Authentication of Volume Data Using Wavelet-Based Foveation

M. S. Kankanhalli¹, E.-C. Chang¹, X. Guan², Z. Huang¹, and Y. Wu¹

Abstract. 3D volume data has been increasingly used in many applications. The digital nature of the data allows easy creation, copying and distribution. However, it also allows ease of manipulation which can enable wilful or inadvertent misrepresentation of the content. For an application like medical imaging, this can have serious diagnostic and legal implications. Thus there is a strong need to establish the integrity of a particular volume data-set. We argue that the traditional data authentication mechanisms like digital signatures or cryptographic methods are not very useful in this context due to their extreme fragility. What is required is a method that can detect the integrity for allowable contentpreserving manipulations. We have developed a novel authentication procedure which is robust against benign content manipulation. The volume data can be robustly authenticated under normal operations such as scaling, resampling and additive Gaussian noise. On the other hand, it offers protection against any malefic or unintentional data manipulation which significantly changes the content of the volume data-set. Such manipulations include cropping, changing of voxel values etc. Our method uses segmentation, wavelet-based foveation, and encryption to achieve this. We have implemented the method and tested its robustness for several manipulations.

Keywords: Volume Data, Authentication, Foveation, Wavelets

1 Introduction

3D volume data has been increasingly used in many applications [11]. Medical imaging is one area which generates an enormous amount of volume data. Recently, there has been increasing awareness about the problem of copyright protection of digital images, video and audio [10]. Researchers have started raising concerns about the copyright protection and piracy of 3D data as well. In fact, there is a growing interest in developing digital watermarking techniques



for 3D mesh as well as volume data [15, 9]. However, all these techniques address the copyright problem, not the integrity problem. In this paper, we address the problem of authenticating 3D volume data i.e. verification of the qenuineness of the data-set. For example, given a medical volume data set which shows a medical condition like a tumor, we do not want the patient to fraudulently alter the data-set so that the tumor is removed and thus mis-represent the medical condition to an insurance company. Similarly, we would not like a medical institution to alter the data-set in order to introduce artifacts which represents some abnormality and make a patient go through unnecessary expensive medical procedures. In such situations, preserving and checking the veracity of a dataset assumes tremendous importance. Even for volume data sets which represent art objects or manufactured objects, the accuracy and integrity of the data-set needs to be preserved. Basically, a secure authentication system can prove that no tampering has occurred during situations where the credibility of the volume data may be questioned. In the two hypothetical scenarios present, there is a need to detect that some illegal manipulation has taken place. On the other hand if some allowable modification (like re-sampling of the data-set) is done, it should detect this manipulation but it should indicate that the data-set is still usable.

We propose that this problem can be addressed by use of a *content-based* digital signature which is robust yet effective. The idea is that at the time of data creation (which is through either some physical scanning device like a CT Scanner or through some software), a content-based digital signature associated with the data-set is simultaneously created. For all further authenticity checks, this data-set can be verified against this digital signature. If there is a mis-match, then the data is considered unreliable and it should not be used.

Two solution possibilities naturally arise when considering this problem. One could argue that either traditional general message authentication techniques can be used or perhaps semi-fragile watermarking techniques could be used. We will now argue why neither of this possibility is applicable for 3D volume data.

Traditional message authentication techniques like hashing-based digital signatures or cryptographic authentication [20] cannot be used because of their extreme fragility. These techniques do not tolerate flipping of even one bit of information of a message. For example, we could use the traditional message digest based digital signature for a volume data-set. Even if one least significant bit of a voxel is changed, the authentication procedure will flag this data-set as unreliable. However, for volume data, certain operations such as scaling, resampling etc. are valid operations in which cases the manipulations are benign. They are not intended to change the significant content of the data-set and thus they do not impact the integrity of the data. One example is the content-based digital signature proposed in [16] which used the histogram of divided data blocks as the content to be hashed. If the voxel values are uniformly deduced by one unit, or more generally, a Gaussian noise with a non-zero mean is added, this signature may fail to authenticate the data. Therefore, we need a novel digital signature method which allows content-preserving manipulations. In other words,

the digital signature must authenticate the data set for such cases. However, if somebody really tampers with the content, e.g. crops out the tumor region, then the digital signature should indicate that data-set has been tampered with and thus is unreliable. Thus traditional digital signatures are not useful but *robust authentication* (robustness to allowable manipulations) is required.

