pirates often reencode content at lower data rates or resolutions to suit their preferred distribution channels.

[v] Collusion attacks are those that involve combining content from separately captured instances of the same content in order to dilute the watermark signals.

[vi] Digital Cinema Systems Specification V1.0, July 20, 2005

[vii] Barni, Bartolini, and DeRosa Perceptual Data Hiding in Still Images, Idea Group Publishing 2005

[viii] Barni, Bartolini Watermarking Systems Engineering, Marcel Dekker, Inc. 2004