4. Hypothesis:

4.1 Video Compression:

Most video compression is lossy — it operates on the premise that much of the data present before compression is not necessary for achieving good perceptual quality. For example, DVDs use a video coding standard called MPEG-2 that can compress around two hours of video data by 15 to 30 times, while still producing a picture quality that is generally considered high-quality for standard-definition video. Video compression is a tradeoff between disk space, video quality, and the cost of hardware required to decompress the video in a reasonable time. However, if the video is over compressed in a lossy manner, visible (and sometimes distracting) artifacts can appear.

Video compression typically operates on square-shaped groups of neighboring pixels, often called macro blocks. These pixel groups or blocks of pixels are compared from one frame to the next and the video compression codec (encode/decode scheme) sends only the differences within those blocks. This works extremely well if the video has no motion. A still frame of text, for example, can be repeated with very little transmitted data. In areas of video with more motion, more pixels change from one frame to the next. When more pixels change, the video compression scheme must send more data to keep up with the larger number of pixels that are changing. If the video content includes an explosion, flames, a flock of thousands of birds, or any other image with a great deal of high-frequency detail, the quality will decrease, or the variable bit rate must be increased to render this added information with the same level of detail.

The programming provider has control over the amount of video compression applied to their video programming before it is sent to their distribution system. DVDs, Bluee-ray discs, and HD DVDs have video compression applied during their mastering process, though Blue-ray and HD DVD have enough disc capacity that most compression applied in these formats is light, when compared to such examples as most video streamed on the internet, or taken on a cell phone. Software used for storing video on hard drives or various optical disc formats will often have a lower image quality, although not in all cases. High-bit rate video codecs with little or no compression exist for video post-production work, but create very large files and are thus almost never used for the distribution of finished videos. Once excessive lossy video compression compromises image quality, it is impossible to restore the image to its original quality.

4.2 Theory of video compression:

Video is basically a three-dimensional array of color pixels. Two dimensions serve as spatial (horizontal and vertical) directions of the moving pictures, and one dimension represents the time domain. A data frame is a set of all pixels that correspond to a single time moment. Basically, a frame is the same as a still picture.

Video data contains spatial and temporal redundancy. Similarities can thus be encoded by merely registering differences within a frame (spatial), and/or between frames (temporal). Spatial encoding is performed by taking advantage of the fact that the human eye is unable to distinguish small differences in color as easily as it can perceive changes in brightness, so that very similar areas of color can be "averaged out" in a similar way to jpeg images (JPEG image compression FAQ, part 1/2). With temporal compression only the changes from one frame to the next are encoded as often a large number of the pixels will be the same on a series of frames.

4.2.1 Lossless compression:

Some forms of data compression are lossless. This means that when the data is decompressed, the result is a bit-for-bit perfect match with the original. While lossless compression of video is possible, it is rarely used, as lossy compression results in far higher compression ratios at an acceptable level of quality.

4.2.2 Lossy compression:

The efficient digital representation of image and video signals has been the subject of considerable research over the past 30 years. Digital video-coding technology has developed into a mature field and products have been developed that are targeted for a wide range of emerging applications, such as video on demand, digital TV/ HDTV broadcasting, and multimedia image/video database services. With the increased commercial interest in video communications, the need for international image- and video compression standards arose.

To meet this need, the Moving Picture Experts Group (MPEG) was formed to develop coding standards. MPEG-1, MPEG-2 and MPEG-4 video coding standards have attracted much attention world-wide, with an increasing number of very large scale integration (VLSI) and software implementations of these standards becoming commercially available. Here MPEG-1 was most basic standard which has been developed, which was not compatible with some services. So MPEG-2 has come later to remove the limitations of MPEG-1 by covering all the services. Later MPEG-4 has been introduced which is totally object based compression standard.

Generally speaking, video sequences contain a significant amount of statistical and subjective redundancy within and between frames. The ultimate goal of video source coding is the bit-rate reduction for storage and transmission by exploring both statistical and subjective redundancies and to encode a "minimum set" of information using entropy coding techniques. This usually results in a compression of the coded video data compared to the original source data. The performance of video-compression techniques depends on the amount of redundancy contained in the image data as well as on the actual compression techniques used for coding. With practical coding schemes a trade-off between coding performance (high compression with sufficient quality) and implementation complexity is targeted. For the development of the MPEG compression algorithms the consideration of the capabilities of "state of the art" (VLSI) technology foreseen for the lifecycle of the standards was most important.

