Image Processing for Surveillance and Security

Deepayan Bhowmik and Mehryar Emambakhsh
Sheffield Hallam University, Sheffield, United Kingdom and
Heriot-Watt University, Edinburgh, United Kingdom

Abstract

Security is a fundamental issue in today's world. In this chapter we discuss various aspects of security in daily life that can be solved using image processing techniques by grouping in three main categories: visual tracking, biometrics and digital media security. Visual tracking refers to computer vision techniques that analyses the scene to extract features representing objects (*e.g.*, pedestrian) and track them to provide input to analyse any anomalous behaviour. Biometrics is the technology of detecting, extracting and analysing human's physical or behavioural features for identification purposes. Digital media security typically includes multimedia signal processing techniques that can protect copyright by embedding information within the media content using watermarking approaches. Individual topics are discussed referring recent literature.

1 Introduction

Surveillance and security are the integral part of today's life particularly when million's of CCTVs are in action covering many public places including, airports, railway stations, high streets, market places, shopping complexes, theatres and many more to name. At the same time, in today's digital world we capture, store and share millions of images, videos and other creative contents. Development in network infrastructure and media compression technology influenced a paradigm shift in entertainment content consumption such as video streaming, and on-demand video services. However all these scenarios pose several challenges in image processing community in addressing digital security in everyday life.

The security and surveillance are often treated separately in the image and signal processing research domain. While computer vision techniques deal with challenges in visual tracking, anomaly detection or biometric identifications, multimedia signal processing community is more concerned with digital media security and digital right management using data hiding techniques (*e.g.*, watermarking). In this book chapter we have attempted to discuss image processing approaches proposed in the literature addressing various digital security related issues, from surveillance to media security. To the best of the authors' knowledge, there is no such attempt made in the literature that brings together such a range of topics under one common theme.

The chapter is broadly categorised in three sections: *a)* visual tracking, *b)* biometrics and *c)* multimedia security. We have restricted our discussion of the individual topics with respect to image processing techniques.

2 Image Processing for Visual Tracking

Automatic detection, tracking and anomaly identification occupy a considerable space in computer vision research and have many applications including intelligent surveillance, human-computer interaction (HCI), human-robot interaction (HRI), augmented reality (AR), medical applications and visual vehicle navigation etc. More importantly, recently there is a great deal of interest in robust visual tracking algorithms due to the increased need of automated video analysis relating safety in public places such as railway station, airport, shopping areas, religious festivals or governmental offices. This poses fundamental challenges in computer vision which involves input from image processing research along with input from machine learning community. In this section we discussed fundamental steps for visual tracking by dissecting the algorithms available in recent literature.

Image analysis for tracking consists of three key steps: 1) detection of interesting moving objects; 2) tracking moving objects in time lapsed frames and 3) analysis of tracked objects for behavioural study such as recognition, prediction and anomaly detection. One can broadly dissect the visual tracking algorithms in four different categories (Yilmaz et al., 2006; Haering et al., 2008; Yang et al., 2011): *a) Object representation*, *b) Feature selection*, *c) Object detection* and *d) Object tracking*. These categories can also be grouped and fitted into a classical image processing pyramid of *low level, medium level, intermediate level* and *high level* algorithms, based on the complexity and the type of data they process (Figure 1).

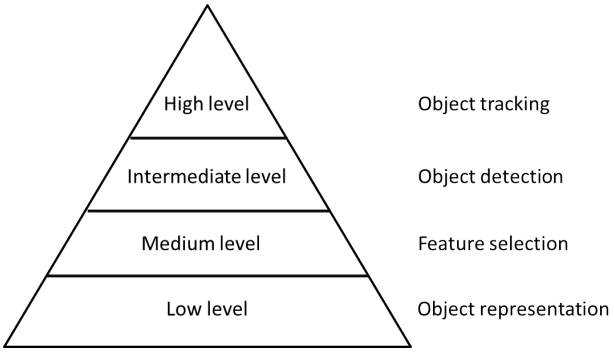


Figure 1: Image processing pyramid of visual tracking algorithms.

2.1 Object Representation

Shape representation

Objects in a scene can be represented by their shape of appearance. Example of shape representations are: ***i) Points*** where objects can be represented by a point, *e.g.*, centroid (Veenman et al., 2001) or by a set of points (Serby et al., 2004) to track objects that engage a smaller region of the image space; ***ii) Primitive geometric shapes*** where objects can be represented by various geometric shapes such as rectangle, ellipse (Comaniciu et al., 2003) or parametric ellipsoid (Limprasert et al., 2013) for tracking purposes; ***iii) Object silhouette and contour*** where the contour of object boundary and the contour region is called as silhouette and used in tracking (Yilmaz et al., 2004). ***iv) Articulated shape models*** where targets are composed of combination of connected body parts, *e.g.*, human subject can be articulated with the torso, head, arm, leg etc... (Husz et al., 2011). Individual body parts are represented using cylinders or ellipses; and ***v) Skeletal models*** where skeletons are extracted from object silhouettes by applying medial axis transform and used in object recognition (Ali and Aggarwal, 2001) or tracking (Schwarz et al., 2012).

Appearance representation

Alternatively the objects can be represented by their appearances combined with the shape model for tracking purposes. Example appearance models are: ***i) Probability densities of object appearance*** features (*e.g.*, color, texture) can estimated for the region of interest from their shape representation. Example probability density estimation algorithms are mixture of Gaussian (Paragios and Deriche, 2002) or histograms (Comaniciu et al., 2003); ***ii) Templates*** are used *e.g.*, for object recognition and tracking, and can be formed from the geometrical shapes or silhouettes (Dufour et al., 2002); and ***iii) Active appearance model*** are generated simultaneously by statistical shape models and appearances, which can generalise the object *e.g.*, face (Edwards et al., 1998). While the shape can be defined by object representation (*e.g.*, contour), the appearance can be defined from colour, texture or gradients.

2.2 Feature Selection

A single feature or combination of features are crucial for visual tracking, in order to uniquely identify a target object in image space or world coordinates. Features can be affected by many factors including viewpoint, occlusion, illumination, texture, or articulation. Commonly used visual features are:

Colour: Various color spaces, *e.g.*, RGB (Red, Green, Blue) and HSV (Hue, Saturation, Value) have been used to represent object features. The object color can be influenced by two physical phenomena (Yilmaz et al., 2006): 1) spectral distribution of the source and 2) surface reflectance of the target object. Comaniciu *et al.* (Comaniciu et al., 2003) used color histograms in mean-shift based object tracking.

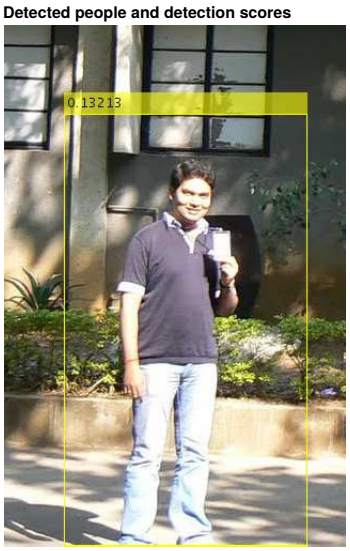


Figure 2: Example of a pedestrian detector using HOG (Dalal and Triggs, 2005).

Edge: Edge features can be generated by capturing sudden intensity changes and are usually less sensitive to illumination changes when compared to using colour feature selection. The Canny edge detector (Canny, 1986) is one of the most used edge detectors to date.

Texture: Texture selection measures the intensity variation of an object's surface to quantify smoothness and regularity (Shotton et al., 2009). Various algorithms have been proposed to investigate texture patterns using filters (Laws, 1980), wavelets (Mallat, 1989) or local binary patterns descriptors (Ojala et al., 2002). Similar to edge feature selection, this is also less sensitive to illumination changes.

Gradient: The directional change in intensity or colour is captured in gradient feature and has been proven to be more robust particularly for human detection *e.g.*, Histogram of Oriented Gradients (HOG) by Dalal and Triggs (Dalal and Triggs, 2005).

Spatio-temporal: Local spatio-temporal features capture the salient and motion characteristics in video and is usually invariant to spatio-temporal shifts, scales or background clutter. For example optical flow (Horn and Schunck, 1981) defines the translation of each pixel in a region and is used as feature in motion-based segmentation and tracking algorithms.

2.3 Object detection

Object detection methods are necessary in order to perform object tracking, and are applied in every frame in the target video. In some cases, temporal cues are used to reduce false detection. Various types of detectors have been proposed in the literature including:

Point detector: Objects are represented by interest points with expressive texture. Commonly used algorithms are invariant to changes in illumination, camera viewpoint, rotation or scaling. Examples of popular point detectors include Harris detector (Harris and Stephens, 1988), KLT (Kanade-Lucas-Tomasi) detector (Shi and Tomasi, 1994), and SIFT (Scale-invariant feature transform) detector (Lowe, 2004).

Background subtraction: These algorithms model the pixels to be either part of the background or part of the foreground. Any considerable change signifies as moving object detection. Example background subtraction algorithms can be found in (KaewTraKulPong and Bowden, 2002) and (Zivkovic and van der Heijden, 2006).

Segmentation: These algorithms partition the image into perceptually (*e.g.*, by colour or texture) similar regions. Segmented regions can be identified as a target object which can later be used in tracking. Various algorithms have been proposed, including mean-shift clustering (Comaniciu and Meer, 2002), segmentation using graph-cuts (Boykov and Funka-Lea, 2006), and active contours (Yilmaz et al., 2004).

Supervised learning: Object detection using sequences of different object views or illuminations can be achieved in supervised learning algorithms as opposed to template matching. Supervised learning is a classification problem. Object detection algorithms are often trained and tested with given object feature set and associated object classes. Adaptive boosting (Adaboost) (Freund and Schapire, 1997) and Support Vector Machines (SVM) (Boser et al., 1992) are two classic examples. An example of a popular pedestrian detector (Dalal and Triggs, 2005) is shown in Figure 2.

