Chapter 3

Related Works

Several previous watermarking works related to our proposed algorithm are reviewed in this chapter. The transform domain watermarking schemes are described in Section 3.1. Next, section 3.2 gives the imperceptible watermarking techniques. The watermarking papers related to the important designing issues include robustness, capacity and detection reliability are then discussed in Section 3.3. Finally, Section 3.4 reviews the geometric distortion robust image watermarking schemes.

3.1 Transform-Domain Watermarks

For different application purposes and design requirements, we select a proper domain for watermark embedding. Usually, the transform domains are preferred to the spatial domain for several reasons as given in Section 2.1. In this section, we focus on the watermarking schemes in three popular transform domains including DCT, DFT and DWT. The DCT-domain watermarking is first described in subsection 3.1.1. Subsection 3.1.2 states the DFT-domain watermarking. Several watermarking techniques applied to DWT coefficients are described in subsection 3.1.3.

3.1.1 Discrete Cosine Transform (DCT)

For the watermarking systems, the DCT domain is widely used mainly due to the great

amount of image and video compression applications (e.g., JPEG, MPEG-2, rectangular border blocks in MPEG-4, H.263) and the computation of perceptual distortion thresholds with ease. For example, a video watermarking scheme developed in [13] employs the spread-spectrum technique where the narrow-band watermark signal is transmitted via the wide-band channel with interference (video signal). The need for full decoding of the MPEG bitstreams is avoided since the watermark is inserted into the DCT coefficients. The drift compensation problem is also solved after video watermark embedding. Another DCT based scheme also avoids the full decoding by selectively discarding high frequency DCT coefficients in the JPEG/MPEG compressed domain [14]. The watermark bits are represented by the pattern of the energy differences between DCT blocks.

Typically, the middle-frequency coefficients are suggested for compression robustness watermarks [15]. A DCT coefficient successfully resists against the JPEG compressions if it is larger than half of the quantization step size before and after embedding [16]. On the other hand, if a DCT coefficient is modified to an integral multiple of a certain step size, which is larger than all the allowable quantization steps used in the JPEG compression, then this watermarked coefficient can be correctly recovered after JPEG compression [17].

In addition to the robust coefficient selection and the proper watermark embedding algorithm, a well-designed detector structure can also improve the detection performance. One example is designing an optimal watermark decoder by statistically modeling the DCT coefficients of the original images with the generalized Gaussian distributions for correlation based detection [18].

To realize the blind detection in a DCT-domain watermarking scheme, the channel interference resulted from the image can be eliminated by exploring the high correlation property of the four subimages on the embedder side. These subimages are generated by downsampling a given image at a factor of 2 both horizontally and vertically. A watermark sample is inserted into two DCT coefficients at the same location of two subimages [19].

Similarly, the blind detection is also achieved by exploiting the interblock (a center block and its nine neighboring ones) correlation for watermarking embedding [20].

3.1.2 Discrete Fourier Transform (DFT)

Watermark embedded in the DFT domain solves the synchronization problem after rotation, scaling and translation attacks. The translation has no impact on the magnitude of the Fourier transform of an image, and the scaling and rotation of an image can be traced. Shifting in the spatial domain converts to a linear shift in the phase component, while the magnitude component is invariant. The scaling in the spatial domain produces a linear shift in the magnitude component, and the rotation without cropping leads to the shifting in both of the phase and magnitude components. Furthermore, the cropping effect blurs the spectrum. Thus, there is no synchronization problem with the cropped images. However, the DFT-domain watermarks may be sensitive to other types of geometric transformation such as local warping. One example of the DFT-domain watermarking is proposed in [21], where the symmetric ring covers the middle-frequency DFT coefficients are separated into sectors. The coefficients located in the same sector are embedded with the same watermark bit for rotation invariance.

Translation, scaling and rotation of an image can be also converted to translation behavior in the log-polar coordinates in Fourier transform domain. In this case, scaling in the spatial domain is translated into the translation along the log-radius axis, and rotation in the spatial domain is translated into the cyclic shift along the angle axis. One example is shown in [22]. Another one called Fourier-Mellin transform is the log-polar mapping followed by the Fourier transform. There is only the scaling effect of the coefficient magnitude in the Fourier-Mellin domain, nevertheless, this domain is invariant to rotation and translation [23].