Dependent on the applications requirements we may envisage "lossless" and "lossy" coding of the video data. The aim of "loss less" coding is to reduce image or video data for storage and transmission while retaining the quality of the original images the decoded image quality is required to be identical to the image quality prior to encoding. In contrast, the aim of "lossy" coding techniques and this is relevant to the applications envisioned by MPEG-l, MPEG-2, and MPEG-4 video standards is to meet a given target bit rate for storage and transmission. Important applications comprise transmission of video over communications channels with constrained or low bandwidth and the efficient storage of video. In these applications high video compression is achieved by degrading the video quality-the decoded image "objective" quality is reduced compared to the quality of the original images prior to encoding. The smaller the target bit rate of the channel the higher the necessary compression of the video data and usually the more coding artifacts become visible. The ultimate aim of lossy coding techniques is to optimize image quality for a given target bit rate subject to "objective" or "subjective" optimization criteria. It should be noted that the degree of image degradation (both the objective degradation as well as the amount of visible artifacts) depends on the complexity of the image or video scene as much as on the sophistication of the compression technique for simple textures in images and low video activity a good image reconstruction with no visible artifacts may be achieved even with simple compression techniques.

One of the most powerful techniques for compressing video is inter frame compression. Inter frame compression uses one or more earlier or later frames in a sequence to compress the current frame, while intra frame compression uses only the current frame, which is effectively image compression.

The most commonly used method works by comparing each frame in the video with the previous one. If the frame contains areas where nothing has moved, the system simply issues a short command that copies that part of the previous frame, bit-for-bit, into the next one. If sections of the frame move in a simple manner, the compressor emits a (slightly longer) command that tells the de-compressor to shift, rotate, lighten, or darken the copy — a longer command, but still much shorter than intra frame compression. Inter frame compression works well for programs that will simply be played back by the viewer, but can cause problems if the video sequence needs to be edited.

Since inter frame compression copies data from one frame to another, if the original frame is simply cut out (or lost in transmission), the following frames cannot be reconstructed properly. Some video formats, such as DV, compress each frame independently using intra frame compression. Making 'cuts' in intra frame-compressed video is almost as easy as editing uncompressed video — one finds the beginning and ending of each frame, and simply copies bit-for-bit each frame that one wants to keep, and discards the frames one doesn't want. Another difference between intra frame and inter frame compression is that with intra frame systems, each frame uses a similar amount of data. In most inter frame systems, certain frames (such as "I frames" in MPEG-2) aren't allowed to copy data from other frames, and so require much more data than other frames nearby.

It is possible to build a computer-based video editor that spots problems caused when I-frames are edited out while other frames need them. This has allowed newer formats like HDV to be used for editing. However, this process demands a lot more computing power than editing intra frame compressed video with the same picture quality.

4.2.3 Current forms of video compression:

Today, nearly all commonly used video compression methods (e.g., those in standards approved by the ITU-T or ISO) apply a discrete cosine transform (DCT) for spatial redundancy reduction. Other methods, such as fractal compression, matching pursuit and the use of a discrete wavelet transform (DWT) have been the subject of some research, but are typically not used in practical products (except for the use of wavelet coding as still-image coders without motion compensation). Interest in fractal compression seems to be waning, due to recent theoretical analysis showing a comparative lack of effectiveness to such methods.

4.3 Video Watermarking:

Digital watermarking includes a number of techniques that are used to imperceptibly convey information by embedding it into the cover data. There has always been a problem in establishing the

identity of the owner of an object. In case of a dispute, identity was established by either printing the name or logo on the objects. But in the modern era where things have been patented or the rights are reserved (copyrighted), more modern techniques to establish the identity and leave it un-tampered have come into picture.

Unlike printed watermarks, digital watermarking is a technique where bits of information are embedded in such a way that they are completely invisible. The problem with the traditional way of printing logos or names is that they may be easily tampered or duplicated. In digital watermarking, the actual bits are scattered in the image in such a way that they cannot be identified and show resilience against attempts to remove the hidden data.

Although steganography and watermarking both describe techniques used for covert communication, steganography typically relates only to covert point to point communication between two parties. Steganographic methods are not robust against attacks or modification of data that might occur during transmission, storage or format conversion. Watermarking, as opposed to steganography, has an additional requirement of robustness against possible attacks.

An ideal Steganographic system would embed a large amount of information perfectly securely, with no visible degradation to the cover object. An ideal watermarking system, however, would embed an amount of information that could not be removed or altered without making the cover object entirely unusable. As a side effect of these different requirements, a watermarking system will often trade capacity and perhaps even some security for additional robustness.