2.4 Object tracking

Object tracking algorithms generate the trajectory of the targets over time in a video sequence. Object detection and tracking are often performed either separately *e.g.*, detecting target objects in every frame and then track by correspondence over frames; or jointly *e.g.*, correspondence is estimated by iteratively updating object information (*e.g.*, points of interest or location) using processing outputs from previous frames. Many tracking algorithms can be found in the literature, and can be grouped into the following categories:

Point tracking: Point tracking is performed by correspondence of detected objects represented by points. Based on the correspondence type, this can again be classified as 1) *deterministic e.g.*, Greedy Optimal Assignment (GOA) Tracker (Veenman et al., 2001) or 2) *statistical e.g.*, Kalman filter (Broida and Chellappa, 1986), Particle Filter (PF) or Sequential Monte Carlo (SMC) (Arulampalam et al., 2002) , Joint Probability Data Association Filter (JPDAF) (Bar-Shalom and Foreman, 1988), Multiple Hypothesis Tracking (MHT) (Blackman, 2004) or Probability Hypothesis Density (PHD) Filter (Vo and Ma, 2006).

Figure 3: Example of color histogram based mean-shift tracking.

Silhouette tracking: Tracking algorithms often consider silhouettes to represent complex object shapes *e.g.*, a human body cannot be described by simple geometric shapes, and corresponding target object regions in successive frames. For example Kang *et al.* (Kang et al., 2004) used histogram of colour and edges as object models and used in tracking.

Kernel tracking: Kernel tracking algorithms depend on the motion of an object often defined by a region in subsequent frames. Mean-shift (Comaniciu et al., 2003), Kanade-Lucas-Tomasi (KLT) feature tracker (Shi and Tomasi, 1994), Support Vector Machine (SVM) tracker (Avidan, 2004) are few examples of kernel tracking. An example of mean-shift tracking (Comaniciu et al., 2003) is shown in Figure 3.

3 Biometrics authentication [1]

Biometrics is a combination of two Greek words (Jain et al., 2004): *bios*, meaning life, and *metrikos*, meaning measure. The main motivation of storing a person's biometric data is to reuse it for the recognition purposes. Therefore, there are generally at least two sessions involved at a biometrics data storage process: *a) data acquisition stage*, in which the raw data is obtained and biometric data is stored; and *b) authentication stage* acquired data for the target subject is compared with the available dataset to check the person's identity. The biometric data obtained at the first stage is usually known as the *gallery*, while the one from the second stage is known as the *probe*.

A good biometric system should be (Jain et al., 2004, 2006):

- Consistent: The observed physical or behavioural biometrics features captured from a subject should not significantly change, when the probe data acquisition is performed, *i.e.*, the data should not lose its similarity with the data stored in the gallery for the same class labels, and should maintain its dissimilarity with the samples from other subjects.

[1] This section is derived from the PhD thesis of the chapter's second author (Emambakhsh, 2014)

- Discriminative: The biometric data should not be similar between different people to maintain its reliability in recognising a subject correctly.
- Easily obtainable: The imaging or in general, data acquisition procedure should be easy and practical for both the subject, whose biometric data is obtained and the biometrics system customer, who utilises the data capture device.
- Robust: Another key feature of a good biometrics is its robustness. The biometric features should not be easily manipulated or changed by the subject. The biometrics system should be prepared to detect such change. Otherwise, the subject might be wrongly classified as another person or might not be correctly verified.
- Secure: As the acquired data is completely private for a subject, it should be very safely and securely stored. Also, the biometrics data should be properly coded and encrypted to avoid any spoofing attacks as much as possible.
- Maintainable: Although the storage costs are annually becoming lower and lower, the size of data should not be too large. This is not just to reduce the storage expenses, but to reduce the processing time of the data.
- Fast: The data acquisition and process procedures should be as quick as possible. Slow processing time is inconvenient for both sides of a biometrics system, i.e the subject and device customer.

There are two widely used approaches to evaluate the performance of a biometrics system. The first one is by using the receiver operator characteristic (ROC) curve. In the case of a verification scenario (or open-set recognition (Scheirer et al., 2013), in which the subject might not necessarily have a corresponding sample in the dataset), based on the matching scores computed over the samples in a biometrics test session, a subject can be selected as legitimate or an impostor. The decision making is usually performed by assigning a threshold on the subject's biometrics matching scores (Jain et al., 2006). All biometric systems produce some degree of errors in their output decisions. Therefore, if the decision making threshold is varied from 0 to 1 over a normalised set of matching scores, a subject's biometrics verification rate can be computed. Plotting these thresholds (false alarm rate (FAR) or false positive rate(FPR)) against the output verification rate (true positive rate (TPR)) or sometimes against FNR, produces the ROC curve. For the former, the more concave the ROC curve means a better verification rate. An important point on an ROC curve is when FAR and TPR are equal. This point is called equal error rate (EER) and is usually reported to show the biometrics system's verification performance. Also, detection error trade-off (DET) is a modified version of an ROC curve used to quantitatively describe the performance of a biometrics system.

The second approach is based on the cumulative matching characteristic (CMC) curve. The CMC curve is usually used for the identification scenarios (or in close-set recognition (Scheirer et al., 2013), in which the test subject has a corresponding sample in the gallery). It is the probability of correctly assigning the right class label to a test sample, when the matching scores are sorted per gallery samples (Jain et al., 2006). Therefore, the classification rate at the i^{th} rank means the

probability of assigning the correct label to the test samples, after the i smallest matching scores are evaluated. For the CMC curves, usually the classification rate of the first rank is reported. It corresponds to the probability of the equality of test subject's label to the smallest matching score.

While fingerprints, irises, pupils, faces, palmprints, voice and gait are the most common biometric modalities, they all suffer from certain issues. For example, fingerprints and palmprints are vulnerable to dirt and scratches (Jain et al., 2004), while iris and pupil recognition performance can be deteriorated by blinking, occlusions or contact lenses (Jain et al., 2006). In addition, gait and voice can be inconsistent and vary over time.

As high resolution cameras have become cheaper, the face has become one of the most reliable biometrics. However, make up, lighting conditions, expressions, pose and occlusions are also commonly known issues associated with face recognition procedures. In order to reduce the sensitivity of face recognition algorithms to lighting and pose, 3D face recognition was introduced. Numerous 3D capturing devices and technologies, such as structured-light based (Savran et al., 2008), Minolta Vivid 900 laser (Phillips et al., 2005) and Photometric Stereo imaging (Zafeiriou et al., 2011), have been utilised to perform data acquisition at biometric sessions.

A typical face recognition algorithm is shown in Fig. Figure 4. The preprocessing section of a face recognition algorithm usually consists of pose correction, denoising and face detection. Then, in the feature detection step, some informative sets of points, which can be consistently detected on the facial surface, are used for feature extraction. After that, the feature space is post-processed. Some of the well-known algorithms used for feature space post-processing are feature selection, mapping, dimensionality reduction algorithms. Multiple sets of features can be fused at this step to increase the classification accuracy.

The next step is denoted as feature space creation, in which the feature space can further be manipulated to make it more robust. There are numerous algorithms for this task, such as principal component or discriminant analysis algorithms. Then, the matching or classification methods are applied to the feature space and, finally, the decision making is performed. Depending the capabilities of a face recognition algorithm, some of these steps might be omitted. For instance, if the feature detection or extraction step is rotation invariant, using an alignment algorithm might not be necessary.

3.1 Nose as a biometric

Although capturing the 3D faces enables solving several issues associated with the 2D face recognition, the algorithms are still sensitive to facial expressions, occlusions, pose and noise. As an example, Figure 5 shows the matching errors for the face of subject 1, in 5 different cases: neutral, occluded, rotated, noisy and with facial expression, demonstrating sensitivity of holistic facial matching algorithms to facial deformities. The matching is performed using the iterative closest point (ICP) algorithm (Best and McKay, 1992).

As a solution to this, 3D face recognition can be performed using the nasal region. The most fundamental reason of choosing the nose region for recognising people is its small variations in different expressions. Compared to other parts of the face, like the cheeks, eyes, forehead and mouth, it is much less changed in non-neutral expressions. Moreover, since the nose tip is usually the closest point to the camera and its convexity is more salient than other parts of the face, the

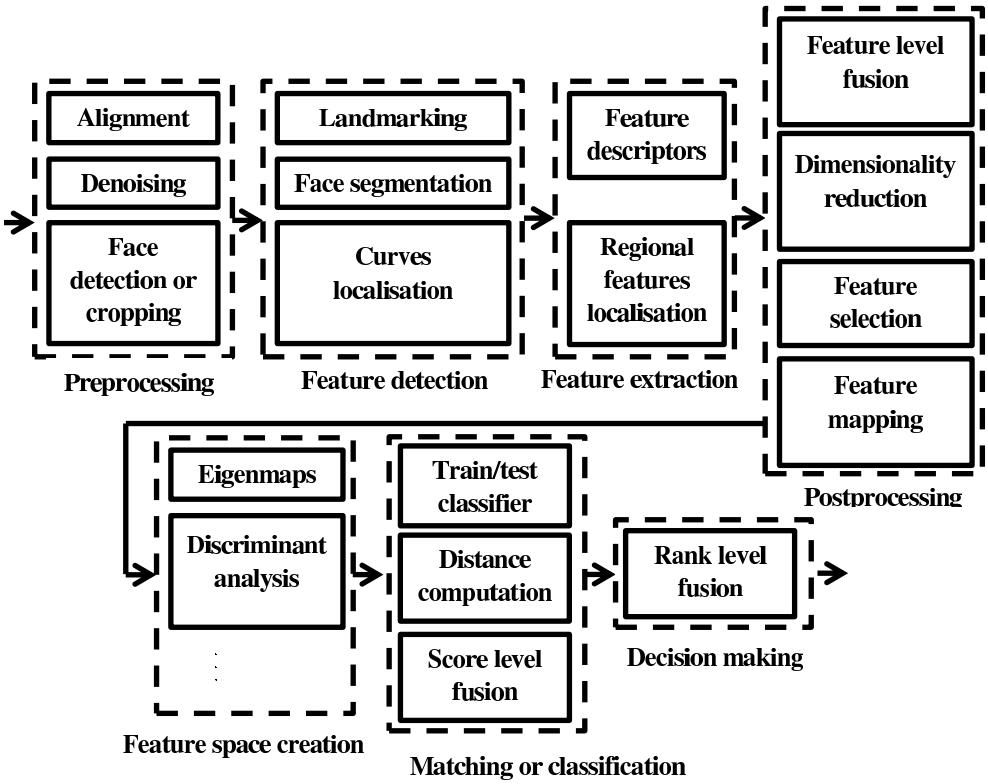


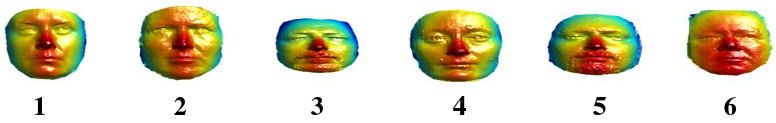
Figure 4: Different steps of a 3D face recognition algorithm (image from (Emambakhsh, 2014)).

nose can be easily segmented on the face. Furthermore, hiding the nose in real biometric sessions is nearly impossible, without attracting suspicion.