The second possibility is the use of semi-fragile watermarks for the purpose of authentication [13]. Many techniques have been developed for 2D digital image data which could perhaps be adapted for 3D volume data. Unfortunately, this is not possible for two reasons. Firstly, the image watermarks usually exploit the characteristics of the human visual system (HVS) in order to hide the secondary watermark information. In case of volume data, the HVS cannot be exploited because we can only visualize the 3D volume data through surface and volume rendering. Secondly, there is an even more serious problem. All watermarking techniques involve the modification of the voxel values for the purpose of embedding the watermark. However, for the case of volume data (and especially related to medical imaging), distortion of the voxel values is not allowed. Even if small perturbations in the voxel values were allowed, there is no watermarking method which can provably bound the distortion of the voxel values. While the Parseval's theorem [6] can guarantee the bounding of the overall watermark signal energy, simultaneously limiting the maximum distortion level in the spatial domain and frequency domain appears to be very difficult. Therefore, watermarking techniques are also not useful for authenticating 3D volume data.

In this paper, we present a new technique for authenticating 3D volume data using a robust content-based digital signature. This signature is derived from the significant features of volume data so that if any of these features are altered significantly, the signature will not match the data-set. The term content-based refers to the fact the important features of the data (whose integrity we are interested in certifying) should be somehow incorporated into the digital signature. The rationale being that if some important content feature is deleted/modified/added, then the digital signature should not match the doctored data-set. The term robust refers to the fact that any manipulation which does not change the significant features should not affect the veracity of the signature. For such benign operations, the digital signature should indeed authenticate the data-set. Common types of operations on volume data-sets are scaling, thresholding, cropping, cut-and-replace a sub-volume, filtering, addition/removal of noise and affine transformations. As long as these operations do not change the content features, they are considered benign. We use a novel wavelet-based foveation technique to accurately and succinctly capture the significant content features. Moreover, the scheme allows a flexible threshold to be set which can determine the extent of the manipulations which can be considered benign.

2 Overview of the Technique

We will now provide an overall description of the method for generating the robust content-based digital signature and the method for authenticating a volume data-set using this digital signature. For the generation of the digital signature, the following steps are required:

- 1. Feature extraction: The basic idea here is to capture the essential features which need to be preserved for authentication. Since the size of a 3D volume data-set is huge, this also allows us to create a compact key derived from the important features. The process is done in three steps:
 - (a) Volume segmentation: The voxels of the input 3D volume data are separated into two classes the significant "foreground" and the relatively less important "background". While we present a method for doing the segmentation in this paper, we recognize that different types of data-sets need their own specialized segmentation technique. Our authentication method is flexible in the sense that it does not really depend on the particular details of the segmentation algorithm used. If required, this part can be customized either for a particular application domain or for an individual volume data-set.
 - (b) Selection of key voxels: In general, the number of voxels in the foreground is quite large. To reduce the amount of data, a few "key voxels" are chosen for the purpose of data reduction.
 - (c) Wavelet-based foveation: To make sure that important content throughout the foreground is captured, we apply the foveation technique which is basically a space-variant filtering technique. We believe it is very important to use this since it summarizes all the important content throughout the foreground with the key voxels as the foci. Thus all significant features are compactly captured. Additionally since it is a many-to-one mapping, it offers security. Thus, this information can be used as a key.
- 2. Encryption: For additional security, public-key cryptography [20] is utilized to encrypt the key derived in the previous step. Basically, the secret key of the owner of the volume data is used to encrypt the feature key obtained. For the purpose of authentication, the public-key of the owner can be used to decrypt this information and the feature key can be thus recovered. Since this step is well-understood, we will not discuss it further in this paper.

For authenticating a particular volume data-set, the following steps are performed:

- 1. Affine transformation parameters recovery: Since, one of the benign manipulations could be the affine transformation of the volume, the transform parameters are computed first.
- 2. Matching: The content features of the transformed volume are compared with the content features of the original data-set (obtained from the digital signature after decryption using the owner's public key). A match value between the original features and the transformed volume features is computed. If this match value exceeds a certain threshold, then the volume is certified as genuine else it is considered untrustworthy.