The working principle of the watermarking techniques is similar to the steganography methods. A watermarking system is made up of a watermark embedding system and a watermark recovery system. The system also has a key which could be either a public or a secret key. The key is used to enforce security, which is prevention of unauthorized parties from manipulating or recovering the watermark. The embedding and recovery processes of watermarking are shown in Figures.
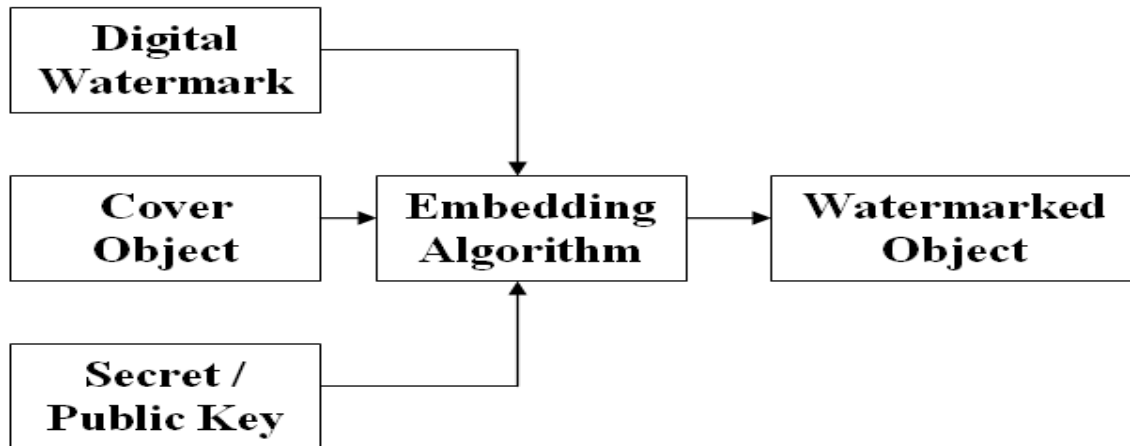
Figure 4.1: Embedding Process – Digital Watermarking

For the embedding process the inputs are the watermark, cover object and the secret or the public key. The watermark used can be text, numbers or an image. The resulting final data received is the watermarked data.
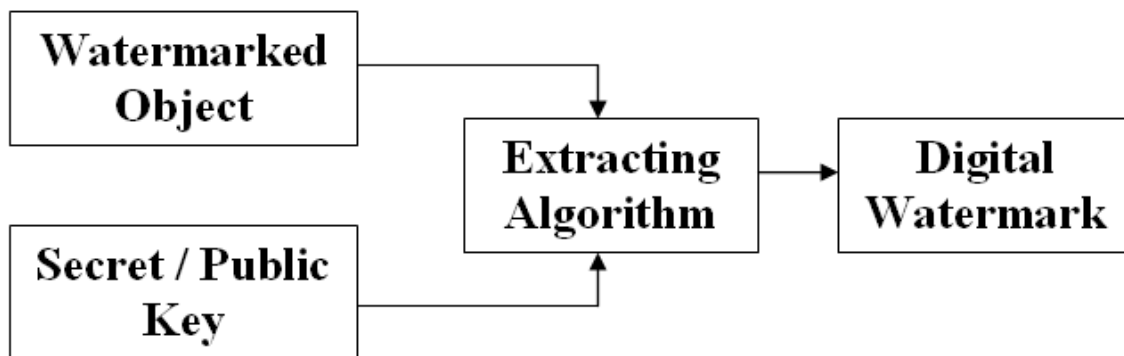

Figure 4.2: Extraction Process – Digital Watermarking

For the extraction process the inputs are the watermarked object and the secret or the public key. The resulting output is the recovered digital watermark.
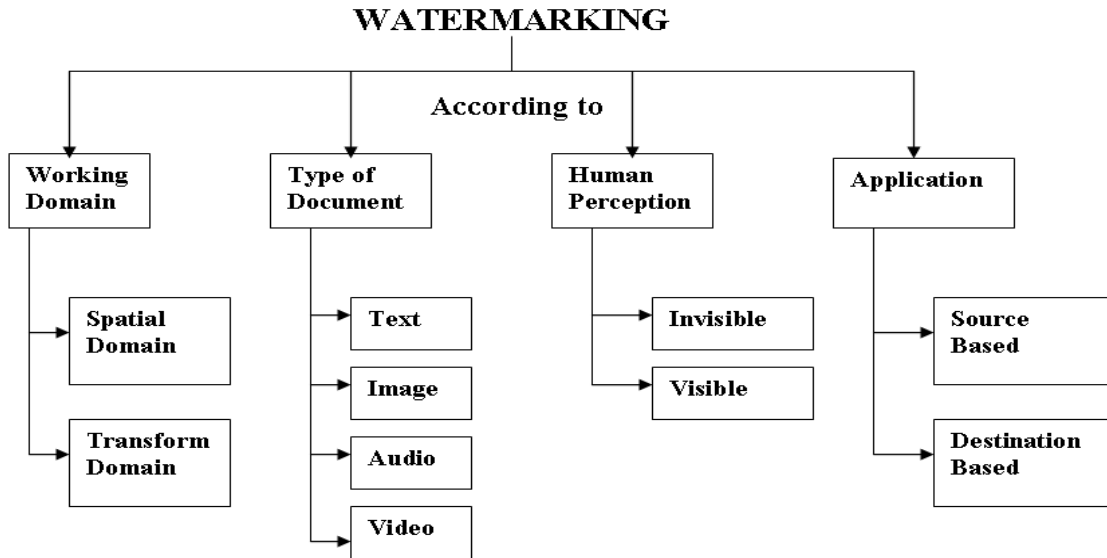
WATERMARKING

