Journal of Theoretical and Applied Information Technology
15th September 2018. Vol.96. No 17

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

# A REVIEW AND OPEN ISSUES OF DIVERSE TEXT WATERMARKING TECHNIQUES IN SPATIAL DOMAIN

**[1]ALI A. ALWAN, [1] SHAHIDAN M.A,[1] NILAM NUR AMIR SJARIF, [2] *MOHAMMED MAHDI HASHIM,[2]MOHD SHAFRY MOHD RAHIM***

[1] Faculty of Advanced Informatics School, University Technology Malaysia, Kuala Lumpur, Malaysia
[2] Faculty of Computing, University Technology Malaysia, Johor Bahru, Malaysia

E-mail: aliraheem1983@gmail.com, mshahidan@utm.my, nilamnur@utm.my, comp.mmh@gmail.com, shafry@utm.my

**ABSTRACT**

Nowadays, information hiding is becoming a helpful technique and fetches more attention due to the fast growth of using the internet; it is applied for sending secret information by using different techniques. Watermarking is one of major important technique in information hiding. Watermarking is of hiding secret data into a carrier media to provide the privacy and integrity of information so that no one can recognize and detect it's accepted the sender and receiver. In watermarking, many various carrier formats can be used such as an image, video, audio, and text. The text is most popular used as a carrier files due to its frequency on the internet. There are many techniques variables for the text watermarking; each one has its own robust and susceptible points. In this study, we conducted a review of text watermarking in the spatial domain to explore the term text watermarking by reviewing, collecting, synthesizing and analyze the challenges of different studies which related to this area published from 2013 to 2018. The aims of this paper are to provide an overview of text watermarking and comparison between approved studies as discussed according to the Arabic text characters, payload capacity, Imperceptibility, authentication, and embedding technique to open important research issues in the future work to obtain a robust method.
**Keywords:** *Information Hiding, Text watermarking, Least Significant Bit (LSB), Different types of watermarking, Spatial Domain***.**

## 1. INTRODUCTION

In these days world, with a fast boom of technology, picture processing is being used in numerous fields and is gambling a very important position. Picture processing has been extensively studied on classical computer systems starting, from image got it to image analysis dealing with improvement, segmentation, transformations, and security [1]. Information hiding mechanisms can be applied to prevent the secret data from the intruder or the malicious modification. Information Security is a procedure of obtaining information secure with more integrity and confidentiality. There are two of information hiding techniques that developed to protect the information and which named Steganography and Watermarking. Both of steganography and watermarking are related to the common concept [2]. Steganography can be defined as an intelligent technique to hide secret data in hosting media, in a manner of complete silence and deceptive which make hosting media carry the secret data, so the presence of the hidden secret data is indiscernible by unauthorized access or the intruder. Stego analysis is detecting the secret data by analysis stage media [3]. Digital Watermark is a digital signal or sample inserted right into digital information, which also can be named as a digital signature. It is a far key procedure in the protection of copyright possession of digital statistics, including text, image, videos, audio, etc., [4]. Digital watermarking is one of the maximum essential strategies uses information hiding so that you can defend any kind of media from the feasible types of attacks. Digital watermarking is described in [in [1] as "a procedure that embeds or inserts more facts, named the watermark or mark, into the unique records to generate the output that is referred to as a watermarked or marked information". The watermarking device goes via 3 levels [tiers [2]: firstly]: first of all, the procedure of era and embedding the watermark within the original media. Then, viable assaults should arise inside the broadcast of the signal via the watermark

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

channel. Eventually, A method of detecting the embedded watermark [4]. Consequently, the demand for an effective technique to keep text privateers and safety is an excessive call for. Digital watermarking is used to clear up those issues. It embeds a signal known as a watermark into digital facts (audio, photos, video, and textual content) without destroying the information value to get the watermarked statistics. The watermark later detected or extracted for use in many programs consisting of: copyright safety, information authentication, records hiding and covert verbal exchange [1]. The watermarking gadget entails main procedures: embedding of the watermark into the authentic statistics and extracting the watermark from watermarked facts or attacked watermarked facts. It is able to be described via some of requirements number of requirements can describe it [1, 2]: robustness, ability, imperceptibility, and authentication. Special domain plays an important role in all the applications of the text watermarking techniques in latest papers so that the relevant literature that have been selected in this study emphasizes on this significant subject only by analyzing the main characteristics and the drawbacks also for the text watermarking techniques. Digital watermarking is supported many applications such as online transactions, Arabic language, Holly Quran etc. [3].

There are several reviews on digital watermarking have been published within recent years, the most important and popular one was published two years ago [6]. it focuses on a review of essential concepts, variety of evaluation measures, security or authentic side of text watermarking scheme and includes the literature that has been published until the time the paper was published. Nevertheless, this review may be considered out of date due there are many of contributions published from that date, these new publications necessarily to be collected within a new review paper. Briefly discussed some other surveys the text watermarking definition, domains as well as techniques in a summarized form without discussing the huge amount of contributions on this area [7] [8], with respect to the review papers. While the difference in our work summarizes the current text watermarking techniques in the spatial domain, also analyzed different problems and the drawbacks of each method that have been innovated from last few years. Text watermarking is the most important and challenging media in the digital watermark is text. This media got more attention for great reason that is the writing and reading using it and the early man aim to develop the text and face the difficulty due to variations among the society and countries. Text media get big history with beginning since the human race [5]. Those necessities utilized in watermark embedding and extraction methods to be carried out in an effective manner. The Arabic language includes many capabilities which are exploited inside the area of statistics hiding inside its branches: steganography and watermarking. Some of the Arabic scripts traits are the existence of factors above or under the letter, diacritics which are equivalent to the vowels in the English language and kashida that's the stretching person [6]. Vowel letters in Arabic can be defined as animated sounds help determine word pronunciation, writing them is the same of normal but different in pronunciation. Choosing of these letters supposed to help of finding the positions to embedding the secret message. Kashida mentioned above used to describe whether the inserting one or zero, which means the absence of kashida in the word means zero otherwise, means one. Kashida (or extension of the letter) do not change the meaning of the word in Arabic language, so present or absence it not effects on the meaning just informative style [7].

The essential criteria that have been adopted in this review for a comparison between approved studies are discussed according to the selected pixel in Arabic text characters, capacity and embedding algorithm. The pixel selection in Arabic text characters is used to achieve the objective of security or authentic such as kashida and Diacritics based technique [9], Arabic text watermarking [10]. while the second criteria refer to the maximum amount of secret massage which can be embedded into the text without reiteration of text quality. Finally, embedding algorithm is used to achieve the objective of text quality (Imperceptibility) which is responsible for keeping the quality of the text same of the original. This is achieved by keeping the bit value same as original as possible. This work will describe carefully only the spatial domain techniques, transform domain does not include. This study achieves a less about Arabic textual content watermarking given that a few researchers are accomplished in this region. A rich of Arabic text assets need to guard its copyrights and hit upon any adjustments in particular in touchy textual content such as the Holly Quran. The researcher required need to make constructed Arabic text watermarking and English text watermarking by using the vowel letters.

This paper intends to discover different textual content watermark strategies in a spatial area for the duration of tested and analyzed to the modern-day

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

methods and identify the challenges and the open research troubles that is thrilling in this region. The structure of the paper as follows: In section 2, Overview of watermarking and the stages of watermarking in section 3. Requirements of digital watermarking in section 4, Types of watermarking in section 5, Techniques of text watermarking in section 6, 7. Other algorithms techniques in section7, 8. Quran text structure in section 8. Section 9 demonstrates the challenges and open research issues. Finally, the conclusion of the paper has been summarized in section 10.

## 2. OVERVIEW OF WATERMARKING

The recent development in the innovative communication, computerized recording, and the storage devices have created the inherent internet situations with the capacity to convey, copy, acquire, and recreate advanced media denied of any quality lossless. In any case, this advancement is an interactive media content distributing enterprises because of unlawful get right of passage to by means of computerized media content material that is requiring a basic and moment wellbeing for licensed innovation rights [2]. Despite the fact that a conventional cryptography approach can take after to keep an unapproved user to get to a virtual media content by utilizing the material contents media encryption process. The generation of the cryptography has a couple of boundaries to completely protect the egghead ownership Rights. In this way, it is clear that distinctive innovation must be achieved in order to gather the scholarly things Rights, track the virtual media substance, and offer the advanced media content confirmation which guarantees lawful got passage to and avert illicit control to the substance thing of digital media. As a dynamic solution, advanced Watermarking is applying to encourage the prerequisites for protecting the egghead Rights of the digital media. Recently, digital watermarking is turning into the current field for the examination bunches that compelling working in this industry. The basic principles of virtual watermarking are the concealing measurements about virtual content as a metadata inside the virtual satisfied material to secure the proprietor. Basically, this idea is created in Germany because of a German expression "waterssermark" is to be had which implies the effect of water on paper [1].