The results of a similar experiment performed over the nasal region is illustrated in Figure 6 for the same subjects. Similarly, the identity of the neutral face has been correctly recognised as the lowest matching error is produced by ICP for subject 1. However, for other problematic cases, the nasal region has been able to lead to the correct identity at lower ranks. For example, for the facial expression case in Figure 6-f, although the matching error is lowest for the wrong subject, the second lowest matching error correspond to the correct subject label. However, for the whole face experiment in Figure 6-f, it occurs in the third rank.

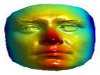
There are a few papers which have focused specifically on 3D nose recognition (Drira et al., 2009; Dibeklioğlu et al., 2009) where the 3D information of the nose region is utilised. Drira *et al.* present another approach for 3D face recognition using the nose region (Drira et al., 2009) demonstrating better performance. Their method utilises the geodesic contours, used in (Bronstein et al., 2005), only on the nose region. After denoising and nose segmentation (by fitting a sphere with radius 100 mm on the nose tip), the concentric contours are found on the nose surface using Dijkstra algorithm (Dijkstra, 1959). Dibeklioglu *et al.* first perform different curvature calculations on the face surface and then segment the nose region (Dibeklioğlu et al., 2009). Then, the ICP algorithm is used for recognition.

Moorhouse *et al.* introduced another approach for 3D nose recognition (Moorhouse et al.,



1 **2** **3** **4** **5** **6**

(a) Gallery consisting of six captures, one each from subjects 1 to 6.



1.79 2.655 3.49 1.895 2.045 1.88

(b) Probe with neutral expression and frontal view and corresponding match errors.

10.66 13.22 **8.89** 13.85 9.74 13.425

(c) Probe with occlusion and corresponding match errors.

5.57 5.92 5.55 5.31 5.84 **5.24**

(d) Probe with different pose and corresponding match errors.

5.11 5.61 **4.91** 5.16 5.00 5.015

(e) Noisy probe and corresponding match errors.

2.53 2.63 2.92 2.24 2.945 **2.22**

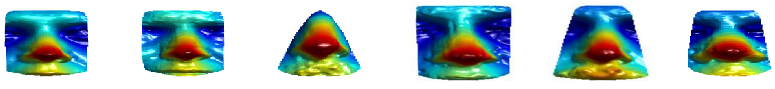
(f) Probe with non-neutral expression and corresponding match errors.

Figure 5: Matching results using five different probe samples of subject 1 in the gallery with variations caused by occlusion, pose, noise and expression (image from (Emambakhsh, 2014)).

2009). Colour information in YCbCr domain is clustered by GMM and curvature calculation is used for landmark detection and nose segmentation. Unlike the previous three 3D nose recognition papers, instead of using laser scanners, photometric stereo imaging is utlised to make the dataset. Four different features are evaluated on the nose surface: geometric ratios, the Fourier descriptors (FD) of the ridge, combination of these features and eigennoses.

A combination of the nasal region, forehead and eyes are used for a 2D/3D face recognition by Mian *et al.* (Mian et al., 2007). A modified ICP algorithm is used for matching, in conjunction with a pattern rejector based on spherical face representation (SFR) and scale-invariant feature transform (SIFT). As a result high recognition ranks are achieved, in particular for neutral probes.

Finally, in the section bellow we describe an example of a highly successful face recognition algorithm based on the 3D nasal curves.



1	2	3	4	5	6

(a) Gallery consisting of six captures, one each from subjects 1 to 6.

0.50	1.14	7.68	1.33	2.55	2.40

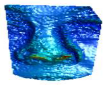
(b) Probe with neutral expression and frontal view and corresponding match errors.

8.66	8.89	11.94	**8.53**	8.96	8.91

(c) Probe with occlusion and corresponding match errors.

1.70	1.915	9.33	2.44	3.82	3.63

(d) Probe with different pose and corresponding match errors.



7.66	7.95	12.25	7.73	8.42	8.47

(e) Noisy probe and corresponding match errors.

1.92	**1.65**	7.62	2.00	3.46	2.79

(f) Probe with non-neutral expression and corresponding match errors.

Figure 6: Matching results using the nasal regions of five different probe samples of subject 1 in the gallery with variations caused by occlusion, pose, noise and expression (image from (Emambakhsh, 2014)).

3.2 Face recognition using the nasal curves

Preprocessing and segmentation

The face is first denoised using a 2.3×2.3 median filtering. Principal component analysis (PCA) is then iteratively performed over the point clouds to correct the pose (Mian et al., 2007). Then the convex areas are located by thresholding the shape index map. The centroid of the largest connected component in the convex map is selected as the nose tip location. The nasal region is then cropped using the method proposed in (Emambakhsh et al., 2013), by intersecting two horizontal and one vertical cylinder.

Nasal region landmarking

The resulting nasal region is shown in Figure 7. In addition to the nose tip $\mathbf{L9}$, three other landmarks are detected: nasal root ($\mathbf{L1}$) and nasal alar groove ($\mathbf{L5}$ and $\mathbf{L13}$). In order to detect $\mathbf{L1}$,

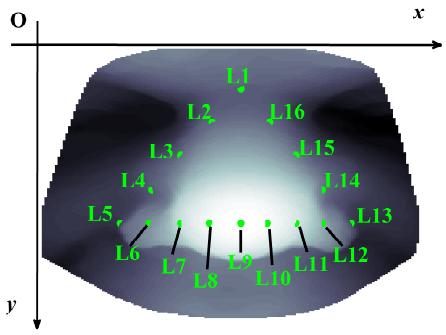


Figure 7: The locations of the landmarks and their names (image from (Emambakhsh et al., 2013)).

several orthogonal planes to the nasal region are intersected with the nasal region. The angle between the i^{th} plane and the y-axis is denoted as α_i. Therefore, the normal vector for the orthogonal plane will be $[\cos\alpha_i, \sin\alpha_i, 0]$, which passes $\mathbf{L9}$. This results in the blue curves shown in Figure 8. For each curve, the global minimum is detected, shown in green in Figure 8. Connecting these points creates a new curve, whose global maximum corresponds to the nasal root $\mathbf{L1}$.

The algorithm proposed in (Segundo et al., 2010) is used to detect $\mathbf{L5}$ and $\mathbf{L13}$. First, an orthogonal plane passing through $\mathbf{L9}$ with normal vector $[0, 1, 0]$ is intersected with the nasal region. Then the minima of the left and right sides of $\mathbf{L9}$ over the intersected curve are detected by computing the first and second differentiation of the curve.

The locations of $\mathbf{L1}$, $\mathbf{L5}$ and $\mathbf{L13}$ are used to create the other landmarks shown in Figure 7. These are simply found by equally dividing the connecting lines between the landmarks and then mapping the middle points to the nasal region.

Nasal curves and feature selection

The found landmarks are used to create the nasal curves shown in Figure 9. The depth map components of each curves is extracted and concatenated to form the feature space. The facial curves are excellent feature descriptors containing rich information in lower dimensional space than the

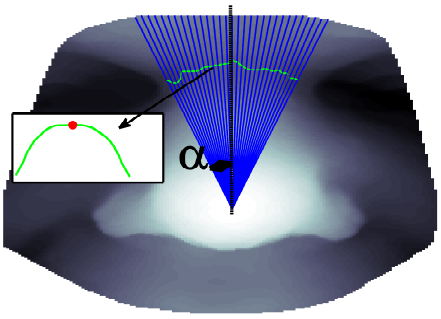


Figure 8: $\mathbf{L1}$ detection procedure: the blue lines are the planes' intersections, the green curve is each intersection's minimum and $\mathbf{L1}$ is given by the maximum value of the minima, shown with a red dot (image from (Emambakhsh et al., 2013)).

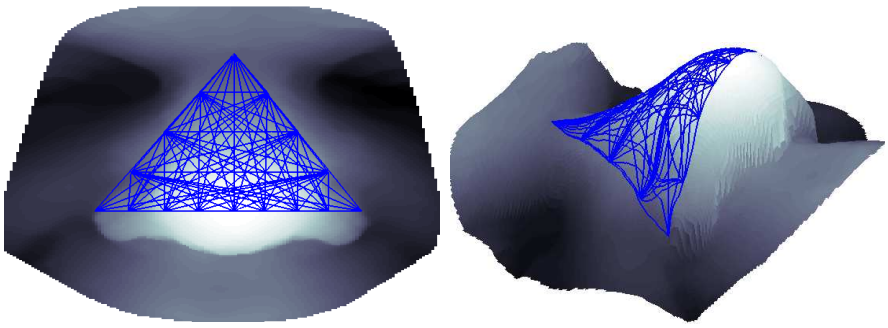


Figure 9: The nasal curves: (a) Frontal and (b) side view (image from (Emambakhsh et al., 2013)).

original 3D point cloud. However, some of these curves are more sensitive to the facial expression while some are more robust.