The watermark search space for synchronization during detection can be much reduced in above cases. However, it is known that there is implementation difficulty in the log-polar

mapping of the Fourier transform. If the image undergoes the log-polar mapping, there will be an accuracy problem associated with log-polar mapping (LPM) of DFT since the inverse log-polar mapping (ILPM) transformation requires image interpolation process. Even when the watermark designed in this domain and undergoes the ILPM, it still results in the watermark distortion. One solution to the above problem is to determine the watermark "locations" in the LPM magnitude spectrum, while the watermark is inserted in the DFT magnitude spectrum after mapping the watermark locations in the LPM domain back to the DFT domain as shown in [24].

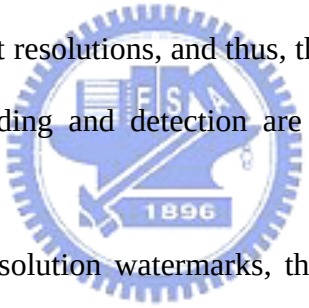
An extra *template* can be embedded in the DFT domain to assist watermark synchronization during the detection process for either image watermarking [25][26] or video watermarking [25]. Essentially, a template is a structured watermark pattern containing no meaningful information. It is secretly defined, and known to both the watermark detector and embedder. It should be invisible and have low interference with the previously embedded watermarks. With the aid of a template, the watermark detector may identify what transformations (rotation, scaling or translation) the image has undergone without referencing to the original image. And thus, the watermark detector can find the original watermark location for watermark extraction. However, a fixed structured template may be identified and destroyed easily. Moreover, it may degrade the overall perceptual quality.

3.1.3 Discrete Wavelet Transform (DWT)

Compared with other transforms such as DCT and DFT, the wavelet transforms has the good property of achieving both the spatial and frequency localization at the same time. It prevents image quality degradation due to the blocking artifacts as those resulted from block-based transforms. Recently, several image (e.g., JPEG2000) and video compression techniques (e.g., Interframe wavelet / 3D subband video coding) also use this technique. Typically, the

DWT-domain watermarking schemes are expected to take the advantage of the multiresolution property of wavelet transforms.

Several DWT based watermarking schemes have been developed. Typically, the lower resolution watermarks and higher resolution ones are embedded into the lower frequency components and higher frequency components of the image, respectively [28]. Since human vision is more sensitive to the lower frequency components and the image energy usually concentrates on the lower frequency components, the lower frequency components are expected to be unaltered by attacks. That is, at least the low-resolution watermark can survive after the common signal processing attacks. Another scheme includes multilevel detection. It inserts the watermark into the high-pass bands at each resolution such that the watermarks are nested from low resolution to high resolution [29]. Such hierarchical structure permits the watermark detection at different resolutions, and thus, the computation complexity is reduced. In [30], the multilevel embedding and detection are also performed based on the scaled versions of watermarks.



Rather than using multiresolution watermarks, the wavelet coefficients of the original images can be grouped into the “super trees” for watermark embedding and blind detection [31]. The wavelet coefficients of a super tree are in the same band (e.g., HL) but grow from low-resolution image coefficients to high-resolution image coefficients. Two super trees are used to embed one watermark bit through quantization where only one tree is quantized according to the watermark bit. On the detector side, the wavelet coefficients of the watermarked image are again grouped and re-quantized. The re-quantization error distribution of the quantized super tree is different from that of the unquantized one. This statistical property is then used for watermark decoding.

Unlike the DFT coefficients, typically the DWT coefficients lack the shift invariance property. The Undecimated Discrete Wavelet Transform (UDWT) has the good property of shift invariance but it requires large computations. Another shift invariant Complex Wavelet

Transform (CWT) is used for image watermarking due to its efficient computation and a modest amount of redundancy as shown in [32]. Regarding the rotation invariance, the Mexican Hat wavelet (Marr wavelet) is considered [33][34] since it has a circularly symmetric frequency response. This wavelet is not separable and it is essentially the Laplacian of a Gaussian function.