Image logos, text messages, and raw water slides can be hidden extra information into host content such as images, audio signals, speech signals, 3D graphics objects, video clips, text and software codes, network flows, XML data, and ontology via utilizing watermarking techniques without any declinations. The watermark must be noticeable from watermarked content notwithstanding when purposeful and inadvertent control is performed on watermarked content [9]. Consequently, the digital watermark can be communicated through two principle forms specified by embedding and extricating the watermark which is illustrated in Figure 1.1.

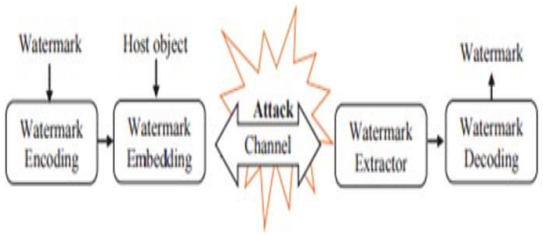


*Figure 1. The concept of digital watermarking.*

## 3. THREE STAGES OF WATERMARKING

**Generation and Embedding:** Pseudo-Random Sequence, M- Sequence and Chaotic Sequence are some sequences employed for the generation of watermark [13]. The combination of watermark signal and the original image can be seen as embedding process.

**Distribution and Possible Attacks.** The distribution process can be understood as the transmission of the signal through the watermark channel. Possible attacks in the broadcast channel might be accidental or Intentional [13].

**Detection:** Detection process allows the owner to be identified and provides information to the intended recipient [14]. There are two types of detection: Blind detection and informed detection.

## 4. REQUIREMENTS FOR DIGITAL WATERMARKING

Digital Watermarking might consider being an interchange of inconvenience on account of the character of virtual watermarking that is contending among payload, transparency, robustness, and security. Depending on different and wide programs of the digital watermarking, the process is planned construction absolutely with respect to those prerequisites and framework innate houses. Figure 1.1 shows the necessities of digital watermarking. As a result, those preconditions

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

contradict each other and to conduct them to meet is troublesome. Inside this enchantment triangle, "the mode of operation" of a particular watermarking machine might be attacked. For instance, immoderate information charging device can also improve its process by decreasing the robustness of the identical processing [11]. These requirements may be different based on two things: type of data that is a watermark and the applications [9]. This disparity between the spatial domain and transform domain is differential inverse the four criteria which are Robustness, Imperceptibility, Capacity, and Authenticity can be described as follows:

### 4.1 ROBUSTNESS OF DIGITAL WATERMARKING

Digital watermarks tend to be strong. Since the durability of a single cost is incorrect, this is due to the fact that each aquatic method is weak against the minimum attack. Moreover, the design of a watermarking approach that is strong for all manipulation methods is unnecessary and unavoidable. For example, in the radio and television broadcast monitoring tool, the watermark is strongly wanted to resist transmission change, which includes A / D and D / A conversions, vertical pressure and horizontal translations, but there may not be a need for a watermark to be strong against the wide adjustments with sizing, rotation, filtering, and distortion may not take the site during broadcast [13]. For instance, authentication handiest calls for fragile watermarking to test whether or not the virtual object has been modified. To finish, durability only requires unexpected packages consisting of proof of ownership, control of cloning, identification, and unavoidable blurring that can be controlled. There may be a theme about eliminating the watermark [14].

### 4.2 CAPACITY OF DIGITAL WATERMARKING

The weight of the hidden bit according to its digital term is specified as the capability or payload and mostly is computed in bit per second (bps). In most cases, the bit of a watermark is naturally regarding the messages ranges which included in the encryption rules set. clearly, range without any delay is based on the size of host records. Meanwhile, in each instance, the host style amount increases, most bits of the watermark are involved in the host information [12].

### 4.3 IMPERCEPTIBILITY OF DIGITAL WATERMARKING

Imperceptibility is characterized as the measure of distortion which is infused by inserting the watermark. It might be explicit through estimating the pleasant or fidelity. Regarding the rating of the validity that evaluates the similarity between the real tools and the watermark, autonomous acceptance process is achieved on a watermark position in order to obtain the high quality of a watermarking technique. Also, there are sincerity and perfect measurements checking via self-measurement and objectivity. The pivotal accurate and reliable implementation of both stability and insufficiency is the self-measurements as well as the physical appearance. Self-testing is led essentially based on a logical field, deciding the connection between the physical world and the self-feeling of individuals in that world [15].

#### 4.3.1 OBJECTIVE MEASUREMENT

The managed spanned words within the implementation goals, the basic notation is extracting the size of the watermark distortion that might be measured with the help of the similarity or deference comparison between the original and watermarked objects [3]. A basic and typical strategy known as peak signal-to-noise ratio (PSNR) can consider being efficient in terms of object size through trying to recognize the distortion as a noise. Distinctive destinations tests for a trademark the digital content related to the cases of audio and speech watermarking. The essential estimation incorporates spectral distortion (SD), mean square error (MSE).

#### 4.3.2 INTELLIGIBILITY MEASUREMENT

Intelligibility technique how conceivable the digital object is. The fundamental clearness is the cost of measurements and the substance of digital media. For instance, the kid talking voice experiences the loss of verbal realities that have no undeniable causes by any means, despite the fact that the excellent sound is great. The major part of packets likes speech, text and natural language, Intelligibility criteria are considered to be an applicable and effective parameter as compared with the imperceptibility [16].

### 4.4. AUTHENTICITY OF DIGITAL WATERMARKING

Journal of Theoretical and Applied Information Technology
15th September 2018. Vol.96. No 17

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Image fragmentation removes a concise version of the image to symbolize its contents used in image authentication. If the image is modified maliciously, the defragmentation must be greatly changed [17]. Unlike the defragmentation properties of encryption such as MD5 and SHA-1 which can be very sensitive to light modifications in input entries, image fragmentation must be solid toward normal image processing. The image must be short and strong to manipulate images every day and be sensitive to manipulation.

## 5. TYPES OF WATERMARKING

There are different techniques can be used and separated into different classes such as Watermarks and watermarking techniques. Watermarking strategies can be partitioned into their classes as indicated by the sort of archive to be watermarked [18].

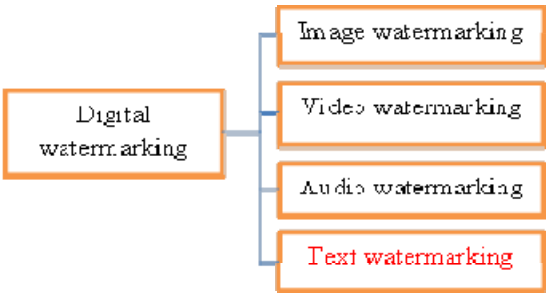


*Figure 2. Types of watermarking.*

Image watermarking: Digital images are mostly used as the cover object for watermarking, as well as that is also more popular over the internet. In this technique, a secret message is embedded in a digital image during an algorithm with the help of secret key to create a stego image. Generally, in this technique pixel intensities are used to hide the secret information [19].

Video watermarking: Hiding secret information in a video format is known as video watermarking. Video files are a consisting collection of images as well as audio. Generally, most of the proposed techniques on images and audio can be implemented to video files too. The use of video watermarking is more eligible instead of the other multimedia files, because of a large amount of information that can be hidden inside video format without noticeable by humans because of the continuous flow of information. Many types of video files can use such as H.264, Mp4, MPEG, AVI or other video formats [20]. Text media get big history with beginning since the human race, in this proposal; we will study this media in detail to be clear [7]. Any visual change in the text strongly change the meaning of the word or sentence, text some time become very short (e.g. social media posts) or long text such as books or Quran. Recently, a race of using internet application like social media and mobile communication allows for much sharing of knowledge need to be secure. Many illicit actions considered in this regard such tampering, illegal copying, and forgery.

## 5. TECHNIQUES OF TEXT WATERMARKING

On the bases of Text watermarking techniques is classified in the Spatial domain. The classification is explained according to the scope of work in Fig. 3. This can be classified as follow:

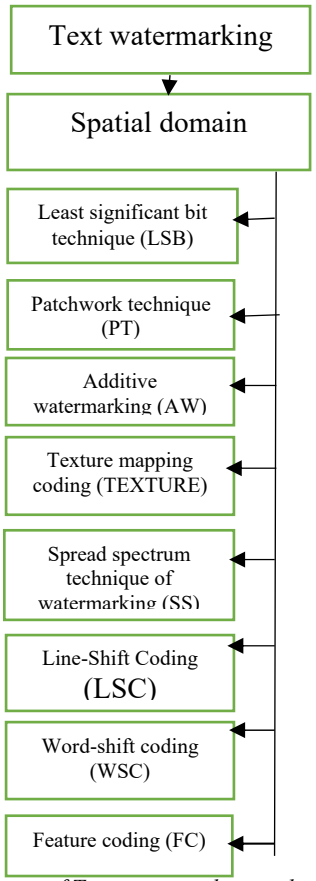


*Figure 3. Classification of Text watermarking techniques*

### 5.1 SPATIAL DOMAIN

Spatial domain focuses mainly on the adjustment of pixel values for one or two randomly

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

selected subgroups of images. This algorithm carries raw data or merges it directly into image pixels. Some spatial domain technology algorithms are LSB, Patchwork, Text Encoding, spread spectrum technique, Line shift coding, word shift coding and feature coding [21].

