Robust Color Image Watermarking and Web-Tracing System Using Digital Wavelet Transform and Mobile Agents

Ziad Al Bkhetan[1] and Nawar Al Awa[2]

[1]Information Systems and Software Engineering department, Faculty of Information Technology Engineering, Damascus University

[2]Systems and Networking department, Faculty of Information Technology Engineering, Damascus University

Abstract *This paper describes a complete system for watermarking and web tracing of color images.The watermark system uses DWT (Discrete Wavelet Transform) that gives the watermark a significant robustness and invisibility. The watermarking system is applied to color images by converting color coding from RGB to YUV space, and applying the watermark on V channel. The system adopts blind detection model that depends on SIFT (Scale invariant feature transform) algorithm for watermark embedding and extraction. These points are used as reference points in both phases (embedding and extraction), because of their high stability against possible attacks. The system is also able to trace images on the web, using watermark mobile agents, sent by the watermark agency to every connected node, contracted with the agency, in order to examine the stored files in them. This system is very efficient in dealing with large-size images across congested networks. It uses blind-model based detection to improve tracing system performance, so that the agent doesn't have to "carry" the original image through its navigation.A portal is also implemented to ensure simple and easy-to-use services.*

Keywords: *Digital Image Watermarking, Discrete-Wavelet-Transform, Scale-Invariant-Feature-Transform, Mobile Agents, Watermark Agency.*

Received April 3, 2012; Accepted August 2, 2012

1. Introduction

Nowadays, the huge developments in multimedia and information technology domain, and the world-wide spread of internet, have made of it an important communication medium that people depend on to exchange digital documents, songs, and videos. However, exposing proprietary files (that contain any type of document, images, or video) publicly, facilitates using them "illegally", without paying copyright fees. As a result, it becomes highly urgent to find ways to protect owners' rights, and trace the "illegal" copies of these files [6].

A lot of copyright protection techniques appeared, like cryptography, in which the files are only readable by the person who has the secret key.

Another technique is digital signature. This approach is used for authentication purpose, where information is added to the original content to detect any change in the document content.

Watermarking is also widely used in this purpose. It is based on adding some information to the original file called watermark, which could be visible or invisible. This information can help not only for copyrights protection, but also to monitor and trace these watermarked files to detect any illegal use.

Research in this domain is looking for increasing reliability, and improving watermark persistence against the deliberate attacks, that try to remove it from files, or introduce intentional damage to make the watermark invalid.

To obtain copyright protection for real and legal owners, we need a fast and effective search technique, in order to search in the internet for the files, and to insure that these files are used by authorized persons and in legal way.

One of the used methods for tracing purpose is called "Network spider" or "Crawler". It mainly depends on the links between website pages in order to navigate simply among pages. Then, it downloads the files to check them locally, i.e. on the local server responsible for this operation [2]. This technique, however, could not be used in our system because it is restrained to a specific kind of files, and because of the huge number of these files. Therefore, another approach is suggested in this paper. It consists of defining a watermark agent, which is a small-sized application, compared to any traced file.

This mobile agent could navigate among internet nodes to check any traced files.

2. Related Work

An image watermarking technique based on discrete wavelet transform (DWT) is suggested in [3]. In this method, the image is decomposed into five levels, and the watermark is taken from the first level. Then, watermarking is applied on every level. In the extraction process, the watermark is checked at every level and according to average value, the watermark is determined. This method is considered as blind model because it does not need the original image for watermark extraction process, and the watermarking process uses secret key. The main advantage of this approach is the adoption of multi-level watermarking, and selection of watermark according to most frequent parts between these levels, and it applies watermark on low frequency sub band which contains sufficient information about the image. Therefore, it may affect the quality of the image. Nevertheless, its main benefit is persistence.

In [4], a discrete wavelet transform is applied on the image, then a cosine transform is applied to watermark, and the watermark is embedded into a high-frequency sub band.

This model is considered as non-blind model, where it needs the original image in extraction process. The watermarking in high-frequency sub band usually causes weak persistence against attacks, but it does not affect the quality of image because these frequencies contain details information.