In order to detect those expression robust curves, the forward sequential feature selection (FSFS) is used to select the most robust feature set. For each selected feature set, the rank-one recognition rate is computed. At each iteration, the feature set combination which generates the highest recognition rate is stored. The one vs. all scenario is performed for the matching step using a city-block (CB) distance computation. The feature selection at different iterations of FSFS is shown in Figure 10. While the distribution of these curves is relatively even over the nasal surface, it is slightly denser on the nasal cartilage, which is less flexible due to its bony structure. This method demonstrated a significant improvement over existing methods using standard datasets (Emambakhsh et al., 2013).

4 Multimedia security

As digital technologies have shown a rapid growth within the last decade, content protection now plays a major role within content management systems. Of the current systems, digital watermarking provides a robust and maintainable solution to enhance media security. Evidence of popularity of watermarking is clearly visible as watermarking research has resulted in 14685 *image & video watermarking* papers published in last 20 years and 1818 (12.4%) alone in 2014-15[2]. This section discusses common watermarking philosophies, applications and recent researches in the domain.

4.1 Definition, properties, applications and attacks

By definition a digital watermark is the copyright or author identification information which is embedded directly in the digital media in such a way that it is imperceptible, robust and secure. The watermarking research is considerably mature by now, after its major inception in mid nineties and offers digital protection to a wide spectrum of application as shown in Figure 11. Cox *et al.* (Cox et al., 2000) listed various applications of watermarking including *broadcast monitoring*, *owner identification*, *proof of ownership*, *authentication*, *transactional watermarking*, *copy control*

[2]Sources: `www.scopus.com`

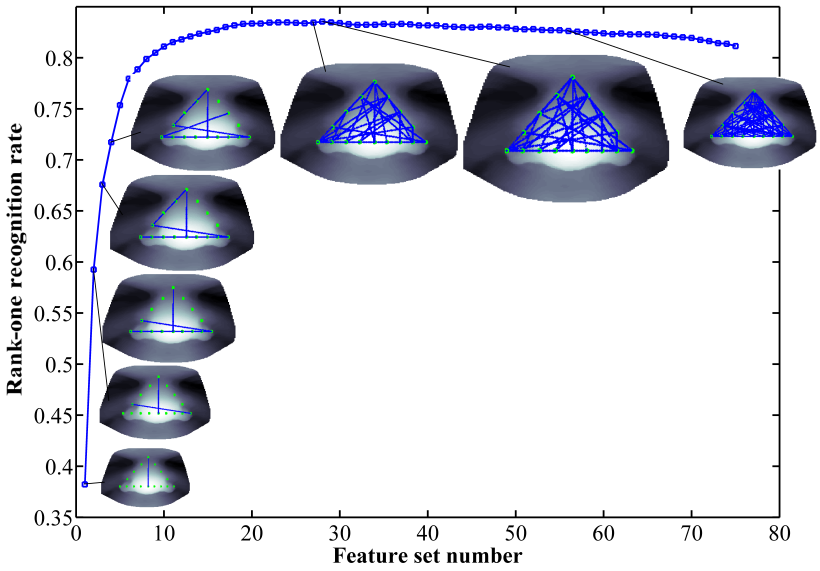


Figure 10: Rank-one recognition rate against the number of nasal curves selected by the FSFS algorithm. The sets of curves for selected feature sets are also shown, with the largest image (second from right) showing the 28 curves that produced the highest recognition rate (Emambakhsh et al., 2013).

and *covert communication*. A few of them are followed in the watermarking industry[3] with few additions such as *Audience measurement* and *improved auditing*. Image quality evaluation methods were proposed in the literature where a watermark is embedded either in the discrete wavelet transform (DWT) (Wang et al., 2007) or discrete cosine transform (DCT) (Nezhadarya et al., 2009) frequency domain and the degradation of the extracted watermark was used to determine the quality without any reference to the original image. Uehira *et al.* (Uehira et al., 2016) proposed a new application of image displays to invisible optoelectronic watermarking systems where a paired installer is used to transfer auxiliary data using watermark code patterns. The authors used flat-panel image display and smart phones to establish an optical link between them, enabling the viewer to receive auxiliary data on their smart devices. An improved management of medical application was proposed in (Tsai et al., 2015) where watermarking techniques were used to embed patient data such as identity (ID), serial number or region of interest to ensure the image association with the correct patient and to indicate its relationship to other images in a series. Yamada *et al.* (Yamada et al., 2016) developed a real-time watermarking system for video-on-demand services where frame images are watermarked, unique to the user, when a server receives request from a user. The system aims to deter piracy. Another application includes a method for providing royalty payments for content distributed via a network (Levy and Stager, 2012).

Digital watermarking comprises elements from a variety of disciplines including image processing, video processing, telecommunication, computer science, cryptography, remote sensing and geographical information systems etc. Watermarking systems are often characterized by a set

[3] `www.digitalwatermarkingalliance.org/applications.asp`

Watermarking Applications		
Sl No.	**Name**	**Description**
1	Broadcast monitoring	Passive monitoring by the automatic watermark detection of broadcasted watermarked media.
2	Copyright identification	Resolving copyright issues of digital media by using watermark information as copyright data.
3	Content authentication	Authentication of original art work, performance and protection against digital forgery.
4	Access control	Access control applications, such as, Pay-TV.
5	Copy control	Disabling copy of CD / DVD etc. by watermarked permission.
6	Packaging and tracking	Transaction tracking and protection against forged consumable items including pharmaceutical products by embedding watermark on packaging.
7	Medical record authentication	Authentication of digitally preserved patient's medical record, including blood sample, X-ray, ECG etc.
8	Insurance / Banking document authentication	Digital authentication of insurance claim, banking, financial, mortgage and corporate documents.
9	Media piracy control	Tracking of the source of media piracy.
10	Ownership identification	Supproting legitimate claim, such as, royalty by the media owner.
11	Transaction tracking	Tracking of media ownership in a buyer-seller scenario.
12	Meta-data hiding	Hiding meta-data within the media instead of a big header.
13	Video summary creation	Instant retrieval of video summary by embedding the summary within the host video.
14	Video hosting authentication	Piracy control by video authentication at video hosting servers, including youtube, megavideo etc.

Figure 11: Watermarking applications.

of common properties and the importance of each property depends on the application requirements. A list of such properties and corresponding example applications(Cox et al., 2002; Barni and Bartolini, 2004) are shown in Figure 12, where last column of the figure shows the associated applications' number from Figure 11.

Based on the embedding method, the watermarking techniques can be categorized as shown in Figure 13. The watermark embedding can be done in the spatial domain or in the frequency domain. The latter have been a much popular choice as frequency decomposition characterizes the host media to represent the human eye characteristics and eye perception towards the media. Therefore frequency domain watermarking can provide better insight to reduce embedding distortion or increase the robustness (Cox et al., 1997). Now, depending on the type of host media, watermarking can be divided into two categories: image and video watermarking. Again, based on the human perception the watermarking schemes can be categorized as visible or imperceptible (invisible) watermarking and the latter can also be categorized as robust, fragile or semi-fragile watermarking. In case of a robust watermarking scheme, the watermark is expected to be sustained even after a compression or any other intentional attack, whereas in the case of a fragile scheme (Fridrich et al., 2000) the watermark information is usually destroyed to any alteration or attack to the media, in order to authenticate the image integrity. A semi-fragile scheme (Lin and Chang, 2000) represents properties from both the above mentioned categories and the watermark information is robust to certain type of attacks while fragile to other type of attacks.

Watermark represents the owner's identity. Hence the selection of the watermark is considered important and varies according to application requirements. Early days of watermarking scheme

General Properties of Digital Watermarking		
Properties	Description	Applications
Imperceptibility	The watermark should not noticeably distort or degrade the host data in order to preserve the quality of the marked document.	1, 2, 3, 4, 14.
Robustness	To measure robustness the watermark must be reliably detectable against signal processing schemes including data compression.	
Fragility	These kinds of watermark are embedded in host data in such a way that they do not survive in the case of any modification even copying.	7, 8.
Tamper-resistance	The tamper-resistance property is focused on the intentional attacks in contrast to robustness.	3, 5, 9, 10, 11, 14.
False positive rate	The probability of identifying an un-watermarked piece of data as containing a watermark by a detector is called the false positive rate.	6.
Data payload	The amount of information present in watermarked media is called data payload.	12, 13.

Figure 12: Watermarking properties and associated applications.

often used a pseudo-random number to embed the watermark and authenticity of the media is examined by the presence or absence of the watermark. In recent literature a message or logo based watermark (Kundur and Hatzinakos, 2004) has been preferred by the researchers and in this case authentication is done by extracting the hidden message or logo to identify the legitimate owner. Figure 14 shows the different types of watermark used in this field.

Main requirements of the watermarking schemes are either 1) to retain the watermark information after any intentional attacks or natural image/video processing operation, or 2) to identify any tampering (fragile watermarking) of the target media. Any process that modifies the host media affecting the watermark information, is called attack on watermarking. Various types of attacks can be grouped together as follows: 1) signal processing, 2) geometric, 3) enhancement, 4) printing-scanning-capturing, 5) oracle, 6) chrominance, 7) transcoding attacks etc. The attack characterization with respect to image and video watermarking and related applications are shown in Figure 15.