**A) LEAST SIGNIFICANT BIT TECHNIQUE (LSB):** LSB technology is one of the simplest strategies to implement. During this method, the watermark bit is entered to a lower full size per pixel bit. The best bits of each pixel are examined to detect watermark information at some point in the extraction or detection technique. On this approach, we assume that the image of the watermark is catted, and the recipient is nevertheless able to obtain the mandatory data since the records are an embedded number of times. This technique is very sensitive to noise and cannot be used for practical purposes. Also, it is not very robust [22]. According to presents a watermarking technique which Least Significant Bits (LSB), its steps and its process with Mat lab images. LSB is applied to watermark for the security of the image. But it is assumed that LSB is not a reliable technique of image watermarking as it works on spatial domain and one can easily identify the secret data in the LSB based watermarked image [23]. Multiple text images have been hidden into a single colored image using modified LSB substitution method. Total 6 text image had been used for hiding purpose. The proposed method has the ability to embed/hide multiple text images in any color image. The encoded text images able to flawlessly recoverable. The key benefit of this coding algorithm is, the embedded text image does not make any noticeable marks on the color image. The drawback of the text data cannot be identified by naked eyes. In this study provides more security because embedded watermark can only identify by reasoning the selected Color Component as each color space has different images, as to increase the capacity and efficiency of the proposed method. [24].

**B) PATCHWORK TECHNIQUE (PT):** Patchwork is a statistical technique that has been developed. In this technique, water stains were introduced based on a census found using the Gaussian distribution. This technique works as follows. Two random patches are selected for migration design [25]. This technique uses a repeating pattern to embed data within an image. Another meaning, the distribution of the image is done in two subsets. One process is selected and applied in a reverse way to these subsets. For example, if one of the subsets is augmented by a factor k, by equivalent quantity the other subset will be reduced [26]. A new idea to study the translations of the Qur'an from the cohesive entity and the lexical coherence point of view, as a means of assessing the accuracy and equivalence of existing translations. Further, this an initial stage of the investigation of the issues, the building of the platform and the identification of some preliminary rules to compare and evaluate the structure of speech translations. This method cannot be used for other types of Quran translation as a general solution because there is no explicit application for each translation. An evaluation factor for existing translations in terms of order of sentences, phrases and different words, which affects the results of the analysis of the text [27].

The researcher mentioned that digital Quran resources for all multimedia types require information security and mechanisms for protection using integrity verification and authentication. Digital watermarking techniques present a very promising solution for ensuring integrity and authenticity of the Quran content. This address security issues concerning for digital online Quran propagation [28]. According to a unique watermarking for texture content image safety and supply authentication. The proposed technique was textually implemented on virtual Quran content photos, the feature of its sensitive nature in its minimal use of portraits/shades and use of best textual content symbols which impose in additional constraints on the seek area for an imperceptible embedding. They had been powerful for giving users a higher self-belief of the authenticity content image safety and real possession of virtual media. Furthermore, the proposed method is the ability for software in other sensitive virtual textual content pictures or graphical content in addition to much

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

less-sensitive content material additionally. Consequently, be used to defend such digital contented from tampering, forgery and illegal content material manipulations [29].

**C) ADDITIVE WATERMARKING (AW):** Spatial domain technique is one of the simplest and most simple ways to embed a watermark. In this, a semi-random noise pattern is added to the pixel intensity of the image. The noise signal is usually a floating point or integer such as -1, 0, 1. In this noise is generated by a key that ensures the possibility of detecting a watermark [30]. Expected a method to insert multiple watermarks into the spatial domain and frequency to make the assets safe and divide the host image into two regions, for example: region A and region B. In Area A, owner information is entered with the help of LSB technology and in Area B; the watermark is introduced with the DCT-DFT additive technology, [31].

In addition, such method is suitable only to localize the image forgery on some of the most significant bits. On the other hand, the block-based concept has issues for the parameter of the block sizes and watermark payloads. Some experiments are required to determine the proper parameter that facilitates acceptable tamper detection while maintaining the image quality. However, the main drawback in block-based concept is being unable to locate the tampered pixels accurately; this might be important for specific applications such as in the military communication [32].

**D) TEXT REMAPPING CODING (TEXTURE):** Texture mapping coding is the technique in which the watermark is hidden in the texture portion of the image. This method is useful only for those images that contain a text pane. Data is hidden within the continuous pattern of the image. Thus, the method is suitable only for those images with variable strength. [33]. This method is suitable only for areas that contain a large number of random image images and cannot be performed automatically. This method hides data within the continuous random patterns of the image [34]. A new innovative method for watermark to facilitate the process of documentation and detection of fraudulent images in the pictures of the Quran. Double layers of implant scheme are provided on wavelets and field to improve the sensitivity of fragile watermarks and protect from attacks. So, a chaotic map is used to obliterate the watermark to make it safe against local attack. The proposed method allows for high loadings of the watermark while maintaining good image quality. The results of the experiment confirm that the proposed methods are fragile and have a high ability to detect tampering, although the area of tampering is very small [35]. According to a method of watermarking textual messages is provided based on a given area of the image; the images used in JPEG format. The overall structure of the proposed algorithm is set in both encryption and decryption. The Watermarking of text messages in bytes of few values is without causing confusion and increasing the size of the image pixels in the image and at a higher rate, respectively. The watermarking in the text because textual data usage is higher than other text due to the low costs of operation, print and the need to less memory has some advantages, In addition, A disadvantage of this method is a loss of JPEG compression method [36].

The developed an embedding algorithm, which uses the unique structure of the "Saravanan" language, has been developed to generate the required watermarks in combination with additional security phrases. The watermark gave an additional level of security by encrypting it with the AES algorithm. The extraction algorithm uses the Saravanan property as a private key to reveal the ownership of the Devanagari text document. The authors also performed performance analysis with a suitable set to analyze potential attack types [37].

**E). SPREAD SPECTRUM TECHNIQUE OF WATERMARKING (SS):** In spread-spectrum watermarking technique is spread over several frequency boxes so that the energy in any box is too small and thus undetectable. Spreading the watermark across the image spectrum ensures a great deal of safety against unintentional or intentional attack as the location of the watermark is not clear and the frequency zone is determined in such a way as to ensure sufficient energy in any single coefficient [38]. It would be practically

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

impossible to detect a well-positioned watermark in the image frequency field. Another author used A fragile watermark is achieved in a helical way which indicates a good numbering scheme for inclusion. The watermark scheme is inconceivable and can contain high capacity for data with good authentication. The average signal-to-noise signal (PSNR) of the combined image is 67.06dB and the average operating time is 1.29 seconds. Then localize and restore the watermark system to be applied in the application of the Holy Quran. The result is hg with up to 100% recovery. The purpose is to verify the validity and validity of Quranic verse images that are usually exchanged among believers in the Internet, specifically in social networking sites and social networking applications. Moreover, the enemies of Islam can attack common images to carry the misleading meaning of the verse [39].

The process of hiding a small text or owner information in color images has been presented in [40]. In this method, the embedding process relies on the insertion of the elements of the text randomly into the color image as noise, where the randomness is achieved using random number producer that is decided to be influenced through the image size as well as the text watermark length. Initially, three color image processing methods; separate color channels, direct conversion to grey and YIQ model, were studied and compared with each other for processing speed first, and the YIQ model was faster than the other models, and henceforth it is adopted for the embedding procedure. the proposed algorithm for embedding and extraction of text watermarks of various content combinations and sizes into vast number of images were satisfactory, as they resulted into an improvement of additional than dual in speed of embedding and extraction than other schemes (such as Least Significant Bit technique, LSB) and fairly satisfactory level of peak signal to noise ratio (PSNR) with low square error values [82].

Other author suggested a novel digital text watermarking algorithm is developed based on Unicode extended characters. The advanced algorithm provides a high-precision technique that is not familiar with watermarking, which is about 99.9% of the similarity factor while maintaining the cognitive quality of the watermarked documents to have a PSNR value higher than 63. The hardness assessment proves that the proposed algorithm tolerates most of the potential attacks and is capable of accurately extracting the watermark High. The power factor rating shows that the proposed algorithm has a distinct watermarking capacity with a load capacity of about 2 bits/word. The main advantage of this method is to reverse the watermark text in its original state before extraction. In addition, abnormalities of the unnatural text with extra space will affect imperceptibly [42].

**F). Line-Shift Coding (LSC)**: Here, each line of the level was shifted slightly or in response to the bit-load value [43]. The activities, objective performance metrics required to consider whether one of the recognized is developing watermarking technique.