In [9], discrete wavelet transform is applied on the image, and then the low-frequency sub bands are selected for watermarking after dividing them into partitions, and a watermark is embedded in each one of them.

This model is also considered as non-blind model. In that approach, watermarking on low-frequency sub bands is used, and the image is divided into a number of blocks. Then, the watermark is embedded in every block. This way is useful especially for extraction process, because the extracted watermark could be built from most frequent and shared information among these blocks.

According to [11], a blind model is used. In this paper, we use the same strategy, but we apply this algorithm on color images after converting color coding for RGB to YUV. Then, we embed the watermark in V channel, and we modify the used filter in order to fit this applied improvement.

We choose this model to facilitate mobile agent work, so it does not need to "carry" the image for extraction process. We merged this algorithm with a tracing system to build a complete system for watermarking and tracing.

3. Suggested Approach

This paper describes a complete protection system for digital files, which is a web application that offers protection services. This application is accessible by any user, and has registration service to make all services available for all registered customers. These offered services are: file- watermarking service, watermarked-file tracing to insure its legal use by authorized users, and reporting service to get a log reports that contain all needed information about these files.

This system consists of two basic parts as described in figure (1):

- Watermarking agency: it is the most important functional part which performs operations like customers management, image watermarking, and tracing management that is charged of redirecting the agent to the nodes, exchanging messages with it, and displaying the reports. This agency contains a database to store all important information and results needed to achieve these operations.
- Watermark mobile agent: This is a small application that navigates among network nodes to check their files.

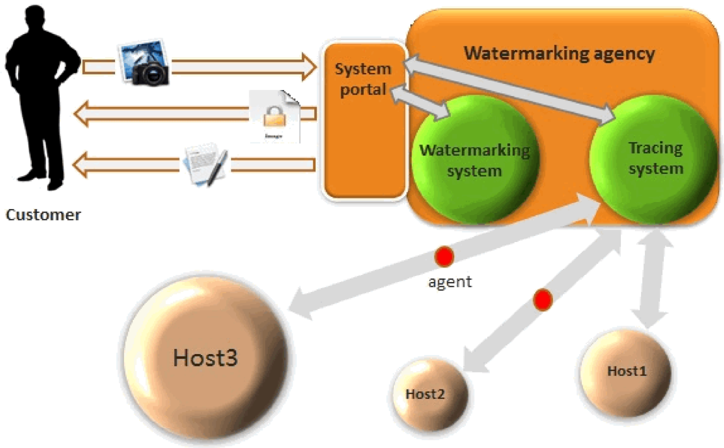

Figure 1. Proposed System.

Our approach aims to develop a system based on watermarking in frequency domain, in order to get invisible watermarks. In detection phase, the approach is based on blind detection model.

A similar watermarking system was applied in previous researches, but for grey images [11]. Our approach could be applied on both grey and color images in effective and strong way, without any effect on watermark requirements such as quality and transparency. This feature is obtained due to the implementation of color coding transformation from RGB to YUV, and applying watermarking on V channel. This improvement allows us to modify the used filter and strength parameter value used in [11] to obtain improved results.

Our approach offers several features that are:

- Direct and simple access for all users regardless of their basic knowledge, and their technical skills.
- Implementation of efficient image watermarking using Discrete Wavelet Transform (DWT), and Scale Invariant Features Transform (SIFT) algorithm that shows good color persistence and immunity when exposed to noise or filtering attacks.
- Implementation of efficient, fast and handy tracing system, taking into account the huge files count over Internet and their size without increasing the stress and load on the network.
- Implementation of mobile watermark agent that could be adapted to host server conditions and to the emergency cases that may occur in it. Therefore, the agent must be able to pause, continue and stop its work and could be notified by all expected events regardless of host server operating system "Window or Linux".
- Implementation of communication system between agent and agency, agent and hosts, agency and hosts, despite of different programming language used in building the shared application in this system.

4. Watermarking and Tracing System

Watermarking systems, that aim to protect owner's copyrights, are based on two different principles: watermarking of images using complex building algorithms in order to make watermark removing very difficult or watermarking images by invisible data with high transparency to make it undetectable, so the attacker can't determine its existence or its location. This latter approach is usually preferred because it maintains image quality, and reduces the probability of watermark removing, because it is difficult to detect an invisible watermark [8].