3.2 Imperceptible Watermarks

One important watermark design issue is the watermark invisibility. By incorporating the human perceptual model into the watermarking system design, the watermark robustness is enhanced and the image fidelity is retained at the same time. Another advantage of this approach is that if the watermarked image spectrum is similar in shape to that of the original image, the attackers cannot easily identify the embedded watermark by using some prior knowledge of the image statistics. Thus, in this section, we first introduce the perceptually imperceptible watermarking techniques in subsection 3.2.1. And next, subsection 3.2.2 describes the statistically imperceptible watermarking techniques.

3.2.1 Perceptual Imperceptibility

Cox et al. have suggested placing the watermark in the perceptually significant frequency components [35]. To enable maximal watermark strength, the perceptual watermarking techniques either exploit the statistically perceptual characteristics (e.g., texture) within a target region (e.g., noise visibility function (NVF) in [36]) or use the perceptual models derived fitted from subjective experiments. There have been many perceptual watermarking techniques developed in the literature. Instead of introducing these schemes individually, we focus on the human visual models which are generally adopted. For more details of the

overviews of the invisible watermarks for the cases of images and videos, readers can refer to [37] and [38].

The imperceptible watermarking design can be achieved with the aid of the masking effects of the human perceptual system. In the case of watermarking, the masking effect means that the watermark visibility is changed due to the existence of the host data. Several psychovisual models (visual threshold models and visual masking models) have been widely used in the image/video applications, such as just noticeable difference (JND) in the spatial domain, contrast sensitivity function (CSF) and pattern masking (contrast masking) in the Fourier transform domain, and frequency masking and texture masking (contrast masking) in the DCT domain. In the following paragraphs, these models are briefly introduced. More details of perceptual masking effects are discussed in [37][39][40].

A simplified perceptual difference threshold called *just noticeable difference* (*JND*) *threshold* is formulated in the spatial domain. A signal is considered imperceptible if its amplitude is smaller then the threshold. This threshold is determined by the luminance of the signal and its surrounding background pattern, and it becomes larger if the surrounding background is brighter.

The *contrast sensitivity function* (CSF) is a function of the spatial frequency and orientation of the target signal. It is the reciprocal of the just visible contrast and is used to measure the contrast sensitivity where the distortions smaller than the threshold in a particular frequency band are unnoticed. This function is derived based on the experiments of sinusoidal wave gratings. The *pattern masking* further considers the relationship between the masking signal and the target signal (masked signal) at different spatial frequencies and orientations. The largest masking threshold occurs if the two signals have the same frequency and orientation.

The *frequency masking* threshold of one AC coefficient depends on its frequency index and the DC coefficient in the same block. Usually, the frequency masking thresholds are

larger for higher frequency components. A more elaborated model named *texture masking* further extends the concept of frequency masking and it is also affected by the image content. It is similar to the texture masking effect in the Fourier transform domain.

For calculating the perceptual distortion thresholds for wavelet based watermarks, the visibility model of the quantization errors of DWT coefficients are proposed [41][42]. In [41], such a threshold is represented as a function of the decomposition level, orientation (e.g., LL, HL, HH, LH), and the display visual resolution (DVR). DVR depends on the viewing distance and the display resolution. A DWT based watermarking scheme further modifying the perceptual model [42] to better match the purpose of watermarking [43]. The perceptual weighting of one wavelet coefficient is computed based on the resolution band, local brightness of the lowpass filtered image and the texture activity in the neighborhoods of a pixel.

3.2.2 Statistical Imperceptibility

The watermark may be subject to the so called *estimation-based attacks* which make use of the statistical techniques to estimate the original host data and watermark from the watermarked data [44]. The estimation can be done based on the stochastic criteria such as maximum likelihood (ML), maximum a posteriori probability (MAP), or minimum mean square error (MMSE). To resist such attacks, the watermark should be designed statistically imperceptible. That is, the watermark is hard to be estimated statistically as stated as follows.