*Figure.4. The watermarking technique*

In line shifting method [44], the lines are moved vertically upward or downward to hide information. It can be applied to images without the need for the original image because of the regular spaces between the lines in the image. Lines displacement process is the most noticeable by the reader, but it has a good strength and resistance to noise due to

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

the length of the lines, so it is more suitable for use with images. It has low capacity where it can hold one bit per line. Figure 5 shows an example of this method.

| وظيفة الأداة واستخدامها | اسم الأداة | الاسم الشائع | شكل الأداة |
|---|---|---|---|
| عرض الصور أو وعاء لباقي الأدوات أو لوحة رسم | خانة الصورة | PictureBox | |
| عرض النصوص الثابتة التي لا يستطيع المستخدم تعديلها | أداة العنوان | Label | A |
| طلب المعلومات من المستخدم أو عرض المعلومات | خانة النص | TextBox | ab| |
| وعاء لباقي الأدوات | الإطار | Frame | |
| ينقر عليه المستخدم لتنفيذ أمر معين | زر الأوامر | CommandButton | |
| اختيار مجموعة من مجموعة (اختيار الألوان المفضلة مثلاً) | خانة التحقق | CheckBox | |
| عرض مجموعة خيارات لاختيار أحدها (اختيار اللغة مثلاً) | زر الخيار | OptionButton | |

*Figure 5. Example of line-shift coding*

The digital watermark is used to determine the copyright ownership and authentication process, through a new method of watermark text is presented using each space between words with pseudo-space. If the character before space is pointed and the watermark mark is one, pseudo space is entered; otherwise, no false space is entered. If the character before space is not specified and the watermark mark is zero, false space is entered in another way, no false space is entered. The results show that proposed method has a higher capacity ratio and imperceptibility than other watermarking techniques, but the used kashida might be easily detected because not secure [45].

Multiple watermark algorithms for mixed Chinese and English texts based on character encoding and attributes have been proposed in [46]. This algorithm can provide greater watermark capacity, and at the same time, keep information with an invisible watermark. It also provides a strong security capability to ensure that watermarks are not destroyed or tampered with by other malicious attackers. In addition, watermark capacity can reach 200%; the special conditions demonstrate a significant watermark capacity for this algorithm. When the text is under multiple attacks, this algorithm has greater ability to extract watermarks and correct errors. It can also be practical for mixed Chinese and English texts. Moreover, the rigidity and safety are not as good as putting the watermark later on the basis of shape. [47]. Recommended a new technique for copyright protection of digital texts via watermarking of encryption note in digital flies provided. The method is based on the placement of hidden character before and after special writing characters and blank lines between paragraphs in the main text with subtle changes and high resistance, the drawback of this method it has been tried to overcome the watermarking resistance in the watermarked text against the attacks of writing changes.

An effective watermarking method for text documents in Hindi has been suggested in [48]. Hindi stands second among all languages around the world. It has a wide availability of its digital contents of different types. The watermark is logically included in the text using "swar" as a special feature of the Hindi language, supported by appropriate encryption. The results are particularly limited and apply to the Hindi language based on the properties of language construction. Successful implementation of the technique may lead to the discovery of a general approach to the watermark of the text. However, this method involves low capacity with high security.

**G). WORD-SHIFT CODING (WSC):** Dividing each line into a group of words. Each group has enough characters. Then, after allowing the bit value of the load, each pair is converted to the right or left. Individual groups are treated as benchmarking [49].

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

| | |
|---|---|
| Watermarking bits | 10010110010110 |
| Original text | البر حسن الخلق والإثم ما حك في نفسك |
| Output text | البر حسن الخلق والإثم ما حك في نفسك<br>0110 1001 10 10 01 |

*Figure 6. Example of word-shift coding*

The words are moved horizontally making spaces expanding between words to hide information. This method needs the presence of the original file or the original image due to the possibility that the file contains a variable number of spaces between adjacent words. It is better and more effective for use with files because they are less attractive to the reader's attention. Its capacity is about one bit per word [50] Figure 6 shows an example of this method. The author suggested that a new steganography method for Arabic text. The method hides secret information bits within Arabic letters by two features, which are the moon and sun letters and the redundant Arabic extension character "-" known as Kashida. The proposed work hides the message in Arabic text using Arabic language features. In Arabic, there are two groups each set of 14 characters, namely the letter sun (solar letters) and the moon letter (moon letters). The secret text is hidden in zeros and 16-bit Unicode characters per character (UTF-8 encoding uses 16 bits to represent one Arabic character). The lower the capacity of one method, the larger the carrier file must be to hide the secret note. The capacity is weighted on the basis of the double aspects that have been the inclusion ratio (ER) and the efficiency ratio (TER). The advantage of applying the concept of Moon and Sun is that it is able to increase the probability of hiding secret bits in any character. However, it is also important to maintain the cognitive aspect while improving the ability. The main drawback is the security of a larger method that must be the carrier file to hide the confidential message [51]. An improved coding method - Kashida with repetitive character characteristics for Arabic text documents, suggested techniques to explore all possible situations where Kashida can be placed before specific Arabic characters that always appear in Arabic scripts and are always connected to other characters from the right. In order to achieve our goal of protecting the copyright of documents and verifying authentication while minimizing vulnerability to security attacks in a document with a watermark. In addition, the two approaches (Method -A and Method-B) depend on the frequency properties of the character frequency. Experimental results show that the lower the capacitance, the greater the susceptibility to perception, in other words, cognitive inversely proportional to amplitude and without validation [52]. [53]. Suggested improving the ability of the Kashida propagation system using white spaces between words in Arabic to hide. The idea combines two ways to gain power by adding white spaces to Kashida as extra redundant pieces to hide. Kashida should be included whenever possible after any Arabic letter, regardless of whether it is dotted or not. The main feature of this method is found in its high capacity since each Arabic letter applies to the diacritic mark and can be concealed in secret. Moreover, the drawback with security is deteriorating because there is a high probability of raising doubts about eavesdropping on the existence of secrets as there is no realization in this work.

A new scheme for word document, in which hiding data in the special attributes of word document objects and performance excellent on the bases of robustness, capacity, and imperceptibility is proposed. They used three properties bookmark, string, and variable for data hiding and also created a layout of the word document for copyright protection through manipulating the word script. The watermarking scheme is excellent on imperceptibility and robustness with large capacity. After applying malicious kinds of attacks, insertion, deletion, replacement and formatting attack, watermarking information still be extracted [54].

In this study, hiding of information is done by using 8 various diacritical characters in Arabic to hide binary bits (0 or 1) in the original cover media. The diacritics are shown or deleted according to its given bits. in this work, since fathah (/) is the most used diacritics in many texts, the fathah diacritic is being utilized to hide bit 1 or 0, This technique is

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

robust but not imperceptible and it is not appropriate for sensitive texts [55]. There are two methods for Arabic text watermarking is proposed by utilizing each word space. Method 1 is specialized to be used in the Arabic text or similar languages as it utilizes the Arabic dotting feature. Method 2 could be used for any language uses spaces to separate their words. The experiment results for method 2 has the largest capacity, but slightly lower imperceptibility than Method 1. The proposed methods are robust against electronic text attacks such as copying and pasting, text formatting and text tampering for tampering ratio up to 84% [56]. Two Arabic text watermarking methods are proposed in this paper by using each word space in the text and this guarantee high insertion capacity. The manipulation of white spaces between words is done by adding small spaces or no width ones instead of using normal space to satisfy imperceptibility requirement. The first method, Method 1 utilizes the dotting feature in a way to improve the capacity of the method presented [57]. According to an invisible technique of watermarking Kashida based for an especially Arabic text document, it is a possible position where Kashida could be put before specific Arabic characters which constantly appear in Arabic scripts and are always related to other letters from the right. The pertinence of this method that includes copyright protection, document-authenticity verification, and document tamper-proofing. Therefore, the decreased the capacity the higher the imperceptibility [58]. An enhance the functions of Kashida within the science of concealment known as "writing texts of Arabic text by extension" to the character of Kashida and comparison. This idea combines two technologies to gain power by adding white spaces to Kashida as extra redundant cover. A feature that has proven to make it more visible in the capacity as expected without security degradation, making it advantageous to help users with replacement information through text documents and making secure connections. The disadvantage of these ideas using the Kashida character is that it cannot be added at the beginning or end of words. It can only be added to joined characters in words [59].

**H). FEATURE CODING (FC):** This approach modifies features a little bit to embed a watermark. [60]. The length of end lines of letters, b, d, h, etc. is altered by extending or shortening it. Line and word shifting could be applied in any language even Arabic, but the feature coding is done on a language based on its features. All of these methods could be applied to an arrangement file or to a binary image of a text. We would discuss some of the studies based on feature coding in the Arabic language. The authors proposed an approach depends on the characteristic of existing of points on the majority of Arabic, Urdu and. Persian letters. These points are used to hide secret binary information. If the secret bit equals to one, the points within the pointed letter are shifted a little upward. If the secret bit equals to zero, the points' location does not change. This method has height capacity where the Arabic language has 15 pointy letters from its all letters (32). It is also used with extension charitable better results [61]. The recovery process is acceptable even in the case of printed text. However, the retyping process would remove all the secret hidden bits. Also, this method requires creating a special font. Figure 7 shows Perpendicular can movement of the points for the letter NOON.