The suggested system depends on implementation of invisible watermarking using blind detection model, where the original image or watermark are not needed for extraction process, but the watermark is needed to compare extracted watermark with the original one. The main advantage of this system is its potential use for color images in frequency domain using DWT, because we aim to implement a strong watermarking algorithm, to guarantee watermark immunity against several possible attacks.

4.1. Filtering Function

Usually Haar wavelet filter is used in DWT-based watermarking system because it decomposes the signal into two parts [7]

- Low frequencies: this part contains approximation information of the image. These frequencies are obtained by passing the image signal through low-pass filter that depends on samples average value.
- High frequencies: this part contains details information of the image, and could be obtained by passing the image signal through high-pass filter that depends on samples difference.

Haar filter allows for reconstructing the signal from its decomposed components without any loss of information, if the used filters satisfy this condition:

$$|H(\omega)|^2 + |G(\omega)|^2 = 1 \qquad (1)$$

H(ω): frequency response of low-pass filter.
G(ω): frequency response of high-pass-filter.

In our approach, the following filtering functions are used:

$$\varphi_{2k}[n] = \begin{cases} 1/\sqrt{2} & , n = 2k, 2k+1 \\ 0 & , otherwise \end{cases} \qquad (2)$$

$$\varphi_{2k+1}[n] = \begin{cases} 1/\sqrt{2} & , n - 2k \\ -1/\sqrt{2} & , n = 2k+1 \\ 0 & , otherwise \end{cases} \qquad (3)$$

Where n = 0,1, 2,….., (original signal sample order).
k = 0,1, 2,….., (original signal sample order)

And we use k as a variable to get the corresponding sample and the next of it, because every output sample depends on the corresponding two samples in the original signal.

The output of filtering function for uni-dimensional signal X(k) is given by:

$$X[2k] = \frac{1}{\sqrt{2}}\left(x[2k] + x[2k+1]\right) \qquad (4)$$

$$X[2k+1] = \frac{1}{\sqrt{2}}\left(x[2k] - x[2k+1]\right) \qquad (5)$$

k: 0, 1, 2, ….. , *L* (signal sample order)
X: output signal, *x:* input signal
But if the signal is bi-dimensional, the filter should be applied for rows then for columns [1].

$$\text{Transform}_{\text{columns}}\,[\text{Transform}_{\text{rows}}\,[\text{Image}]] \qquad (6)$$

The image in our system is decomposed in a way to get the following frequency sub-bands:
1 – Low-low frequencies LL.
2 – Low-high frequencies LH.
3 – High-low frequencies HL.
4 – High-high frequencies HH.

4.2. Watermarking Algorithm

The watermark must be of small size that can't affect image quality, and sufficiently large to hold all needed information for tracing operation

In our system, we used the following watermark information:

- Secret number to distinguish the real owner of image.
- IP address of authorized device using the image.
- Permissions associated with the image like printing, copying, or displaying.
- Expiry date of using the image.
- Sensitivity field which determines its importance, and its degree of sensitivity. According to this field, the agent takes the appropriate action upon detection of illegal use. These actions could range from simple notification till file destroying.

4.3. Embedding and Extraction Processes

These operations are explained in figure 2. First, the watermark is converted into bipolar signal, and then divided into two partitions with the same length L1, L2.

We select the central part of the image for watermark embedding because this part contains the most important information in the image, and it allows avoiding image cropping.

Then, color coding is converted from RGB to YUV using the following equations [12]:

$$Y = 0.299 * R + 0.587 * G + 0.114 * B \quad (7)$$

$$U = -0.147 * R - 0.289 * G + 0.436 * B \quad (8)$$

$$V = 0.615 * R - 0.515 * G - 0.1 * B \quad (9)$$

We use V channel for embedding and we apply DWT on this channel using Haar wavelet filter [1].

The image signal is therefore decomposed into four sub-bands in the second level. These sub bands are LL, LH, HL, and HH.

The sub-band HH contains image details, thus embedding a watermark in it will make it very sensitive to several attacks like noise and filtering attacks.