In [45], this attack is analyzed under the assumptions that the original host data and watermark are independent, zero-mean, stationary, and colored Gaussian random processes. It concludes that if the power spectrum of the watermark is proportional to that of the original host data (power spectrum condition), the watermark is difficult to be estimated in the mean-squared error sense (Wiener filtering). Another work uses the image denoising methods

for the estimation-based attack purpose [36]. To have the watermark resist this attack, the watermark should match a noise visibility function (NVF), which is derived from the image statistics (mean and variance) for evaluating the region smoothness.

3.3 Robustness, Capacity and Detection Reliability

Several works studying the watermark capacity issue have been published using the theoretical analysis approach [3][46][47]. Clearly, there are tradeoffs between the achievable watermarking rate, allowable distortion, and robustness against attacks [3]. It has been reported that the transform–domain watermarking techniques can offer a higher capacity under specific attacks (such as compression) [46][48]. The perceptual watermark capacity in different transform domains has been analyzed in [49]. In [47], the capacity constrained by reliable statistical detection is calculated. In [50], the minimum number of coefficients in discrete wavelet domain with spread spectrum watermark embedding is theoretically analyzed using the human visual model and a probabilistic detection model. In [51], two optimization schemes based on genetic algorithms are proposed. In the first scheme, the best watermark embedding positions are searched to enhance the content visual quality under guaranteed robustness. The second one aims at application-specific data capacity, fidelity, and theoretically optimal robustness against certain types of attacks.

The analysis on the visibility of embedded watermarks becomes quite complicated when both attacks and watermarks coexist. In [52], the author suggests that the joint distortion due to watermarking and the attack (compression) on the original host data should be kept lower than the just noticeable difference (JND) of the human perceptual system. Therefore, the watermark capacity is also constrained by the human visual threshold. The work in [53] takes account of such issue and theoretically estimates the maximum number of bits can be embedded watermark for JPEG-to-JPEG image watermarking schemes. In a JPEG-to-JPEG

watermarking scheme, watermarking is applied to the JPEG compressed image, and then, the watermarked image is again JPEG-compressed for JPEG standard compliance. The authors adopt a modified Watson's human visual system (HVS) model for computing the just noticeable difference (JND) of the DCT coefficients.

Another possibility to insert watermark while maintaining acceptable distortions is considering the watermark embedding and the attack effect at the same time. For this, the achievable rate pairs of quantization rate (related to the size of the source codebook) and watermarking rate (related to the number of possible watermarks) based on the joint compression and watermarking techniques are theoretically analyzed as shown in [54]. That is, the original host data and the watermark are jointly encoded as a representation vector in a source codebook. In the analysis, the average per-symbol quadratic distortion between the original host data and the compressed watermarked data has to satisfy the distortion constraint. The additive Gaussian noise attacks and the non-blind detection are assumed.

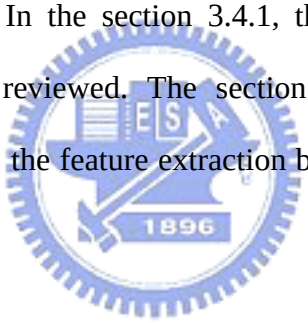
Informed embedding modifies the host data coefficients in accordance with both the host data and the watermark, rather than the watermark alone. To achieve the watermark robustness and high capacity at the same time, one example applies the informed watermark embedding in an iterative sense is proposed in [55]. In each iteration, the watermark embedder modifies the fixed sets of DCT coefficients (12 low-frequency AC components) to prevent the embedded watermark bits are not decoded as invalid ones even after the attack of noise addition.

It is observed that increasing the number of watermarked coefficients does not necessarily benefit the detection performance. To improve the detection performance with limited watermark capacity, one possible solution is embedding multiple watermarks under the consideration of possible types of attacks as shown in [56]. A cocktail watermarking scheme inserts two complementary watermarks by applies both the negative modulation (decreasing the magnitude of transformed coefficients) and positive modulation (increasing

the magnitude of transformed coefficients) such that at least one can survive after different attacks.

3.4 Geometric Distortion Robust Watermarking

These attacks on watermarks can roughly be classified into geometric distortions and noise-like signal processing. Geometric distortions are difficult to tackle. They induce synchronization errors between the extracted watermark and the original watermark during the detection process, even though the watermark still exists on the watermarked image. Nowadays, several approaches that counterattack geometric distortions have been developed. These schemes can be roughly divided into invariant transform domain based, moment-based and feature based algorithms. In the section 3.4.1, the invariant transform domain based watermarking techniques are reviewed. The section 3.4.2 describes the moment based watermarking schemes. Finally, the feature extraction based methods are discussed in section 3.4.3.