3 Creation of the Content-Based Signature

Given a volume data-set, we first extract the feature points (key voxels) after performing segmentation. The number of the feature points must be small in order to guarantee the acceptable small size of the signature. Based on the extracted feature points, a weighted norm is selected. The volume is then lossily compressed using this weighted norm as the measure. The description of the weighted norm and the compressed data are encrypted using public-key cryptography and they constitute the robust content-based digital signature associated with this volume data-set.

3.1 Feature Extraction of Volume Data

In this subsection, we describe the process of extracting from the original volume data a small number of voxel groups that represent the important information (features). It consists of two steps: volume segmentation and key voxel sampling. Volume segmentation is to identify and demarcate into foreground/background the voxel in the original volume data. Note that the foreground voxels can belong to different sub-categories (like bone, skin, soft tissues, etc.). This results in the segments (a connected sub-volume) with each one representing an important feature of the data. It is similar to image segmentation, a common technique used in computer vision. Usually the number of voxels in each segment is too large to be directly used for the signature. Thus, key voxel sampling is used to derive a few key voxels from each segment for the foveation process which effectively summarizes the significant content of the data-set, and will be detailed in the next subsection.

We propose a segmentation method based on the voxel value analysis and bounding box information of the isosurfaces. It can be summarized as follows. We assume that the voxel values are scalar for the ease of description.

- 1. Partition of the voxel values by data analysis. First, all the voxel values are sorted in the non-descending order. Second, partition the sorted list using the threshold value. The threshold value is specified by the user in our current implementation. However, heuristics can be applied if the domain knowledge is known for the particular class of volume data. For many volume data sets, the density values of significant content components are distinguishable even though the voxels representing them are closely connected to each other. Sometimes, they may perhaps even have similar voxel values in which case domain knowledge could be utilized for distinguishing them. For example, human CT/MRI volumes can be partitioned by using the density values as well as anatomical knowledge.
- 2. Isosurfacing. From the partition, we derive a set of voxel values that partition different parts. These voxel values are used to derive the same number of sets of the isosurfaces.
- 3. Segmentation. One segment of voxels can be formed if they are bounded as a closed sub-volume by (1) one isosurface, (2) several isosurfaces, or (3) one or several isosurfaces with the one or several border planes of the volume. It can be efficiently derived using the scan conversion algorithm, an extension of the standard scanline

- algorithm used in the rasterization and hidden-surface elimination, to derive and accumulate the intervals bounded by the isosurfaces and border planes iteratively.
- 4. Feature extraction. It is a process of selection of key voxels. A 3D Gaussian mask is applied on the volume several times as lowpass filtering. Due to the large size of volume data, we simulate the 3D Gaussian filtering as a windowed lowpass filtering dimension by dimension. In the highly blurred resulting volume, the key voxels are chosen to be local maximum voxels which are above a predefined threshold. The key voxels are then used as the input to the foveation procedure.

If the size of volume data is N^3 , sorting in the first step takes $O(N^3 \log N)$ time. It is $O(N^3 + \log h)$ for isosurfacing in step 2 where h is the number of different extreme values (min or max) [5]. It is $O(N^3)$ for the scan conversion in step 3. So the overall time complexity is $O(N^3 \log N)$. One example as a result of this procedure is shown in Fig 1.

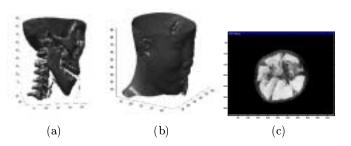


Fig. 1. Examples of volume segmentation results: (a) the skull bone, (b) the skull muscle and soft tissue, and (c) the internal part of a tomato.

3.2 Content-based Weighted Norm

We now briefly present the idea behind summarizing the volume using a space-variant wavelet based filter. The basic idea is to summarize and compress the important content information. Most 2D/3D imaging systems use a norm (usually the Euclidean 2-norm) to measure their performance. However, the 2-norm treats each pixel/voxel equally. However, in most real-life data, it is possible to determine some regions that are more interesting for the application at hand. For example, through feature detection, we can find significant voxels in a given data-set. In such cases, a weighted norm is more appropriate. The weighted norm $\|\cdot\|_w$ for the volume V(x,y,z) with a weighting function w is given by:

$$||V||_w^2 = \sum_{x,y,z} w(x,y,z)V(x,y,z)^2,$$

where $w(\cdot,\cdot,\cdot)$ is the weighting function. In our authentication system, a content-based weighted norm is used for measurement of the distortion caused by allowable and illegal operations. The weight of each voxel is determined through a

combination of segmentation and feature detection schemes. In the digital signature creation process, the original volume data is lossily compressed under the weighted norm. The highly compressed data S, together with the description of weighting function W, forms the signature (S,W). This signature can then be further encrypted. Because the description of the weighting function is part of the signature, to satisfy compactness, w cannot contain the full information of the original data set. We now describe in detail the whole procedure.