While watermarking schemes in general are evaluated in terms of imperceptibility (Asikuzzaman et al., 2014; Koz and Alatan, 2008), robustness against various intentional (Fallahpour et al., 2014) and unintentional (*e.g.*, compression (Sttz et al., 2014), filtering (Zhu et al., 2014) or geometric (Zhang et al., 2011)) attacks and fragility (Chan et al., 2015), security of the watermarking schemes are also reported in (Bianchi and Piva, 2013; Adelsbach et al., 2004). The security of the watermark can be defined as the ability to properly conceal the watermark information in such way that it is secret to the unauthorized users. The security of the watermarking schemes are usually implemented using two different approaches (Adelsbach et al., 2004):

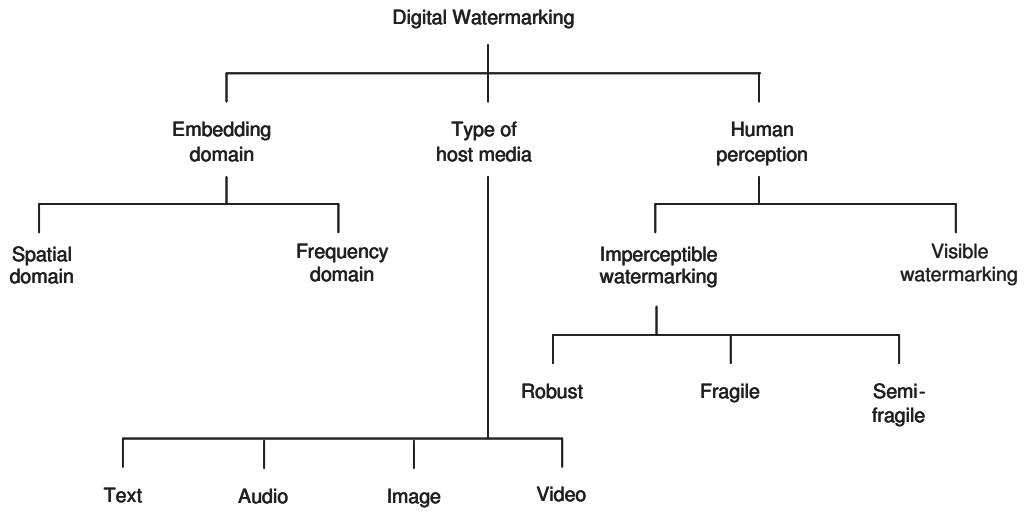


Figure 13: Types of watermarking techniques.

- *Asymmetric watermarking* which uses two different keys for watermark embedding and detection and
- *Zero-knowledge watermark detection* using cryptographic techniques where the watermark detection process is substituted by cryptographical protocol.

Cryptographical scrambling of the watermark logo is also used in order to secure the watermark (Kundur and Hatzinakos, 2004) in addition to the other security measures, such as, key based coefficient selection, random filter parameter selection etc. These are particularly useful when the attacker has access to the watermark detector.

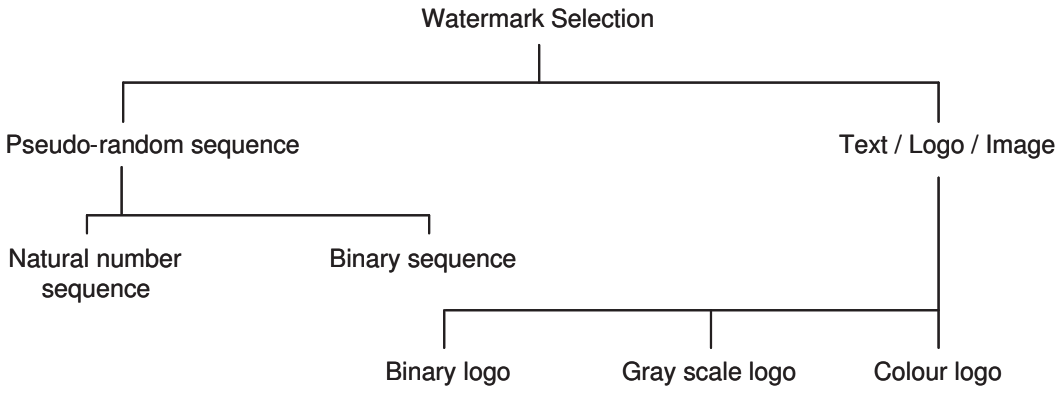


Figure 14: Watermark types.

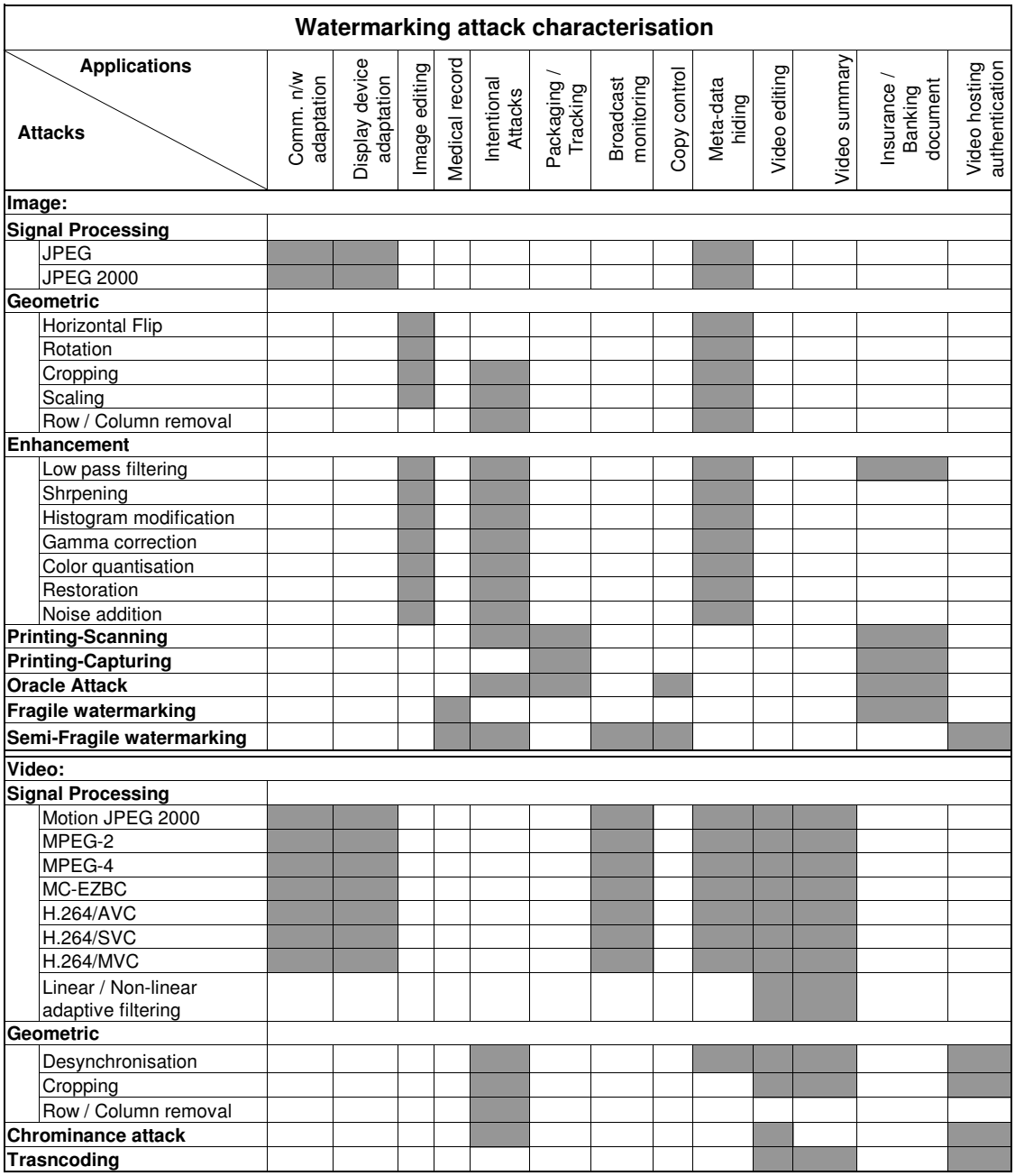


Figure 15: Attack characterization.

4.2 Watermarking process

The watermarking procedure, in its basic form, consists of two main processes: 1) Embedding and 2) Extraction and authentication. At this point, for simplicity, we describe these processes with reference to the image watermarking.

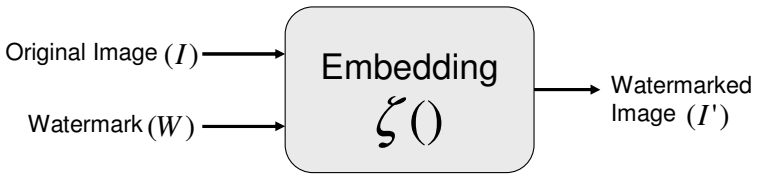


Figure 16: Watermark embedding process.

Embedding

This process insert or embed the watermark information within the host image by modifying all or selected pixel values (spatial domain); or coefficients (frequency domain), in such a way that the watermark is imperceptible to human eye and is achieved by minimizing the embedding distortion to the host image. The system block for the embedding process is shown in Figure 16 amd can be expressed as:

$$I' = \zeta(I, W), \tag{1}$$

where I' is the watermarked image, I is the original host image, W is the watermark information and $\zeta()$ is the embedding function.

The embedding function can further be categorized in sub-processes: 1) forward transform (for frequency domain), 2) pixel / coefficient selection, 3) embedding method (additive, multiplicative, quantization etc.) and 4) inverse transform.

Finally the performance of the watermark embedding is measured by comparing the watermarked image (I') with the original unmarked image (I) and is calculated by various metrics: 1) peak signal to noise ratio (PSNR), 2) weighted PSNR (wPSNR) (Kwon and Lee, 2001), 3) structural similarity measure (SSIM) (Wang et al., 2004), 4) just noticeable difference (JND) (Watson, 1993) and 5) subjective quality measurement (Koumaras, 2008).

PSNR: This is one of the most commonly used visual quality metric which is based on the root mean square error (RMSE) of the two images with dimension of $X \times Y$ as in Eq. (2) and Eq. (3).

$$PSNR = 20 \log_{10} \left(\frac{255}{RMSE} \right) \quad dB. \tag{2}$$

$$RMSE = \sqrt{\frac{1}{X \times Y} \sum_{m=0}^{X-1} \sum_{n=0}^{Y-1} (I(m,n) - I'(m,n))^2}. \tag{3}$$

wPSNR: On contrary to error measurement in spatial domain as in the previous metric, this metric measures PSNR in wavelet transform domain with weighting factors at different frequency decomposition level. The host and processed images are firstly wavelet decomposed and then squared error is computed at every subband. Finally wPSNR is calculated using cumulative squared error with weighting parameters for each subband. The weights for various subbands are adjusted in such way that wPSNR has the highest correlation with the subjective score.