But the LL sub band contains the most important information in the image. As a result, if the watermark is embedded in this sub band, the image quality will be affected. Consequently, we choose LH and HL sub bands to embed the watermark. This embedding is applied using these equations [11]:

$$f'HL(x,y) = fHL(x,y) + thHL + fHL(x,y).alfa.watermark1(l) \quad (10)$$

$$l = 1,2,\ldots,L1$$

$$f'LH(x,y) = fLH(x,y) + thLH + fLH(x,y).alfa.watermark2(l) \quad (11)$$

$$l = 1,2,\ldots,L2$$

Where:
f_{HL}(x,y): original coefficient at point (x, y) in HL sub band.
f'_{HL}(x,y): modified coefficient at point (x, y) in HL sub band.
f_{LH} (x,y): original coefficient at point (x, y) in LH sub band.
f'_{LH}(x,y): modified coefficient at point (x, y) in LH sub band.
Alfa: strength parameter.

We chose three different points with large characteristic scale using SIFT [5] algorithm. These points are considered as reference points in each sub band for embedding and extraction algorithm.
Color coding is converted from YUV into RGB using the following equations [12]:

$$R = Y + 1.14 * V \quad (12)$$

$$G = Y- 0.395 * U - 0.581 * V \quad (13)$$

$$B = Y + 2.032 * U \quad (14)$$

Similarly, the extraction algorithm uses SIFT algorithm to detect the watermark and compare it with the original one in order to get similarity average. If the similarity value is greater than certain threshold, it would be taken (figure 3).

4.4. Tracing System

In order to exploit the real benefits of a watermarking system in protecting the owner copyrights, an effective tracing system should be used to complete watermarking system. Internet contains a lot of nodes that include huge number of images, and we need to search for images that are illegally used. Internet "spider" technique could be deployed to get the images and download them in the local server, then these images are checked to detect the watermark. This technique suffers from the overload it induces on the network, because of the huge number of files to be downloaded. Alternatively, we use in our system a watermark mobile agent, which is an application that navigates among Internet nodes searching for files to check. As a result, it moves one time only to a node to check all files stored in it. In addition, it has a small size, compared to image files, so it is considered as the best solution for these cases [2].

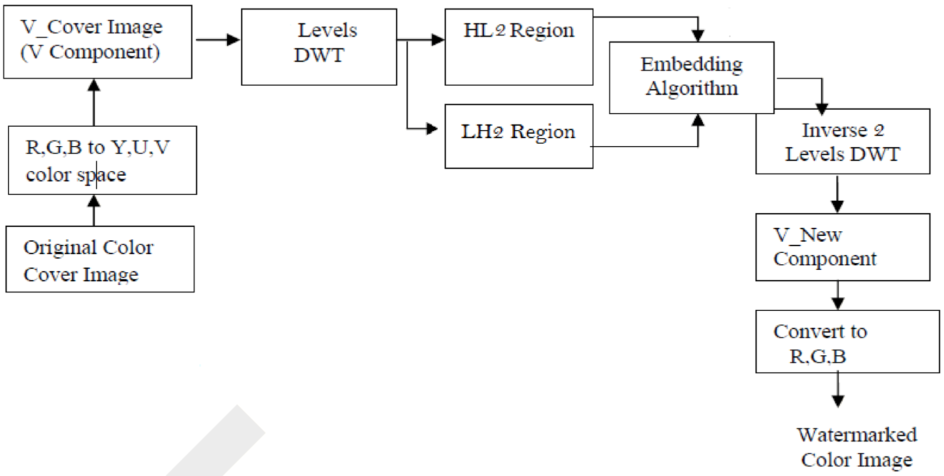

Figure 2.Suggested embedding algorithm steps.