3.3 Wavelet-based Foveation Technique

Our visual system has a space-variant nature where the resolution is high in a point (fovea) but falling off towards the peripheral [17]. This distribution of resolution provides a fast and simple way of reducing information in the visual field, without sacrificing the size of the visual field and the resolution around the fovea. As the biological visual system is highly effective, this space-variant nature has inspired the design of many computer vision systems which resembles the biological foveated vision [3,1,18], video conferencing [2,7], and visualization systems [12].

The foveated volume is obtained from a uniform resolution volume through a space-variant smoothing process where the width of the smoothing function is small near the fovea but gradually increases towards the peripheral. The process of going from a uniform volume to a foveated volume is known as foveation. The foveation of a function $V: \mathbf{R}^d \to \mathbf{R}$ is determined by a smoothing function $g: \mathbf{R}^d \to \mathbf{R}$, and a weight function $w: \mathbf{R}^d \to \mathbf{R}_{>0}$.

$$(TV)(\mathbf{x}) := \int_{\mathbf{R}^d} V(\mathbf{t}) w(\mathbf{x}) g(w(\mathbf{x}) \| \mathbf{t} - \mathbf{x} \|_2) d\mathbf{t}. \tag{1}$$

The weighting function w depends upon three parameters and takes the form

$$w(\mathbf{x}) = \left(\alpha \|\mathbf{x} - \gamma\|^d + \beta\right)^{-1}.$$
 (2)

We call α the rate as it determines how fast resolution falling off, call γ the fovea as it determines the point of highest resolution, and call β the foveal resolution as it determines the resolution at the fovea. Both α and β are nonnegative and the smoothing function g is normalized so that $\int_{-\infty}^{\infty} g(\mathbf{x}) d\mathbf{x} = 1$. In general, we could replace the weighting function by any non-negative function. This generalization is useful when we are interested in volumes with multiple foveae. Given two weighting functions w_1, w_2 , the blended w_3 is

$$w_3(\mathbf{x}) = \max\{w_1(\mathbf{x}), w_2(\mathbf{x})\}. \tag{3}$$

Foveated volumes can also be treated as the approximation of an volume using a fixed number of bit, using a weighted norm as the underlying measure. This weighted norm can be derived from (1) and has the form,

$$||V||_{w} = \int_{\mathbf{R}^{d}} \frac{V(\mathbf{x})}{w(\mathbf{x})} d\mathbf{x}, \tag{4}$$

where the weighting function w is the function in (2).

Wavelet bases have important applications in mathematics and signal processing due to their ability to build sparse representation for large classes of functions and signal [14]. It is a natural choice for foveated volume due to their locality in space and frequency. Interesting, the choice of the weighting function (2) gives a self-similarity across scales [4], which is illustrated in Fig 2. This property leads to a simple but fast extraction algorithm [4].

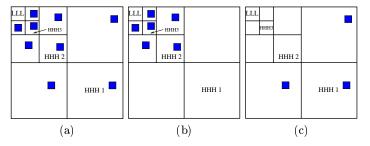


Fig. 2. Allowable Lowpass Filtering: (a) Original wavelet coefficients (C_w) ; (b) After allowable lowpass filtering (C'_w) ; (c) Remaining coefficients (C_w^*) . $C_w - C'_w = C_w^*$.

3.4 Extracting the Coefficients

Recall that the first part of the signature (S,W) is the highly compressed volume. To obtain S, one could first compute the foveation (1) with respect to the multi-foveae weighting function, and then compress the foveated volume using a known lossy or lossless compression technique for uniform volumes. Because computing (1) directly is computational intensive, we use the approximation (5).

$$(T^{\text{fov}}I) \approx \text{IDWT}(M \text{ DWT}(I)).$$
 (5)

In our implementation, S is extracted from the volume by quantizing the wavelet coefficients M DWT(I), followed by a lossless compression using gzip. For an intuitive illustration, we use a 2D image to show its compression result (Fig 3). The (S, W) can then be encrypted and be stored as the digital signature for that image.