SSIM: This quality measurement metric assumes that human visual system is highly adapted for extracting structural information from a scene. Unlike PSNR, where average error between two images taken into consideration, SSIM focuses on a quality assessment based on the degradation of structural information. The structural information in the scene is calculated using local luminance and contrast rather an average luminance and contrast.

JND: In this metric the host and test images are DCT transformed and Just Noticeable Differences are measured using thresholds. The thresholds are decided based on 1) luminance masking and 2) contrast masking of the transformed images. The threshold for luminance pattern relies on the mean luminance of the local image region, whereas the contrast masking is calculated within a block and particular DCT coefficient using a visual masking algorithm.

Subjective: Although various objective metrics have been proposed to measure the visual quality, often by modeling the human visual system or subjective visual tests, the subjective test offers best visual quality measurement. Subjective tests procedures are recommended by ITU (Koumaras, 2008) which defines the specification of the screen, luminance of the test room, distance of the observer from the screen, scoring techniques, test types such as double stimulus continuous quality test (DSCQT) or double stimulus impairment scale test (DSIST) etc. The tests are carried out with multiple viewer and the mean opinion score (MOS) represents the visual quality of the test image. However the subjective tests are often time consuming and difficult to perform and hence researchers prefer objective metrics to measure the visual quality.

Among these metrics, due its simplicity, the most common method of evaluating the embedding performance in watermarking research is PSNR. It is also observed that most of the metrics behaves in a similar fashion when compared with any embedding distortion measured by PSNR of $35dB$ or above.

Extraction and authentication

As the process name suggested, it consists of two subprocess: 1) extraction of watermark and 2) authentication of the extracted watermark. The watermark extraction follows a reverse embedding algorithm, but with a similar input parameter set. Now based on the watermark extraction criteria any watermarking method can be categorized in: 1) non-blind type and 2) blind type. For the first category, a copy of the original un-watermarked image is required during extraction whereas in the latter case, the watermark is extracted from the test image itself. The extraction process can be written in the simplified form as:

$$W' = \varpi(I', I), \tag{4}$$

where W' is the extracted watermark, I' is the test image, I is the original image and $\varpi()$ is the extraction function.

Once the watermark is extracted from the test image, the authentication is performed by comparing with the original input watermark information. Common authentication methods are defined by finding the closeness between the two in a vector space, by calculating the similarity correlation or Hamming distance. A complete system diagram of extraction and authentication process

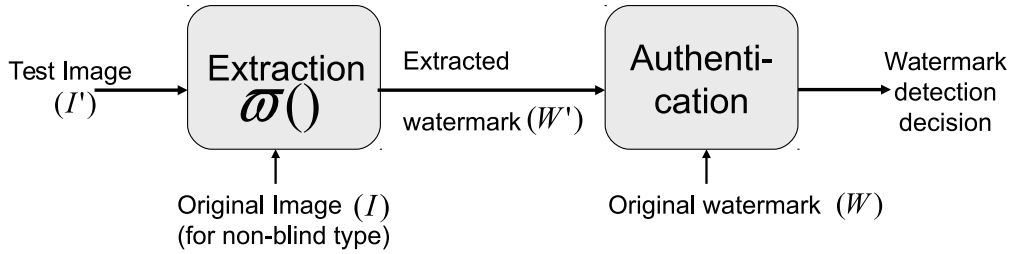


Figure 17: Watermark extraction and authentication process.

is shown in Figure 17. An example of image watermark embedding and extraction processes are shown in Figure 18(a) and Figure 18(b), respectively. The figure also shows the difference image demonstrating how the watermark information is distributed within the textured region so that it is imperceptible to human vision.

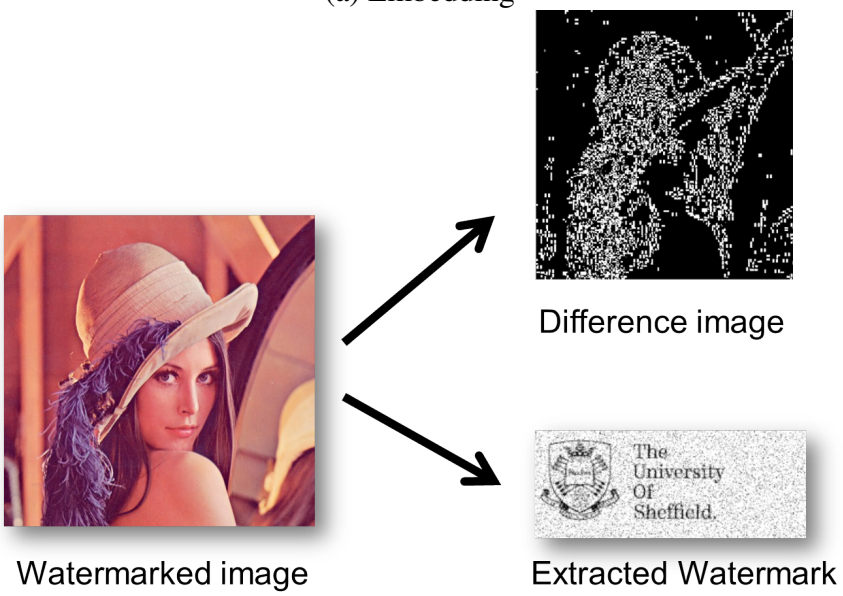
4.3 Recent research trends

The visual quality of host media (often known as imperceptibility) and robustness are widely considered as the two main properties vital for a good digital watermarking system. They are complimentary to each other and hence challenging to attain the right balance between them. Frequency-based watermarking, more precisely wavelet domain watermarking, methodologies are highly favoured in the current research era. The wavelet domain is also compliant within many image coding, *e.g.*, JPEG2000 (Taubman and Marcellin, 2002) and video coding, *e.g.*, Motion JPEG2000, Motion-Compensated Embedded Zeroblock Coding (MC-EZBC) (Chen and Woods, 2004), schemes, leading to smooth adaptability within modern frameworks. Due to the multi-resolution decomposition and the property to retain spatial synchronisation, which are not provided by other transforms (the Discrete Cosine Transform (DCT) for example), the Discrete Wavelet Transform (DWT) provides an ideal choice for robust watermarking (Bhowmik and Abhayaratne, 2016; Piper et al., 2005; Bhowmik and Abhayaratne, 2008; Soheili, 2010; Abhayaratne and Bhowmik, 2013; Feng and Yang, 2005; Xia et al., 1998; Xie and Arce, 1998; Barni et al., 2001; Bhowmik et al., 2016; Kim and Moon, 1999; Marusic et al., 2003; Zhang and Mo, 2001).

When designing a watermarking scheme there are numerous features to consider, including the wavelet kernel, embedding coefficients and wavelet subband selection. Each of these particular features can sufficiently impact the overall watermark characteristics (Bhowmik and Abhayaratne, 2014) and is largely dependant upon the target application requirements.

a) Wavelet Kernel Selection: An appropriate choice of wavelet kernel must be determined within the watermarking framework. There have been previous studies to show that the performance of watermark robustness and imperceptibility is dependant on the wavelet kernels (Bhowmik and Abhayaratne, 2008, 2009). The orthogonal Daubechie wavelets are a favourable choice with many early watermarking schemes (Feng and Yang, 2005; Xia et al., 1998; Xie and Arce, 1998; Barni et al., 2001; Kundur and Hatzinakos, 2004), although the later introduction of bi-orthogonal wavelets within the field of digital watermarking has increased their usage (Kim and Moon, 1999; Marusic et al., 2003; Zhang and Mo, 2001).

(a) Embedding



(b) Extraction

Figure 18: An image watermark embedding and extraction example.

b) Host Coefficient Selection: Various approaches exist to choose suitable transform coefficients for embedding a watermark. In current methods, coefficient selection is determined by the threshold values based upon the coefficient magnitude (Kim and Moon, 1999) or a pixel masking approach based upon HVS (Barni et al., 2001) or the median of 3 coefficients in a 3×1 overlapping window (Xie and Arce, 1998) or simply by selecting all the coefficients (Xia et al., 1998; Feng and Yang, 2005; Kundur and Hatzinakos, 2004).

c) Wavelet Subband Selection: The choice of subband bears a large importance when determining the balance between robustness of the watermark and imperceptibility. Embedding within the high frequency domain subbands (Huo and Gao, 2006; Barni et al., 2001; Feng and Yang, 2005; Kundur and Hatzinakos, 2004; Xia et al., 1998) can often provide great imperceptibility but with limited watermark robustness capabilities. Contradictory schemes embed data only within the low frequency subbands (Xie and Arce, 1998; Zhang and Mo, 2001) aimed towards providing a high

robustness. Spread spectrum domain embedding (Dey et al., 2012; Kim and Moon, 1999; Chen et al., 2003) modifies data across all frequency subbands, ensuring a balance of both low and high frequency watermarking characteristics. The number of decomposition levels is also an important factor. Previous studies have researched watermarking schemes using two (Huo and Gao, 2006; Feng and Yang, 2005; Xia et al., 1998), three (Marusic et al., 2003; Zhang and Mo, 2001) and four or more (Barni et al., 2001; Kundur and Hatzinakos, 2004) wavelet decomposition levels.

An extension of image watermarking to video is the easiest option for any video watermarking scheme. Frame-by-frame video watermarking (Hartung and Girod, 1998; Cox et al., 1997) and 3D wavelet based video watermarking schemes (Campisi, 2005; Kim et al., 2004) are available in the literature. However a direct extension of the image watermarking schemes without consideration of motion, produces flicker and other motion related mismatch. The watermarking algorithms along with MCTF, which decomposes motion information, provides a better solution to this problem (Vinod and Bora, 2006; Bhowmik and Abhayaratne, 2012).