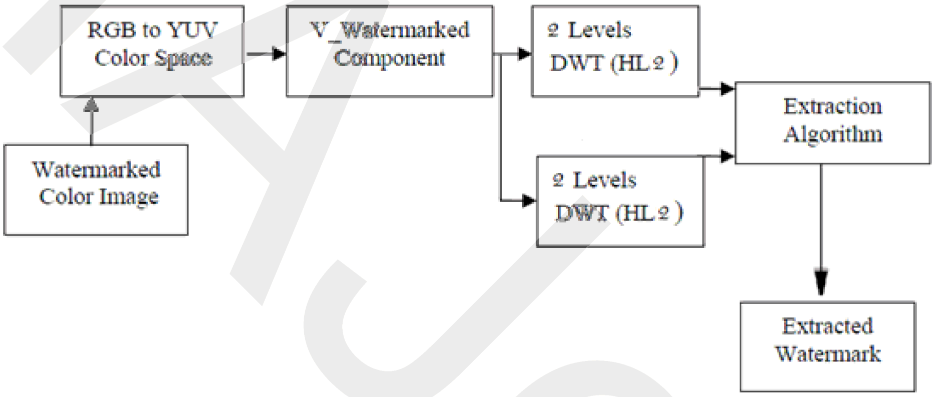

Figure 3..Suggested extraction algorithm steps.

4.5. Agent Implementation

The watermark mobile agent is implemented to be able to work under severe server conditions without influencing its performance. This agent should be executed regardless of the operating system used in the host nodes. The agent must be adapted to long -time work conditions, so it should be able to pause, continue and stop Moreover, it should be unaffected with stopping the server or restarting it. In that case, it should not redo the work from the beginning. Therefore, it has to store the current status, and continue its work according to that status. On the other hand, the agent should react according to notification message received from host server.

In our system, we use Windows services for Windows operating system, and Linux service for Linux operating systems. Both of them depend on the same principle and technique. The service could start the work when system reboots and runs in background while the system is running. We can then control its function, and change its status to “pause, stop, or continue”. It could receive notifications from the host server. Using this technique, the agent exploits its time efficiently without any effects on the work of other users.

4.6. Communication between System Components

We build the system in way that achieves communications despite of different operating systems, platforms and the programming languages used in developing the application. Web services are used to achieve that objective, especially in sending reports. The message is encrypted before sending it to achieve more security.

5. Tests and Results

We test the system on many color images, which are different in size and type, using many random watermarks which have multi lengths without using error correction code (ECC). This ECC could be integrated in future work to improve the results.

We choose strength parameter Alfa –the parameter that is used in embedding equation with value equals “2” to achieve good persistence and transparency.

5.1. Results of Watermarking

Table (1) describes results that compare our algorithm to current existing algorithms.

Notice that the power of the algorithm mentioned in [3] against noise attacks is due to embedding the watermark in low-frequency sub bands, and to the characteristics of the noise signal as it contains high frequency components. As a result, no effect is noticed on watermark.

However, these attacks are the weak point of the method mentioned in [4], because it depends on watermarking in high frequency sub bands, so the watermark will be affected by the noise and low-pass filters.

We noticed too that the suggested approach achieved the best results in most situations compared to these algorithms, because it depends on LH, HL sub bands that are considered as best solution compromising transparency and persistence. In addition, using YUV coding instead of RGB increases the efficacy of our algorithm.

Table 1.The extraction rate of watermarks, when attacked by different means.

Type of attack	[3]	[4]	[9]	[11]	our approach
Gaussian Noise	0.98	0.86	____	1	0.99
Salt & pepper	0.99	0.86	0.93	0.94	0.983
Median filter	____	0.91	0.91	0.87	0.99
Low pass filter	____	0.91	____	___	1
JPEG		0.85	0.91	1	0.98

Table (2) reveals results after applying some attacks on the images.

Table 2. results of applying some types of attack on the images.

Attack	Img0	Img3	Img2	Img4	Img1
G	0.99	1	0.98	1	1
S	0.98	1	0.98	0.99	1
L	1	1	1	1	1
M	0.98	1	0.98	1	1
SG	0.97	1	0.97	1	1
LM	0.98	0.99	0.98	0.99	0.99
SGLM	0.93	0.99	0.95	0.98	0.98

G: Gaussian Noise attack.
S: Salt and pepper noise attack.
L: Low-pass filter attack.
M: Median Filter attack.

Figure 4 describes watermark detection average after applying different attacks on the watermarked images, compared to used approaches, and the Figure 5 describes our algorithm results for different tested images (showed in Annex 1).