3.4.1 Invariant Transform Domain Based Techniques

Watermark embedded in invariant-transform domains generally means that the synchronization problem is solved even with rotation, scaling and translation attacks. The DFT domain has this good property; therefore, it is widely used in the watermarking systems as discussed in 3.1.2. Note that in the case of DFT domain, the symmetric points on the lower half DFT plane have to be also altered to the exact same values in order to produce a real-valued image after the spectrum modification on the upper half DFT plane for watermark embedding.

3.4.2 Moment Based Techniques

The watermark detection process is similar to the pattern recognition process in computer

vision, but the original images may not be available to the watermark detector. Moments of objects have been widely used in pattern recognition. Higher order moments are more sensitive to noise and some normalization schemes have been designed to tolerate noise [57]. A watermarking system employing image normalization with respect to orientation and scaling is proposed in [58]. If the image normalization process is applied to the entire image, it will be sensitive to cropping and local region distortion. Another moment based watermarking scheme [59] hides watermarks by modifying image content iteratively to produce the mean value of several invariant moments in a predefined range. The watermarked image is a linear combination of the original image and a weighted nonlinear transformation of the original. The weight is computed such that the mean of the watermarked image invariants is a predefined number. The watermark detector verifies the presence of the watermark by checking the mean value of these moments. This scheme can resist orthogonal transformations and general affine transformation, but it is sensitive to cropping and aspect ratio changes. In [60], the watermark rotation, scaling and translation invariance is achieved by using the Zernike moments and image normalization techniques. Zernike moments are good for their insensitivity to image noises and information content.

3.4.3 Feature Based Techniques

The extracted features of image content can be used as reference points for both watermark embedding and detection [61]-[63]. In [63], the Harris detector and the Achard - Rouquet detector are used for feature extraction. The extracted feature points are robust to geometrical transformation as rotation, translation, scaling or even little morphing. These points often locate around image corners, edges and deep textures. Several triangles are constructed using the feature points, and the watermark is embedded independently inside each triangle. At detector, the extracted triangles are warped to a reference pattern to correlate with the original

watermark reference triangle pattern. Simulation results show that this scheme is less effective for images with mainly textures.

Image orientation means the main image energy spreads over a particular direction. Once an image is geometrically distorted, the orientation should change accordingly. Therefore, if a watermark pattern is designed and embedded aligned to the image orientation, at watermark detection, the watermark can be extracted by estimating the new image orientation without referencing to the original image. In [64], such a rotation invariant watermarking scheme is proposed through the use of the 2-D real and imaginary Gabor functions for orientation estimation.

Another scheme that resists the scaling and rotation attacks uses Radon transformation [65]. According to the characteristic values extracted by the Radon transformations with proper curve selection, a reference watermark is geometric transformed to conform to these characteristic values before both embedding and detection. Thus, the synchronization problem does not exist during watermark detection.

In [61], the watermarking of facial images is based on the localized salient facial features such as eye and mouth. These features are extracted with the shape template matching, and they are rotation, translation and scaling invariant.

In [62], authors suggest extracting feature points by the Mexican Hat wavelet scale interaction method. These points are connected to form a Voronoi diagrams for watermark embedding, and they experimentally show that it is very robust to high quality JPEG compression [66]. Although these feature points are rotation-invariant, the embedded watermarks in the Voronoi diagrams are not rotation-invariant and thus still have to be searched in the rotated images.

Many image watermarking schemes have been proposed to resist the global affine transforms such as rotation, scaling and translation. But there are few papers dealing with the *random bending distortions*. The freedom of random bending is unlimited. Such attacks may

cause almost unnoticeable perceptual distortion, but it can defeat most of the existing image watermarking schemes. A possible solution is proposed in [67] where the random distorted image is corrected based on a regular mesh model. The mesh node represents some position on an image. To estimate the image distortions, they find the best match of the corresponding mesh node intensity between the original and distorted images.