*Figure 7. Perpendicular movement of the points for the letter NOON*

According to use a capacity and security issues of text steganography by employing LZW compression technique and color coding-based approach. Moreover, uses the forward mail platform to hide the secret data. This procedure first compresses secret and then hides behind the compressed secret information into the email addresses and also in the cover memo of the email. The secret data bits are embedded in the message (or cover text) by making it colored by a color coding table. The planned technique not only produces a high embedding capacity but also decreases computational difficulty. Moreover, the security of the proposed method is significantly improved by employing stage keys and there is no

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

imperceptibility [62]. The proposed technique for English language using natural language watermarks is enhanced by making the technique more robust against tampering attacks. Moreover, a watermark key can be created using a combination of the count of nouns, pronouns, modal verbs and conjunctions along with author name. Therefore, encryption methods are used to enhance the security level which gives better results [63].

## 6. OTHER ALGORITHMS TECHNIQUES

There are some other algorithms techniques that are following below: -

### 6.1 XOR WATERMARKING TECHNIQUES

The XOR watermarking algorithm was proposed to be used to eliminate LSB watermarking defects in watermarking [64]. The cover is processed with watermarking and embedded data using the XOR logical process. Watermarks created using LSB Watermarking are easy to detect. The incentive here is to increase its security. The XOR method is displayed with watermarking in Figure 6. In this algorithm, XNOR (not XOR) can be used operatively instead of the XOR trigger. Plotted with watermarking in the spatial area of the color image using the Border, LSB and XOR algorithms with watermarking. The authentication algorithm is developed using a watermark. "Watermarking XOR technology" and "LSB watermarking technology" were found useful. The most important feature of watermarking XOR technology is: It is impossible to extract watermark without the original image in LSB technology with watermarking; in order to make this technology useful, the watermark needs to be converted to QR code. The QR code consists of 0 and 1 only. These values are closest to black on an 8-bit image and the human eye cannot be separate the difference between these two values. The results of the experiment indicate that the watermark can be included in the page or character limits. If the proposed techniques are combined with artificial intelligence systems, the security of the Qur'an will be increased [65]. Watermark with LSB techniques can easily be easily watermark-extracted by attackers. The very important feature of the Xor method with watermarking is that; it is impossible to extract the watermark without the original image in the LSB technique with watermarking; although the watermark can be extracted no matter what the picture. XOR technology is used with watermarking to remove LSB technology defects with watermarking [66]. Suggested a digital watermarking approach to protect and secure the most sensitive text by XOR-ing method only those Quranic letters that have certain diacritics related to it. Special features are used to convert the standard Unicode system to UTF-8. This process is that it can be applied to sensitive text since it does not alter the nature of the text. The proposal is promising with little complexity of memory and time complexity compared to existing approach.

### 6.2 ZERO WATERMARKING

Used a zero-watermark text method based on the features of sentences and linear weight equations used to identify important sentences. Using the entropy clause, the relevance of the text of the room and the length of the sentence, use the detailed weighting to obtain the final weight of each sentence. Which are encrypted and registered with a trusted third party called the CA? In order to enhance durability and have been secured less than the proposed method, some very important words are chosen to build the watermark [67]. In order to get a solution for authentication and copyright protection of digital contents and text documents utilizing a zero-watermarking algorithm for copyright protection of text documents. This method integrates the occurrence frequency of articles and vowels characters in the text to keep it. This technique a zero-watermarking approach provides a robust solution for text watermarking problem. The results illustrate that procedure was concluded more robust even when the watermark length was shorter; also, they were highly protected, and efficient with minimal computational requirements. The watermark remains resilient after attacks which make the watermark more active and robust. Therefore, no capacity of this algorithm [68]. According to the Chinese phonetic alphabet (CPAZW) to determine the discrepancy between rigidity and concealment by including watermark information without modifying any host

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

information. This is the first algorithm to conduct research on the structure of Chinese phonetic alphabets, transforming complex characters into the simplest alphanumeric alphabet in Chinese, a technology in which there are no watermarks based on the high concealment of Chinese phonetic alphabets. This algorithm succeeded in solving the contradiction between concealment and power. The technology with no watermarks depends on the high concealment of the Chinese phonetic alphabet. Denial of the methods mentioned, such as the small size of the data to be merged and the weak ability to attack. [69]. An adaptive algorithm with a watermark that is not text-based. The algorithm can be used to protect all digital text content from manipulating and manipulating illegal content by including the watermark logo of the original publisher in the same document corresponding to the cover document to create a special key, which is compared with a character key from another style document to demonstrate authenticity and ownership in text and watermark media Multiple. This reduced decryption time has an advantage since the encoding process (and greater delay) is processed for each text document only once, while the decryption process is usually tested multiple times to verify by many online clients [70]. Consumed a zero technique with water which proves that physiognomies are Arabic characters without changing the text. The watermark key depends on the properties of the verse/host. A key is created for each verse of the Qur'an where the initial phase of the name/number of Surat and number of numbers/sec is verified. Therefore, this method provides the minimum requirements for hardware resources, which reduces the cost of implementation [71]. General zero watermark recognition protocol. Watermarks were produced using the Uninterruptible Logistics Map of powerful text structures that are output from plain text. Watermarks are included in plain text by substituting synonyms based on the method of adjacent. The link value is calculated to reveal the property [72].

### 6.3 SVD SINGULAR VALUE DECOMPOSITION TECHNICS

An improved new proposal based on individual value analysis (SVD) is proposed to watermark the data. It suggests that the content of text images published on the Internet and works on images of Quranic textual texts should be validated as an ideal case study for compassionate compassion. The proposed approach has strong robustness and security against the most serious attacks on digital image-based content that includes engineering attacks on content. The watermark can be extracted in most of the following cases for different types of known attacks. In order to analyze performance results. Therefore, it can be easily applied to protect and document another sensitive digital image content [73].

### 6.4 NATURAL LANGUAGE

A watermarking scheme on the basis of Natural Language. It was a fantastic idea to introduce Natural Language into the contrast of encryption. Till now, the text watermarking is applied for English, Turkish, Chinese, and Arabic language text by different methods. This study contains novel watermarking techniques for English language text documents. The focus is on grammatical principles in conjunction, modal verbs and pronouns to generate encrypted watermark message. It uses AES (Advanced Encryption Algorithm) for encryption and decryption. Hence his method is quite effective and can be considered for future development process [74]. The German text, the scheme offers a number of advantages. First, it is blind and, therefore, can detect and extract watermarks in the obscurity of non-watermark text. Second, part of our embedding styles preserves the structure and concept of the original text. Thirdly, the techniques could be adapted to other languages (English) and therefore non-binding in German. The assessment, based on 1,645 evaluations by 131 people, revealed several interesting notices on our system. There are also a number of obvious restrictions represented by low payload and imperceptibility, our plan is not appropriate for short texts at this stage, and last but not least, author mention that robustness was very little addressed [75] [76].

An invisible digital watermark relies on the text information contained in the web page. Watermarks

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

are based on grammatical and semantic principles, which are encrypted and converted to white spaces using binary characters controlled before they are embedded in a Web page. Hyper Text Markup Language (HTML) is used as a cover file to embed watermarks that have been designed. The proposed method has been validated against several attacks to find a very good durability [77]. According to used new-defined characters generated by the TrueType, the function is used as a watermark; this watermark is embedded into the host Microsoft Word document according to some rules. The experimental results it can be applied to both Chinese and English language. In addition, this method has completely imperceptibility, high robustness against many kinds of attacks and can determine the tamper locations. The technique can be used in the copyright protection for Microsoft Word documents. Though this method has efficient performance, it should be optimized [78] [83].

## 7. QURAN TEXT STRUCTURE

Quran consists of 114 suras (as sections), which have clear limits in the script. In overall, the lengthier suras seem earlier in the Quran while the shorter ones seem later. Each sura is formed from some verses. Neither the number of verses in suras nor the word count of verses is similar. The Arabic Quran corpus consists of 77784-word tokens and 19287-word types [65]. Arabic language of Quran differs from Modern Standard Arabic (MSA) which is utilized in writing and formal spoken. The main difference is in grammar and lexicon. Based on this difference, it can be argued that even methods and tools for analyzing modern Arabic text may need to be changed to analyze the Quranic discourse. Therefore, for Quran text watermarking that we can use same tools and techniques that we utilized in Arabic text watermarking [66].