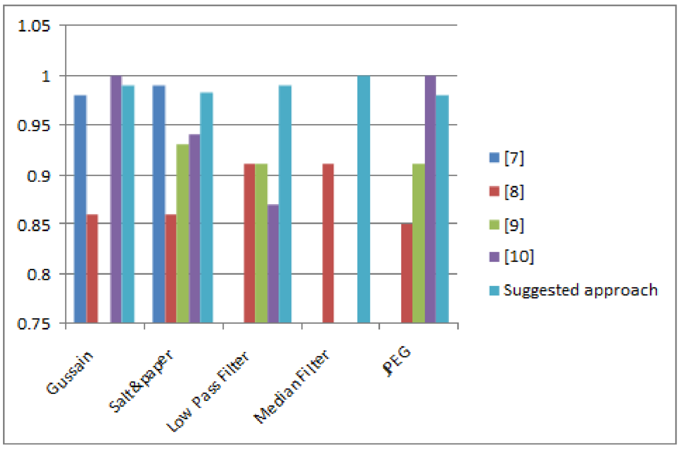

Figure 4. Watermark detection rate after some attacks.

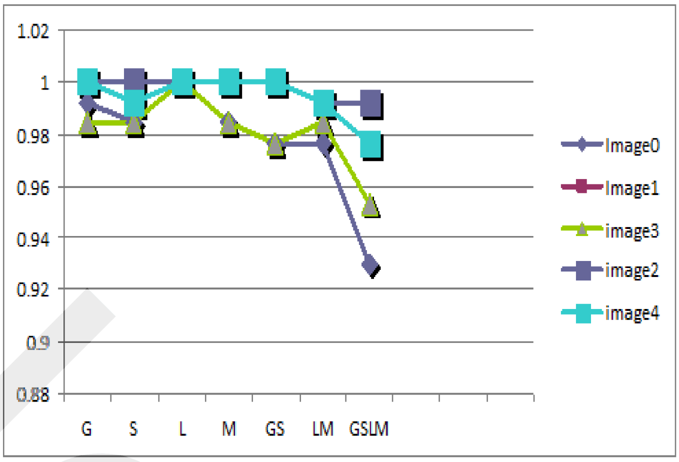

Figure 5. Results of watermark detection for some tested images (suggested approach).

5.2. Results of Tracing

The used agent, instead of network spider, could be implemented efficiently for search process. We can define the complexity indicator by the following relation:

$$\text{Complexity} = \text{files count} * \text{size of every file} \quad (15)$$

In spider case, when the local server does all tasks sequentially, this causes slow execution. Moreover, a direct connection must be provided during the operation.

In watermark agent case in non-blind based system, i.e. a system that needs the original image for extraction process, the complexity indicator could be calculated as:

$$\text{Complexity} = \text{agent navigation count} * (\text{agent size} + \text{image size} \quad (16)$$

Movements count is significantly less than the file count.

To reduce this complexity, we suggest, in our system, using blind detection model. The complexity could then be calculated by:

$$\text{Complexity} = \text{agent navigation count} * \text{agent size} \quad (17)$$

In this case, the work could be distributed on the agency and hosts nodes, so tasks are done in parallel. This means that the system performance will be increased. No need for direct connection with agency in this case, as the agent can do its work offline then contact the agency to send reports.

This model increases system efficiency; especially in congested networks. The agent size is small compared to image size.

In terms of performance, SIFT algorithm represents a bottleneck, because it needs to build scale space that is composed of four octaves, each of them contains five images. The execution requires working with twenty images for each image check.

6. Conclusion and Future Work

In this research, we focused on building an effective algorithm for color images watermarking to protect owners copyrights.

This algorithm depends on scale invariant feature transform to determine references points for embedding and extraction algorithm.

Moreover, we developed a tracing system based on navigating mobile agents among network nodes.

The suggested system could be used regardless of the operating system. Using web services for communication process allows the system to be available for any environment or any programming language used in developing the agency and related nodes. This also enables calling any method behind firewall.

The test results show the algorithm strength and robustness against several attacks. It demonstrates the capability of the tracing system to achieve the expected goals.

In the future, we will work on developing this system to increase its robustness against geometrical attacks, and will build the agent dynamically to enhance its search capabilities.