5 Conclusions

In this chapter, we discussed various image processing techniques used in visual surveillance and security applications covering visual tracking, biometric application and digital media security. Individual categories provided a general introduction to the topic, discussed the commonly used techniques by dissecting the approaches proposed in the literature. Suitable examples are also provided for better understanding. The comprehensive bibliography is useful for the readers interested in further research.

Bibliography

Charith Abhayaratne and Deepayan Bhowmik. Scalable watermark extraction for real-time authentication of JPEG 2000 images. *Journal of real-time image processing*, 8(3):307–325, 2013.

A. Adelsbach, S. Katzenbeisser, and A. R. Sadeghi. Cryptography meets watermarking: Detecting watermarks with minimal or zero knowledge disclosure. In *Proc. European Signal Processing Conference (EUSIPCO)*, volume 1, pages 446–449, 2004.

A. Ali and J. K. Aggarwal. Segmentation and recognition of continuous human activity. In *Proc. IEEE Workshop on Detection and Recognition of Events in Video*, pages 28–35, 2001.

M.S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Trans. on Signal Processing*, 50(2):174–188, Feb 2002.

M. Asikuzzaman, M. J. Alam, A. J. Lambert, and M. R. Pickering. Imperceptible and robust blind video watermarking using chrominance embedding: A set of approaches in the dt cwt domain. *IEEE Transactions on Information Forensics and Security*, 9(9):1502–1517, Sept 2014.

S. Avidan. Support vector tracking. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 26(8):1064–1072, Aug 2004.

Y. Bar-Shalom and T. Foreman. *Tracking and Data Association*. Academic Press Inc., 1988.

M. Barni and F. Bartolini. *Watermarking Systems Engineering (Signal Processing and Communications, 21)*. CRC Press, Inc., Boca Raton, FL, USA, 2004. ISBN 0824748069.

M. Barni, F. Bartolini, and A. Piva. Improved wavelet-based watermarking through pixel-wise masking. *IEEE Trans. Image Processing*, 10(5):783–791, May 2001.

Paul J Best and Neil D McKay. A method for registration of 3-D shapes. *IEEE Transactions on pattern analysis and machine intelligence*, 14(2):239–256, 1992.

Deepayan Bhowmik and Charith Abhayaratne. A generalised model for distortion performance analysis of wavelet based watermarking. In *Digital Watermarking*, pages 363–378. Springer, 2008.

Deepayan Bhowmik and Charith Abhayaratne. Embedding distortion modeling for non-orthonormal wavelet based watermarking schemes. In *Proc. SPIE Wavelet App. in Industrial Processing VI*, volume 7248, page 72480K (12 Pages), 2009.

Deepayan Bhowmik and Charith Abhayaratne. 2D+t wavelet domain video watermarking. *Advances in Multimedia*, 2012:6, 2012.

Deepayan Bhowmik and Charith Abhayaratne. On Robustness Against JPEG2000: A Performance Evaluation of Wavelet-Based Watermarking Techniques. *Multimedia Syst.*, 20(2):239–252, 2014.

Deepayan Bhowmik and Charith Abhayaratne. Quality scalability aware watermarking for visual content. *IEEE Transactions on Image Processing*, 25(11):5158–5172, Nov 2016. doi: 10.1109/TIP.2016.2599785.

Deepayan Bhowmik, Matthew Oakes, and Charith Abhayaratne. Visual attention-based image watermarking. *IEEE Access*, PP(99):1–1, 2016. doi: 10.1109/ACCESS.2016.2627241.

T. Bianchi and A. Piva. Secure watermarking for multimedia content protection: A review of its benefits and open issues. *IEEE Signal Processing Magazine*, 30(2):87–96, March 2013.

S.S. Blackman. Multiple hypothesis tracking for multiple target tracking. *IEEE Aerospace and Electronic Systems Magazine*, 19(1):5–18, Jan 2004.

Bernhard E. Boser, Isabelle M. Guyon, and Vladimir N. Vapnik. A training algorithm for optimal margin classifiers. In *Proc. Fifth Annual Workshop on Computational Learning Theory*, COLT '92, pages 144–152, 1992.

Yuri Boykov and Gareth Funka-Lea. Graph cuts and efficient N-D image segmentation. *International Journal of Computer Vision*, 70(2):109–131, 2006.

Ted J. Broida and R. Chellappa. Estimation of object motion parameters from noisy images. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, PAMI-8(1):90–99, Jan 1986.

Alexander M. Bronstein, Michael M. Bronstein, and Ron Kimmel. Three-dimensional face recognition. *International Journal of Computer Vision*, 64:5–30, 2005.

Patrizio Campisi. Video watermarking in the 3D-DWT domain using quantization-based methods. In *Multimedia Signal Processing, 2005 IEEE 7th Workshop on*, pages 1–4. IEEE, 2005.

John Canny. A computational approach to edge detection. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, PAMI-8(6):679–698, Nov 1986.

H. T. Chan, W. J. Hwang, and C. J. Cheng. Digital hologram authentication using a hadamard-based reversible fragile watermarking algorithm. *Journal of Display Technology*, 11(2):193–203, Feb 2015.

Peisong Chen and John W Woods. Bidirectional MC-EZBC with lifting implementation. *Circuits and Systems for Video Technology, IEEE Transactions on*, 14(10):1183–1194, 2004.

T.-S. Chen, J. Chen, and J.-G. Chen. A simple and efficient watermarking technique based on JPEG2000 codec. In *Proc. Int'l Symp. on Multimedia Software Eng.*, pages 80–87, 2003.

D. Comaniciu and P. Meer. Mean shift: a robust approach toward feature space analysis. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 24(5):603–619, May 2002.

D. Comaniciu, V. Ramesh, and P. Meer. Kernel-based object tracking. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 25(5):564–577, May 2003.

I. J. Cox, J. Kilian, F. T. Leighton, and T. Shamoon. Secure spread spectrum watermarking for multimedia. *IEEE Trans. Image Processing*, 6(12):1673–1687, Dec. 1997.

I. J. Cox, M. L. Miller, and J. A. Bloom. Watermarking applications and their properties. In *Information Technology: Coding and Computing, 2000. Proceedings. International Conference on*, pages 6–10, 2000.

I. J. Cox, M. L. Miller, and J. A. Bloom. *Digital watermarking*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2002. ISBN 1-55860-714-5.

N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *Proc. IEEE CVPR*, volume 1, pages 886–893, 2005.

N. Dey, M. Pal, and A. Das. A session based blind watermarking technique within the nroi of retinal fundus images for authentication using dwt, spread spectrum and harris corner detection. *International Journal of Modern Engineering Research*, 2:749–757, 2012.

H. Dibeklioğlu, B. Gökberk, and L. Akarun. Nasal region-based 3D face recognition under pose and expression variations. In *3rd International Conference on Advances in Biometrics*, pages 309–318, 2009.

E.W. Dijkstra. A note on two problems in connexion with graphs. *Numerische Mathematik*, 1: 269–271, 1959.

H. Drira, B.B. Amor, M. Daoudi, and A. Srivastava. Nasal region contribution in 3D face biometrics using shape analysis framework. In *3rd International Conference on Advances in Biometrics*, pages 357–366, 2009.

R.M. Dufour, E.L. Miller, and N.P. Galatsanos. Template matching based object recognition with unknown geometric parameters. *IEEE Trans. on Image Processing*, 11(12):1385–1396, Dec 2002.

G.J. Edwards, C.J. Taylor, and T.F. Cootes. Interpreting face images using active appearance models. In *Proc. IEEE International Conference on Automatic Face and Gesture Recognition*, pages 300–305, Apr 1998.

M. Emambakhsh, A.N. Evans, and M. Smith. Using nasal curves matching for expression robust 3D nose recognition. In *6th IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, pages 1–6, 2013.

Mehryar Emambakhsh. *Using the 3D shape of the nose for biometric authentication*. PhD thesis, University of Bath, 2014.

M. Fallahpour, S. Shirmohammadi, M. Semsarzadeh, and J. Zhao. Tampering detection in compressed digital video using watermarking. *IEEE Transactions on Instrumentation and Measurement*, 63(5):1057–1072, May 2014.

X. C. Feng and Y. Yang. A new watermarking method based on DWT. In *Proc. Int'l Conf. on Computational Intelligence and Security, Lect. Notes in Comp. Sci. (LNCS)*, volume 3802, pages 1122–1126, 2005.

Yoav Freund and Robert E Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55(1):119 – 139, 1997.

J. Fridrich, M. Goljan, and A. C. Baldoza. New fragile authentication watermark for images. In *Proc. IEEE ICIP*, volume 1, pages 446 –449, 2000.

Niels Haering, PterL. Venetianer, and Alan Lipton. The evolution of video surveillance: an overview. *Machine Vision and Applications*, 19(5-6):279–290, 2008.

Chris Harris and Mike Stephens. A combined corner and edge detector. In *Proc. of Fourth Alvey Vision Conference*, pages 147–151, 1988.

Frank Hartung and Bernd Girod. Watermarking of uncompressed and compressed video. *Signal processing*, 66(3):283–301, 1998.

Berthold K.P. Horn and Brian G. Schunck. Determining optical flow. *Artificial Intelligence*, 17: 185–203, 1981.

F. Huo and X. Gao. A wavelet based image watermarking scheme. In *Proc. IEEE ICIP*, pages 2573–2576, 2006.

Z.L. Husz, A.M. Wallace, and P.R. Green. Tracking with a hierarchical partitioned particle filter and movement modelling. *IEEE Trans. on Systems, Man, and Cybernetics, Part B: Cybernetics*, 41(6):1571–1584, Dec 2011.

A.K. Jain, A. Ross, and S. Pankanti. Biometrics: a tool for information security. *IEEE Transactions on Information Forensics and Security*, 1:125–143, 2006.

A.N. Jain, A. Ross, and S. Prabhakar. An introduction to biometric recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(1):4–20, 2004.