*Table1: Summarization of the main characteristic features of Text watermarking Techniques in Spatial Domain*

| Techniques | Algorithm | Drawbacks |
|---|---|---|
| LSB | Kashidas[20] | The algorithm circularly embeds the watermark message N times in the host document. |
| | Devanagari language 'Saravanan' [6] | This technique belongs to the open-space method which follows structural approach and does not affect the value and meaning of the document |
| | (LSB) substitution method [22] | The receiver side extracted text images are much similar to their original images at the transmitter end. |
| | (LSB) based hash function [54] | Which is a blind technique as image-related information is never sent to receiver separately |
| | LSB [29] | LSB offers more amount of distortion and is less secure due to sequential mapping |
| PT | Integrity verification and authentication [25] | Not capacity and imperceptibility |
| | Text image protection and source authentication [26] | Low imperceptibility |
| | Quran translation [27] | There is not an annotated application for every translation |
| | A zero-text watermarking algorithm [70] | Not capacity with low imperceptibly |
| AW | Word space in Arabic text [28] | The pseudo-space is very small space used to separate two parts of the same word |
| | Multiple watermarks [29] | The watermark is added to the low sub-band of DWT transformation |
| | Quantum watermark strategy [30] | The block-based concept is being unable to locate the tampered pixels accurately |
| | Features points [26] | The data in the LSB bits are visually insignificant also it is not robust to attacks methods. |

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

| | | |
|---|---|---|
| Texture | A novel watermarking [52] | Though the tampered area is very small |
| | 'swar' (vowel [35] | The limitations like lower robustness, deficient imperceptibility and unsuitable security of existing digital text watermarking algorithm |
| | JPEG format [34] | Not stored in the JPEG image pixel data and no change is explored in the frequency coefficients |
| | Quadtree [54] | This algorithm is only suitable for those areas with a large number of arbitrary texture images. |
| | Human visual system [55] | Data augmentation or authorized replication, copyright protection has become an exigent challenge |
| SS | Spiral manner [13] | An attack the shared images to load a misguiding meaning of the verse. |
| | Hiding a small text [38] | No Authentication |
| | Novel digital text watermarking [39] | Low capacity because of the limit dataset |
| | A spatial light modulator (SLM)[58] | A pattern imposed on the SLM, which can disperse the incoming spectrum into a range of varying deflection angles. |
| | Schur decomposition [64] | This algorithm not only guarantees the invisibility of watermarking but also has strong robustness in the operations of common image processing and geometric attacks. |
| LSC | A pseudo-space [42] | It is also easy to perform and does not need complex calculation. |
| | Multiple watermarking algorithms [43] | The limit of the dataset and low security and imperceptibility |
| | Diacritics and kashida techniques [14] | Used as the basis for hiding secret bits. Not secret massage |
| | Lines displacement [16] | It has low capacity where it can hold one bit per line |
| | Watermarking of encryption message [44] | To overcoming the watermarking resistance in the watermarked text against the attacks of writing changes and loss of the large volume of the data. |
| WSC | Natural language components and UMARAM encryption technique [49] | Low imperceptibility. |
| | The characteristics of Arabic language [8] | The lower the technical capacity, the larger the carriers file to hide the confidential message. |
| | Kashida[33] | Unable to locate the tampered pixels accurately; this might be important for specific applications such as in the military communication. |
| | The special attributes of word document [18] | The integrity of our context does not change. |
| | Word space [5] | The semantic similarity of words, such as the techniques based on how net. |
| | An enhanced-Kashida encoding method [53] | Low capacity and without authentication. |
| | Extension 'Kashida' Character [54] | In this robustness such that the output font cannot be a gauge. |
| | Kashida[51] | There is a small percentage and the stego-object is not altered much compared with the original cover text. |
| FC | LZW compression [59] | Low Authentication also reduces computational complexity |
| | Natural language [51] | Low Authentication and not capacity. |
| | Text image protection and source authentication [18] | Low imperceptibility. |
| | Novel digital text watermarking [57] | Low capacity because of the limit dataset. |

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

| | | |
|---|---|---|
| Other Techniques | XOR-ing [63] | This algorithm does not depend on any format and can keep it. |
| | XOR and LSB techniques [64] | XOR watermarking technique is used to eliminate the drawbacks of the LSB watermarking technique. |
| | Zero-watermarking method [62] | The text is not associated with word order, the sentence length is calculated out liquidation function words. |
| | Chinese phonetic alphabets zero watermarking algorithm [66] | The little size of data to be embedded and lessen anti-attack capability. |
| | An adaptive algorithm for text-based zero-watermarking [49] | Computational-times with low capacity and security. |
| | A zero-watermarking [2] | Not capacity and imperceptibility. |
| | Singular value decomposition (SVD) [34] | Not capacity. |
| | Natural language [74] | Low authentication and not capacity. |
| | A natural watermark scheme (NLW) [60] | Not security. |
| | Natural language components and UMA RAM encryption technique [55] | With the small volume of attack and robustness. |
| | HTML (Hyper Text Markup) language 75] | Not robust against the basic amendment and deletion attacks. |

*Table2: Summarization of the main characteristic features*

| Ref | Capacity | Authentication | Imperceptibility |
|---|---|---|---|
| [20] | low | -- | high |
| [6] | low | high | low |
| [22] | low | -- | high |
| [54] | -- | high | low |
| [29] | -- | low | high |
| [25] | -- | Moderate | -- |
| [26] | Moderate | high | low |
| [27] | low | low | -- |
| [70] | - | high | low |
| [28] | high | high | -- |
| [29] | low | high | -- |
| [30] | low | -- | low |
| [26] | Moderate | low | low |
| [52] | low | high | low |
| [35] | low | high | -- |
| [34] | low | -- | low |
| [54] | low | low | -- |
| [55] | high | low | -- |
| [13] | high | Moderate | high |
| [38] | high | -- | Moderate |
| [39] | low | high | high |
| [58] | low | low | -- |
| [64] | low | high | -- |
| [42] | high | -- | high |
| [43] | high | low | low |
| [14] | high | low | -- |
| [16] | low | low | -- |
| [44] | low | Moderate | low |
| [49] | -- | high | -- |
| [8] | Moderate | -- | low |
| [33] | high | low | -- |
| [18] | high | low | high |
| [5] | high | -- | Moderate |
| [53] | low | -- | high |
| [54] | Moderate | low | -- |
| [51] | low | -- | Moderate |
| [59] | high | low | -- |
| [51] | -- | low | Moderate |
| [18] | Moderate | high | low |
| [57] | low | high | high |
| [63] | low | -- | -- |
| [64] | low | -- | low |
| [62] | Moderate | low | -- |
| [66] | low | Moderate | Moderate |
| [49] | low | low | -- |
| [2] | -- | high | -- |
| [34] | -- | high | Moderate |
| [74] | -- | low | Moderate |
| [60] | Moderate | -- | low |
| [55] | -- | high | -- |
| [75] | -- | Moderate | -- |

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

*Table3: The Usefulness &Drawbacks of Text watermarking Techniques in Spatial Domain*

| Method | Usefulness | Drawbacks |
|---|---|---|
| **LSB** | Resistance to geometric attacks such as removal of inner distance, scaling, rotation, simplicity, and lack of computational complexity and conceptual clarity. | Fail in facing cropping attacks, compression, and a low-pass filter. Robustness restriction, capacity limitation in data storage and low resistance |
| **PT** | High level of robustness against the more type of attacks | It can hide only a little amount of information |
| **AW** | To add pseudo-random noise style to the strength of image pixel. | The noise is generated by a key, such that the connection, between the numbers of different keys, will be lessening |
| **Texture** | In this technique hides data inside continued random texture styles of a picture. | This method is suitable for those areas with a big number of arbitrary texture images |
| **SS** | In this technique, by using energetic signal propagation a high resistance can be achieved. | The blind watermarking method in order to embed the watermark doesn't use the host signals. Do not specifically protect the value of DC blocks |
| **LSC** | The word shift encoding for the length of the lines. It's Very clear. | Have regular line spacing between neighboring lines within a paragraph. |
| **WSC** | The spacing between words is different by using horizontally shifting the place of the words within text lines, therefore, watermarking the document unambiguously. | The spacing between words in the unencoded text. |
| **PC** | In this method alter the conditions of particular letters by using either lengthening or shortening. | These end lines are changed by using extending or shortening their lengths by one (or more) pixels, further. Not changing the end line characteristic. |

## 8. CURRENT PROBLEMS, CHALLENGES, AND OPEN RESEARCH ISSUES.

There are many works on the text watermarking in the spatial domain, few of these works focusing on better text quality, while others aimed at payload capacity or provide more security method. The aim of those works is to get a robust and imperceptible method. In addition, some of these works having impractical assumptions to secure the method, while the others can be used only for checking the effect of the method on different attacks [80]. According to the related work, the challenges and open issues in this important research is to accept the existence of the text watermarking in the spatial domain and dealing with this fact by detecting the weak points and drawbacks during a new robust and imperceptible technique. This open issue motivates the researchers to suggest an effective solution addressing this concern [1]. Finally, all the methods presented in this work cannot be applied to all Arabic texts. Therefore, more techniques have been using with sensitive Arabic texts such as religious and official documents, including the Holy Quran, where it is not allowed to modify the text's position, in other words, and watermark technology. Algorithms or techniques can be used in other fields and topics not only for Arabic texts and watermarking techniques.

The summarization of the main characteristic features and the drawbacks of the text watermarking techniques in the spatial domain have been shown in Table 1, 2 and Table 3. Respectively.

## 9. CONCLUSION

Information hiding is becoming an extensive field that fetches a serious research interest. This is the reason that text watermarking is earning more attraction in the direction of securing data through network communication. This study summarizes the current text watermarking techniques in a spatial domain, also analyzed different problems and the drawbacks of each method that has been innovated in the last few years. The implemented watermarking techniques are different in their propositions. Hence, few of these works are dealing with the text quality, while others are working on data hiding capacity, imperceptibility, authentication. The limitation of this study only spatial domain and Arabic text watermarking. All

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

these techniques can give more efforts and vital for future research regarding the text watermark. Finally, the open issues of this work motivate the researchers to suggest an effective solution addressing this concern by text watermarking using Arabic text characters to get a high capacity, imperceptibility, and Authentication.