According to

Working Domain	Type of Document	Human Perception	Application

Spatial Domain	Text	Invisible	Source Based
Transform Domain	Image	Visible	Destination Based
	Audio		
	Video		

Figure 4.3: Types of Watermarking

Spatial Domain Watermarking: In this method the pixel information of the two-dimensional image is altered so as to embed the hidden data. Three Different techniques are defined in the spatial domain watermarking

Transform Domain Watermarking: Transform domain watermarking techniques apply some invertible transforms to the host image before embedding the watermark. Then, the transform domain coefficients are modified to embed the watermark and finally the inverse transform is applied to obtain the marked image. The transforms commonly used for watermarking purposes are:

Invisible Watermarking: In this technique the watermark is embedded in the cover object in such a way that it cannot be perceptually visible.

Visible Watermarking: In this technique the watermark is superimposed on the cover object in such a way that it can be perceptually visible.

Source Based Watermarking: Source-based watermark are desirable for ownership identification or authentication where a unique watermark identifying the owner is introduced to all the copies of a particular image being distributed. A source-based watermark could be used for authentication and to determine whether a received image or other electronic data has been tampered with.

Destination Based Watermarking: In Destination based each distributed copy gets a unique watermark identifying the particular buyer. The destination based watermark could be used to trace the buyer in the case of illegal reselling.

For the testing of the results the parameters which I will use as a part of visual quality matrix are as listed below. Which basically includes the parameters like SNR, PSNR, Correlation Quality, Mean Square Error, Normalized cross correlation etc.

4.4 Visual Quality Matrices – Pixel Based Matrices:

Signal to Noise Ratio (SNR)	$$SNR = \frac{\sum_{x=1}^{M} \sum_{y=1}^{N} f'^2(x,y)}{\sum_{x=1}^{M} \sum_{y=1}^{N} \{f(x,y) - f'(x,y)\}^2}$$
Peak Signal to Noise Ratio (PSNR)	$$PSNR = 10 \times \log \frac{255^2}{MSE}$$
Mean Square Error (MSE)	$$MSE = \frac{1}{M \times N} \sum_{x=1}^{M} \sum_{y=1}^{N} \left\{ \left(f(x,y) - f'(x,y)\right)^2 \right\}$$
Normalized Mean Square Error (MSE)	$$NAE = \frac{\sum_{x=1}^{M} \sum_{y=1}^{N} \left(f(x,y) - f'(x,y)\right)^2}{\sum_{x=1}^{M} \sum_{y=1}^{N} \left(f(x,y)\right)^2}$$
Normalized Absolute Error(NAE)	$$NAE = \frac{\sum_{x=1}^{M} \sum_{y=1}^{N} \|f(x,y) - f'(x,y)\|}{\sum_{x=1}^{M} \sum_{y=1}^{N} \|f(x,y)\|}$$
Structural Content (SC)	$$SC = \frac{\sum_{x=1}^{M} \sum_{y=1}^{N} f^2(x,y)}{\sum_{x=1}^{M} \sum_{y=1}^{N} f'^2(x,y)}$$

Average Difference (AD)	$$AD = \frac{\sum_{x=1}^{M}\sum_{y=1}^{N}(f(x,y) - f'(x,y))}{M \times N}$$
Maximum Difference (MD)	$$MD = Max\{|(f(x,y) - f'^{r}(x,y)|\}$$
Normalized Cross Correlation (NCC)	$$NCC = \frac{\sum_{x=1}^{M}\sum_{y=1}^{N}f(x,y) \times f'(x,y)}{\sum_{x=1}^{M}\sum_{y=1}^{N}f^{2}(x,y)}$$