P. KaewTraKulPong and R. Bowden. An improved adaptive background mixture model for real-time tracking with shadow detection. In Paolo Remagnino, GraemeA. Jones, Nikos Paragios, and CarloS. Regazzoni, editors, *Video-Based Surveillance Systems*, pages 135–144. Springer US, 2002.

Jinman Kang, I. Cohen, and G. Medioni. Object reacquisition using invariant appearance model. In *Proc. of International Conference on Pattern Recognition (ICPR)*, volume 4, pages 759–762, Aug 2004.

J. R. Kim and Y. S. Moon. A robust wavelet-based digital watermarking using level-adaptive thresholding. In *Proc. IEEE ICIP*, volume 2, pages 226–230, 1999.

Seung-Jin Kim, Suk-Hwan Lee, Kwang-Seok Moon, Woo-Hyun Cho, In-Taek Lim, Ki-Ryong Kwon, and Kuhn-Il Lee. A new digital video watermarking using the dual watermark images and 3D DWT. In *TENCON 2004. 2004 IEEE Region 10 Conference*, pages 291–294. IEEE, 2004.

H. G. Koumaras. Subjective video quality assessment methods for multimedia applications. Technical Report ITU-R BT.500-11, Geneva, Switzerland, april 2008.

A. Koz and A. A. Alatan. Oblivious spatio-temporal watermarking of digital video by exploiting the human visual system. *IEEE Transactions on Circuits and Systems for Video Technology*, 18 (3):326–337, March 2008.

D. Kundur and D. Hatzinakos. Toward robust logo watermarking using multiresolution image fusion principles. *IEEE Trans. Multimedia*, 6(1):185–198, Feb. 2004.

O. Kwon and C. Lee. Objective method for assessment of video quality using wavelets. In *Proc. IEEE Int'l Symp. on Industrial Electronics (ISIE 2001).*, volume 1, pages 292–295, 2001.

K. Laws. *Textured image segmentation*. Phd thesis, Electrical Engineering, University of Southern California, 1980.

Kenneth L Levy and Reed R Stager. Digital watermarking applications, August 21 2012. US Patent App. 13/590,940.

W. Limprasert, A. Wallace, and G. Michaelson. Real-time people tracking in a camera network. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, 3(2):263–271, Jun 2013.

C. Y. Lin and S. F. Chang. Semifragile watermarking for authenticating JPEG visual content. In *Proc. SPIE Security, Steganography, and Watermarking of Multimedia Contents*, volume 3971, pages 140–151, 2000.

David G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.

S.G. Mallat. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 11(7):674–693, Jul 1989.

Slaven Marusic, David BH Tay, Guang Deng, and Marimuthu Palaniswami. A study of biorthogonal wavelets in digital watermarking. In *Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on*, volume 2, pages II–463. IEEE, 2003.

Ajmal S Mian, Mohammed Bennamoun, and Robyn Owens. An efficient multimodal 2d-3d hybrid approach to automatic face recognition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 29(11):1927–1943, 2007.

A. Moorhouse, A.N. Evans, G.A. Atkinson, J. Sun, and M.L. Smith. The nose on your face may not be so plain: Using the nose as a biometric. In *3rd IET International Conference on Crime Detection and Prevention (ICDP)*, pages 1–6, 2009.

E. Nezhadarya, Z. J. Wang, and R. K. Ward. Image quality monitoring using spread spectrum watermarking. In *Proc. IEEE ICIP*, pages 2233–2236, Nov 2009.

T. Ojala, M. Pietikainen, and T. Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 24(7):971–987, Jul 2002.

Nikos Paragios and Rachid Deriche. Geodesic active regions and level set methods for supervised texture segmentation. *International Journal of Computer Vision*, 46(3):223–247, 2002.

P.J. Phillips, P.J. Flynn, T. Scruggs, K.W. Bowyer, Jin Chang, K. Hoffman, J. Marques, Jaesik Min, and W. Worek. Overview of the face recognition grand challenge. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 947– 954, 2005.

Angela Piper, Reihaneh Safavi-Naini, and Alfred Mertins. Resolution and quality scalable spread spectrum image watermarking. In *Proceedings of the 7th workshop on Multimedia and security*, pages 79–90. ACM, 2005.

Arman Savran, Nese Alyüz, Hamdi Dibeklioğlu, Oya Çeliktutan, Berk Gökberk, Bülent Sankur, and Lale Akarun. Bosphorus database for 3D face analysis. In *Biometrics and Identity Management*, volume 5372, pages 47–56. Springer Berlin / Heidelberg, 2008.

W.J. Scheirer, A. de Rezende Rocha, A. Sapkota, and T.E. Boult. Toward open set recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(7):1757–1772, 2013.

Loren Arthur Schwarz, Artashes Mkhitaryan, Diana Mateus, and Nassir Navab. Human skeleton tracking from depth data using geodesic distances and optical flow. *Image and Vision Computing*, 30(3):217 – 226, 2012.

Maurício Pamplona Segundo, Luciano Silva, Olga Regina Pereira Bellon, and ChauãC Queirolo. Automatic face segmentation and facial landmark detection in range images. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 40(5):1319–1330, 2010.

D. Serby, E.K. Meier, and L. Van Gool. Probabilistic object tracking using multiple features. In *Proc. International Conference on Pattern Recognition*, volume 2, pages 184–187, Aug 2004.

J. Shi and C. Tomasi. Good features to track. In *Proc. IEEE CVPR*, pages 593–600, 1994.

Jamie Shotton, John Winn, Carsten Rother, and Antonio Criminisi. TextonBoost for image understanding: Multi-class object recognition and segmentation by jointly modeling texture, layout, and context. *International Journal of Computer Vision*, 81(1):2–23, 2009.

M. R. Soheili. Blind Wavelet Based Logo Watermarking Resisting to Cropping. In *Proc. 20th International Conference on Pattern Recognition*, pages 1449–1452, 2010.

T. Sttz, F. Autrusseau, and A. Uhl. Non-blind structure-preserving substitution watermarking of h.264/cavlc inter-frames. *IEEE Transactions on Multimedia*, 16(5):1337–1349, Aug 2014.

D. S. Taubman and M. W. Marcellin. *JPEG2000 Image Compression Fundamentals, Standards and Practice*. Springer, USA, 2002.

Jer-Min Tsai, I-Te Chen, Ying-Fong Huang, and Cheng-Che Lin. Watermarking technique for improved management of digital medical images. *Journal of Discrete Mathematical Sciences and Cryptography*, 18(6):785–799, 2015.

K. Uehira, K. Suzuki, and H. Ikeda. Does optoelectronic watermark technology migrate into business and industry in the near future? applications of optoelectronic watermarking technology to new business and industry systems utilizing flat-panel displays and smart devices. *IEEE Transactions on Industry Applications*, 52(1):511–520, Jan 2016.

Cor J. Veenman, Marcel J. T. Reinders, and Eric Backer. Resolving motion correspondence for densely moving points. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 23(1):54–72, Jan 2001.

Pankajakshan Vinod and PK Bora. Motion-compensated inter-frame collusion attack on video watermarking and a countermeasure. *IEE Proceedings-Information Security*, 153(2):61–73, 2006.

Ba-Ngu Vo and Wing-Kin Ma. The Gaussian mixture probability hypothesis density filter. *IEEE Transactions on Signal Processing*, 54(11):4091–4104, Nov 2006.

S. Wang, D. Zheng, J. Zhao, W. J. Tam, and F. Speranza. An image quality evaluation method based on digital watermarking. *IEEE Transactions on Circuits and Systems for Video Technology*, 17 (1):98–105, Jan 2007.

Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Processing*, 13(4):600–612, April 2004.

A. B. Watson. Visual optimization of dct quantization matrices for individual images. In *Proc. American Institute of Aeronautics and Astronautics (AIAA) Computing in Aerospace*, volume 9, pages 286–291, 1993.

X. Xia, C. G. Boncelet, and G. R. Arce. Wavelet transform based watermark for digital images. *Optic Express*, 3(12):497–511, Dec. 1998.

L. Xie and G. R. Arce. Joint wavelet compression and authentication watermarking. In *Proc. IEEE ICIP*, volume 2, pages 427–431, 1998.

Takaaki Yamada, Michiro Maeta, and Fuminori Mizushima. Video watermark application for embedding recipient id in real-time-encoding vod server. *Journal of Real-Time Image Processing*, 11(1):211–222, 2016.

Hanxuan Yang, Ling Shao, Feng Zheng, Liang Wang, and Zhan Song. Recent advances and trends in visual tracking: A review. *Neurocomputing*, 74(18):3823–3831, Nov 2011.

A. Yilmaz, Xin Li, and M. Shah. Contour-based object tracking with occlusion handling in video acquired using mobile cameras. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 26 (11):1531–1536, Nov 2004.

Alper Yilmaz, Omar Javed, and Mubarak Shah. Object tracking: A survey. *ACM Computer Survey*, 38(4):13–es, Dec 2006.

Stefanos Zafeiriou, Mark Hansen, Gary Atkinson, Vasileios Argyriou, Maria Petrou, Melvyn Smith, and Lyndon Smith. The Photoface database. In *IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 132–139, 2011.

H. Zhang, H. Shu, G. Coatrieux, J. Zhu, Q. M. J. Wu, Y. Zhang, H. Zhu, and L. Luo. Affine legendre moment invariants for image watermarking robust to geometric distortions. *IEEE Transactions on Image Processing*, 20(8):2189–2199, Aug 2011.

Z. Zhang and Y. L. Mo. Embedding strategy of image watermarking in wavelet transform domain. In *Proc. SPIE Image Compression and Encryption Tech.*, volume 4551-1, pages 127–131, 2001.

X. Zhu, J. Ding, H. Dong, K. Hu, and X. Zhang. Normalized correlation-based quantization modulation for robust watermarking. *IEEE Transactions on Multimedia*, 16(7):1888–1904, Nov 2014.

Zoran Zivkovic and Ferdinand van der Heijden. Efficient adaptive density estimation per image pixel for the task of background subtraction. *Pattern Recognition Letters*, 27(7):773 – 780, 2006.