## ACKNOWLEDGEMENTS

The authors would like to thank Ministry of Higher Education (MOHE) and Universiti Teknologi Malaysia (UTM) for their educational and financial support. This work is conducted at Advances Informatics School (AIS) under Cyberphysical Sytems Research Group (CPS RG) and funded by Universiti Teknologi Malaysia (GUP Tier 1: QK130000.2538.18H4).

## REFERENCES:

[1] Sonka, M., Hlavac, V., & Boyle, R. "Image processing, analysis, and machine vision". Cengage Learning. (2014).

[2] Liu, Y., Zhu, Y., & Xin, G. "A zero-watermarking algorithm based on merging features of sentences for Chinese text". Journal of the Chinese Institute of Engineers, 38(3), (2015): 391-398 .

[3] Mahdi Hashim, M. O. H. A. M. M. E. D., Mohd Rahim, And Mohd Shafry. "Image Steganography Based On Odd/Even Pixels Distribution Scheme And Two Parameters Random Function." *Journal Of Theoretical & Applied Information Technology* 95.22 (2017).

[4] Muhammad, Khan, et al. "A novel magic LSB substitution method (M-LSB-SM) using multi-level encryption and achromatic component of an image." Multimedia Tools and Applications 75.22 (2016): 14867-14893.

[5] RIZZO, S. G., BERTINI, F., & MONTESI, D. (2016, AUGUST). TEXT AUTHORSHIP VERIFICATION THROUGH WATERMARKING. IN INTELLIGENCE AND SECURITY INFORMATICS CONFERENCE (EISIC), 2016 EUROPEAN (PP. 168-171). IEEE.

[6] Salloum, S. A., AlHamad, A. Q., Al-Emran, M., & Shaalan, K. (2018). A Survey of Arabic Text Mining. In *Intelligent Natural Language Processing: Trends and Applications* (pp. 417-431). Springer, Cham.

[7] Rizzo, S. G., Bertini, F., Montesi, D., & Stomeo, C. (2017, July). Text Watermarking in Social Media. In *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017* (pp. 208-211). ACM.

[9] Ahuja, R., & Bedi, S. S." Video watermarking scheme based on candidates I-frames for copyright protection". Indonesian Journal of Electrical Engineering and Computer Science, 5(2), (2017): 391-400.

[10] Alotaibi, R. A., & Elrefaei, L. A. "Arabic Text Watermarking: A Review". arXiv preprint arXiv:1508.01485 .(2015).

[11] Patil Nitin N, J B Patil. "Implementation of a Novel Watermarking Technique for Devanagari Text." International Journal of Information and Electronics Engineering 5(5), 2015.

[12] Nematollahi, M. A., Vorakulpipat, C., & Rosales, H. G." Digital Watermarking". Springer Singapore. (2017).

[13] Shaker, A. A., Ridzuan, F., & Pitchay, S .A. "Text Steganography using Extensions Kashida based on the Moon and Sun Letters Concept". international journal of advanced computer science and applications, 8(8), (2017): 286-290 .

[14] Rhazlane, S., El Ouazzani, A., Harbi, N., Kabachi, N., Badir, H., Tanger, E. N. S. A., & Tangier, M. "Data Alteration: A Better Approach to Securing Cloud Data with Encryption". In Proceedings of the EDA Conference Revue des Nouvelles Technologies de l'Information, (2017). (pp. 2-3).

[15] Tehranchi, B., Petrovic, R., Winograd, J. M., & Angelico, D. A. U.S. Patent No. 9,704,211. Washington, DC: U.S. Patent and Trademark Office, (2017).

[16] Khaleghparast, R. "Image Rightful Ownership Watermarking Method for the Cloud Environment". (Doctoral dissertation, Auckland University of Technology), (2017).

[17] Nematollahi, M. A., Vorakulpipat, C., Gamboa-Rosales, H., Martinez-Ruiz, F. J., & Jose, I. "Digital Speech Watermarking Based on Linear Predictive Analysis and Singular Value Decomposition". Proceedings of the

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

National Academy of Sciences, India Section A: Physical Sciences, 87(3), (2017): 433-446.

[18] Nematollahi, M. A., Vorakulpipat, C., & Rosales, H. G. "Security Enhancement of Digital Watermarking". In Digital Watermarking, Springer, (2017). Singapore. (pp. 191-203).

[19] Nematollahi, M. A., Vorakulpipat, C., & Rosales, H. G. "Natural language watermarking". In Digital Watermarking, Springer, (2017). Singapore. (pp. 103-119).

[20] Jothimani, S., & Betty, P. "Image authentication using global and local features". In Green Computing Communication and Electrical Engineering (ICGCCEE), 2014 International Conference on IEEE, (2014, March). (pp. 1-5).

[21] Patil, N. N., & Patil, J ."Implementation of a Novel Watermarking Technique for Devanagari Text". International Journal of Information and Electronics Engineering, 5(5),*(2015) :* 352 .

[22] Zakariah, M., Khan, M. K., Tayan, O., & Salah, K. "Digital Quran Computing: Review, Classification, and Trend Analysis". Arabian Journal for Science and Engineering, 42(8), (2017): 3077-3102.

[23] Rizzo, Stefano Giovanni, Flavio Bertini, and Danilo Montesi. "Content-preserving text watermarking through Unicode homoglyph substitution." Proceedings of the 20th International Database Engineering & Applications Symposium. ACM, 2016.

[24] Hashim, Mohammed, Et Al. "A Review And Open Issues Of Multifarious Image Steganography Techniques In Spatial Domain." Journal Of Theoretical & Applied Information Technology 96.4 (2018).

[25] Alginahi, Yasser M., Muhammad N. Kabir, and Omar Tayan. "An enhanced Kashida-based watermarking approach for increased protection in Arabic text-documents based on frequency recurrence of characters." International Journal of Computer and Electrical Engineering 6.5 (2014): 381.

[26] Vasudev, Rajiv. "A review on digital image watermarking and its techniques." Journal of Image and Graphics 4.2 (2016): 150-153.

[27] Das, Ujjal Kumar, Shefalika Ghosh Samaddar, and Pankaj Kumar Keserwani. "Digital Forensic Enabled Image Authentication Using Least Significant Bit (LSB) with Tamper Localization Based Hash Function." Intelligent Communication and Computational Technologies. Springer, Singapore, 2018. 141-155.

[28] Upadhyay, Anamika, and Nirupama Tiwari. "A survey on QuadTree based digital watermarking." watermark 6.12 (2017).

[29] Singh, Pooja Rajvir. "A Review on Digital Image Watermarking and its Methods." (2017).

[30] Tabrizi, Arash Amini, and Rohana Mahmud. "Issues of coherence analysis on English translations of Quran." Communications, Signal Processing, and their Applications (ICCSPA), 2013 1st International Conference on. IEEE, 2013.

[31] Alginahi, Yasser M., Omar Tayan, and Muhammed N. Kabir. "An adaptive zero-watermarking approach for authentication and protection of sensitive text documents." (2013).

[32] Laouamer, Lamri, and Omar Tayan. "An enhanced SVD technique for authentication and protection of text-images using a case study on digital Quran content with sensitivity constraints." Life Science Journal 10.2 (2013): 2591-2597.

[33] Alotaibi, Reem A., and Lamiaa A. Elrefaei. "Utilizing word space with pointed and un-pointed letters for Arabic text watermarking." Computer Modelling and Simulation (UKSim), 2016 UKSim-AMSS 18th International Conference on. IEEE, 2016.

[34] Joshi, Kamaldeep, Rajkumar Yadav, and Ashok Kumar Yadav. "An Additive Watermarking Technique in Gray Scale Images Using Discrete Wavelet Transformation and Its Analysis on Watermark Strength." World Academy of Science, Engineering and Technology, International Journal of Computer, Electrical, Automation, Control and Information Engineering 10.7 (2016): 1428-1433.

[35] Yang, Yu-Guang, et al. "Analysis and improvement of the watermark strategy for quantum images based on quantum Fourier

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

transform." Quantum Information Processing 12.8 (2013): 2765-2769.

[36] Agarwal, Monika. "Text steganographic approaches: a comparison." arXiv preprint arXiv:1302.2718 (2013).

[37] Rashid, Aaqib. "Digital Watermarking Applications and Techniques: A Brief Review." International Journal of Computer Applications Technology and Research 5.3 (2016): 147-150.

[38] Khalil, Mohammed S., et al. "Two-layer fragile watermarking method secured with chaotic map for authentication of digital holy Quran." The Scientific World Journal 2014 (2014).

[39] Pourabasi, Elmira. "Watermarking of Text Messages in Bytes of Little Value without Causing Confusion." International Journal of Computer Science and Information Security 14.5 (2016): 446.