References

[1] Antoine. J. P, Murenzi. R, and Vandergheynst. P,“Two-dimensional directional wavelets in image processing”, Int. J. Imaging Sys. and Tech., 7:152-165, 1996.

[2] Chess. D, Harrison. C, and Kershenbaum. A, “Mobile agents: Are they a good idea?” *In: Mobile Object Systems: Towards the Programmable Internet*, Springer-Verlag. Lecture Notes in Computer Science No. 1222, 1997.

[3] Gilani. S.A.M, Hameed. K, and Mumtaz. A, “Digital Image Watermarking in the Wavelet Transform Domain”, *International Journal of Applied Science, Engineering and Technology* www.waset.org, 2005

[4] Jiansheng. M, Sukang. L and Xiaomei. T, “A Digital Watermarking Algorithm Based On DCT and DWT”, *Proceedings of the 2009 International Symposium on Web Information Systems and Applications (WISA’09)* Nanchang, P. R. China, 2009,

[5] Lowe. D, “Distinctive image features from scale invariant keypoints”, *International Journal of Computer* Vision 2(60): 91-110, 2004

[6] Memon. N, Sencar. H. T, “Watermarking and ownership problem: a revisit”, *in Proceedings of the 5th ACM workshop on Digital Rights Management, Alexandria, VA, USA, pp. 93-101*, 2005.

[7] Nguyen. T, Strang. G, “*Wavelets and Filter Banks, Wellesley-Cambridge Press*”, 1996.

[8] Seitz. J, "*Digital Watermarking for Digital Media*", University of Cooperative Education Heidenheim, Germany, Information Science Publishing, 2005

[9] Soheili. M. R, “A Robust Digital Image Watermarking Scheme Based on DWT”, *Department of Computer Engineering, Faculty of Engineering, Tarbiat Moallem University*, Tehran, Iran, 2010.

[10] Terzija. N “Robust digital image watermarking algorithms for copyright protection”, *Faculty of Engineering, University of Duisburg-Essen*, 2006.

[11] Terzija. N “Robust digital image watermarking algorithms for copyright protection”, *the University of Duisburg-Essen, to obtain the academic degree of Doctor of Engineering Science*, 2006.

[12] https://en.wikipedia.org/wiki/YUV

Nawar Al-Awa is an associate professor at the faculty of Information Technology Engineering in Damascus University. He obtained the Ph.D. degree in Image processing from National Polytechnic Institute of Grenoble (INPG) in France in 1995, and got his engineering degree in 1991 from the National School for Electronics and Radio-electricity in Grenoble (ENSERG) in France. His research interests include: parallel system design, embedded devices, and image applications.

Ziad Al-Bkhetan received his Bachelor of Education in Information Systems and Software Engineering, Faculty of Information Technology Engineering, Damascus University in year 2011. Now He is an application developer at Alcatel lucent in Jordan.

Annex 1: Some images used in tests

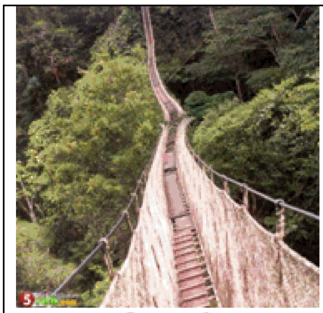

Image 0

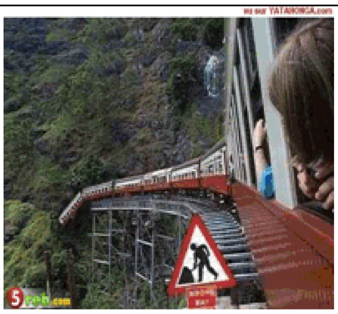

Image 1

Image 2

Image 3

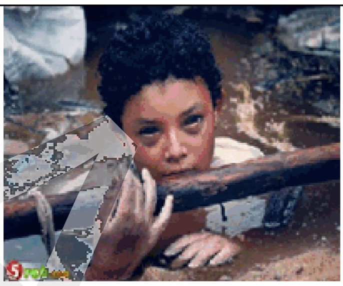

Image 4

Image 5