[40] Patil, Nitin Namdeo, and Jayantrao Bhaurao Patil. "Performance Analysis of a Novel Text Watermarking Technique for Devanagari Text." Proceeding of International Conference on Intelligent Communication, Control and Devices. Springer, Singapore, 2017.

[41] Singh, Amit Kumar, et al. "Robust and imperceptible spread-spectrum watermarking for telemedicine applications." Proceedings of the National Academy of Sciences, India Section A: Physical Sciences 85.2 (2015): 295-301.

[42] Hisham, Syifak Izhar, et al. "Localization watermarking for authentication of text images in quran with spiral manner numbering." Advances in Information Technology for the Holy Quran and Its Sciences (32519), 2013 Taibah University International Conference on. IEEE, 2013.

[43] Kumar, Sanjay, and Ambar Dutta. "A novel spatial domain technique for digital image watermarking using block entropy." Recent Trends in Information Technology (ICRTIT), 2016 International Conference on. IEEE, 2016.

[44] Al-maweri, Nasraddin Ahmed Salem, et al. "Robust Digital Text Watermarking Algorithm based on Unicode Extended Characters." Indian Journal of Science and Technology 9.48 (2016).

[45] Zakariah, Mohammed, et al. "Digital Quran Computing: Review, Classification, and Tren Analysis." Arabian Journal for Science and Engineering 42.8 (2017): 3077-3102.

[46] Kushwah, Vineet, Sumit Tiwari, and Manvendra Gautam. "A review study on digital watermarking techniques." International Journal of Current Engineering and Scientific Research 3.1 (2016): 189-193.

[47] Alotaibi, Reem A., and Lamiaa A. Elrefaei. "Utilizing word space with pointed and un-pointed letters for Arabic text watermarking." Computer Modelling and Simulation (UKSim), 2016 UKSim-AMSS 18th International Conference on. IEEE, 2016.

[48] Rui, Xu, Chen XiaoJun, and Shi Jinqiao. "A multiple watermarking algorithm for texts mixed Chinese and English." Procedia Computer Science 17 (2013): 844-851.

[49] Ahvanooei, Milad Talebi, Seyed Hashem Tabasi, and Sajad Rahmani. "A novel approach for text watermarking in digital documents by zero-width interword distance changes." DAV International Journal of Science 4.3 (2015): 550-558.

[50] Mittal, Misha, and Dinesh Kumar. "Word Sense Disambiguation Approaches for Indian Languages: A Survey." International Journal of Computer Applications & Information Technology 9.1 (2016): 182.

[51] Raut, Samruddhi S., and A. R. Mune. "A Review Paper on Digital Watermarking Techniques." International Journal of Engineering Science 10460 (2017).

[52] Alotaibi, Reem A., and Lamiaa A. Elrefaei. "Utilizing word space with pointed and un-pointed letters for Arabic text watermarking." Computer Modelling and Simulation (UKSim), 2016 UKSim-AMSS 18th International Conference on. IEEE, 2016.

[53] Shaker, Anes A., Farida Ridzuan, and Sakinah Ali Pitchay. "Text Steganography using Extensions Kashida based on the Moon and Sun Letters Concept." INTERNATIONAL JOURNAL OF ADVANCED COMPUTER

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

SCIENCE AND APPLICATIONS 8.8 (2017): 286-290.

[54] Alginahi, Yasser M., Omar Tayan, and Muhammad N. Kabir. "A zero-watermarking verification approach for Quranic verses in online text documents." Advances in Information Technology for the Holy Quran and Its Sciences (32519), 2013 Taibah University International Conference on. IEEE, 2013.

[55] Al-No fair, Safia Meteb, Manal Mohammed Fattani, and Adnan Abdul-Aziz Gutub. "Merging Two Steganography Techniques Adjusted to Improve Arabic Text Data Security." Journal of Computer Science & Computational Mathematics (JCSCM), 6 (3) (2016): 59-65.

[56] Khan, Esam Ali. "using arabic poetry system for steganography." (2014).

[57] Kumar, Avi, and Keerthi Kumar KM. "Enhanced LSB Algorithm for Stegano Communication." Journal of Web Development and Web Designing 1.1, 2, 3 (2017).

[58] Alotaibi, Reem A., and Lamiaa A. Elrefaei. "Improved capacity Arabic text watermarking methods based on open word space." Journal of King Saud University-Computer and Information Sciences (2017).

[59] Al-nofair, Safia, Manal Fattani, and Adnan Gutub. "Capacity Improved Arabic Text Steganography Technique Utilizing 'Kashida'with Whitespaces." The 3rd International Conference on Mathematical Sciences and Computer Engineering (ICMSCE2016). 2016.

[60] Alginahi, Yasser M., Muhammad N. Kabir, and Omar Tayan. "An enhanced Kashida-based watermarking approach for increased protection in Arabic text-documents based on frequency recurrence of characters." International Journal of Computer and Electrical Engineering 6.5 (2014): 381.

[61] Hakak, Saqib, et al. "Approaches for preserving content integrity of sensitive online Arabic content: A survey and research challenges." Information Processing & Management (2017).

[62] Kabir, Muhammed N., Omar Tayan, and Yasser M. Alginahi. "Evaluation of watermarking approaches for Arabic text documents." International Journal of Computer Science and Information Security 11.3 (2013): 49.

[63] Khadim, Umair, et al. "Information hiding in text to improve performance for worddocument." International Journal of Technology and Research 3.3 (2015): 50.

[64] Malik, Aruna, Geeta Sikka, and Harsh K. Verma. "A high capacity text steganography scheme based on LZW compression and color coding." Engineering Science and Technology, an International Journal 20.1 (2017): 72-79.

[65] Cheema, Prabhjot Kaur, and Kamaljit Kaur. "Improved Text Watermarking Technique." International Journal of Advanced Research in Computer Science 5.5 (2014).

[66] Mussa, Sarah Abdul-Ameer, and Sabrina Tiun. "Word sense disambiguation on english translation of holy quran." Bulletin of Electrical Engineering and Informatics 4.3 (2015): 241-247.

[67] Kamaruddin, Nurul Shamimi, et al. "A Review of Text Watermarking: Theory, Methods, and Applications." IEEE Access 6 (2018): 8011-8028.

[68] Shih, Frank Y. Digital watermarking and steganography: fundamentals and techniques. CRC press, 2017.

[69] Tuncer, T., F. Ertam, and E. Avci. "A watermarking application for authentication of Holy Quran." Advances in Information Technology for the Holy Quran and Its Sciences (32519), 2013 Taibah University International Conference on. IEEE, 2013.

[70] Kamaruddin, Nurul Shamimi, Amirrudin Kamsin, and Saqib Hakak. "Associated diacritical watermarking approach to protect sensitive arabic digital texts." AIP Conference Proceedings. Vol. 1891. No. 1. AIP Publishing, 2017.

[71] Huang, Jiafeng, et al. "Sentiment analysis of Chinese online reviews using ensemble learning framework." Cluster Computing (2018): 1-16.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[72] Singh, Prabhishek, and R. S. Chadha. "A survey of digital watermarking techniques, applications and attacks." International Journal of Engineering and Innovative Technology (IJEIT) 2.9 (2013): 165-175.

[73] Zhu, Ping, et al. "A text zero-watermarking algorithm based on Chinese phonetic alphabets." Wuhan University Journal of Natural Sciences 21.4 (2016): 277-282.

[74] Tayan, Omar, and Yasser M. Alginahi. "A review of recent advances on multimedia watermarking security and design implications for digital Quran computing." Biometrics and Security Technologies (ISBAST), 2014 International Symposium on. IEEE, 2014.

[75] Laouamer, Lamri, and Omar Tayan. "An enhanced SVD technique for authentication and protection of text-images using a case study on digital Quran content with sensitivity constraints." Life Science Journal 10.2 (2013): 2591-2597.

[76] Jia, Weijia, et al. "Blind detection of spread spectrum flow watermarks." Security and Communication Networks 6.3 (2013): 257-274.

[77] Makbol, Nasrin M., and Bee Ee Khoo. "A new robust and secure digital image watermarking scheme based on the integer wavelet transform and singular value decomposition." Digital Signal Processing 33 (2014): 134-147.

[78] Kaur, Harmandeep, Er Simarjeet Kaur, and M. Tech. "Text watermarking using techniques dct and dwt: A review." International Journal of Computer Application and Technology1.1 (2014): 1-6.

[79] Halvani, Oren, Martin Steinebach, and Lukas Graner. "Towards Imperceptible Natural Language Watermarking for German." Proceedings of the 5th ACM Workshop on Information Hiding and Multimedia Security. ACM, 2017.

[80] Mir, Nighat. "Copyright for web content using invisible text watermarking." Computers in Human Behavior 30 (2014): 648-653.

[81] Zhang, Shi-ru, et al. "New digital text watermarking algorithm based on new-defined characters." Computer, Consumer and Control (IS3C), 2014 International Symposium on. IEEE, 2014.

[82] Katzenbeisser, Stefan, and Fabien Petitcolas. Information hiding. Artech house, 2016.

[83] Upadhyay, Anamika, and Nirupama Tiwari. "A survey on Quad Tree based digital watermarking." watermark 6.12 (2017).