Note that gzip is a general lossless compression tool, which does not exploit properties of volumes, especially the coherence of wavelet coefficients across space and scale. Thus it is not the best technique for our application. A possible improvement can be done by incorporating the well-known zero-tree algorithm [19] into our scheme.

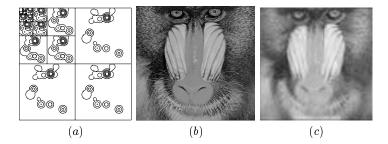


Fig. 3. (a) The mask M for the weighting function. (b) The original image (262Kbytes). (c) The compressed image (4Kbytes) using the mask M.

4 Implementation and Experiment Results

The first phase of the authentication process is detection of the allowable affine translation applied to the volume data-set. For consistency, the norm used in the detection is the weighted norm whose weighting function is part of the signature. That is, the detection finds the affine transformation T_{\min} such that $h_0 = \|T_{\min}(S) - V\|_w$ is minimum. Through our preliminary experiment, we find that such T_{\min} can be accurately determined for translation and rotation. In the rest of this section, we assume that no affine transformation has been applied to the data-set.

In the second phase of authentication, the similarity value $h_0 = ||S - V||_w$ is compared with a predetermined threshold A_0H , where A_0 is a normalizing factor depending only on the size and mean of the volume data-set. If h is smaller, then the volume is declared to be authenticated. Otherwise, it is rejected and hence considered unreliable. The choice of the H depends on the level of allowable attacks. It can be determined analytically by assuming a certain distribution on the voxel, or through experiment conducted a-prior to the signature creation. In our experiments, we choose H = 0.08, which is analytically determined by assuming that the allowable low-pass will filter out only the first level wavelet coefficients as illustrated in Figure 2.

We did experiments on two volume data sets with 256 gray levels, SKULL $(64 \times 64 \times 64)$ and TOMATO $(128 \times 128 \times 64)$. In the selection of key voxels, we used a windowed lowpass filtering for five times with the window size 9 and the threshold 1.5. The resulting numbers of key voxels are 25 for SKULL and 124 for TOMATO. The sizes of the signatures are 8K and 19K bytes respectively.

Five experiments were done with these two volume data sets. The first three experiments examine the signature robustness under global manipulation like low-pass filtering, sharpening, and lossy compression, whereas the last two experiments consider local manipulation like cropping and localized modification. In the first experiment, the volume data-sets are subjected to low pass filtering. The low filtering is achieved by a rectangle window. From Figure 4(a), the signature remains authentic even under a 7x7x7 rectangular filtering. In the second

experiment, Figure 4(b), Gaussian white noise is added. To test the robustness of our signature under lossy compression, we applied zero-thresholding to the volume data-set. That is, given a threshold T, all wavelet coefficient C satisfying |C| < T are replaced by zeros. The results for different T is shown in Figure 5.

Figure 6 (a) shows the robustness after the voxels in the center region are replaced by zeros, and Figure 6(b) shows the robustness after the volume is cropped.

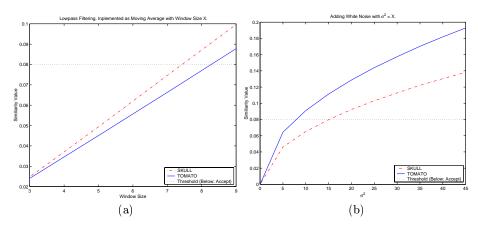


Fig. 4. Results for Lowpass Filtering (a) and Addition of White Noise (b)

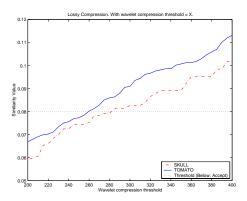


Fig. 5. Results for Lossy Compression

Modification of information of a volume data can take three forms. One is to modify a particular part of the voxel values to other values, e.g., set all voxel values to zero (Fig 6 (a)). The second is to crop the volume to a subset of the

original one (Fig 6 (b)). Removal is acceptable as long as important information remains. Since our method of important feature extraction does not aim at a particular region of the volume, it is enough to give false-signature alarm when too much information has been removed. This can be seen in Fig 6. In real-world applications, users can define the regions-of-interest for feature extraction, for example, tumors or abnormal bones. The third modification is the addition of previously non-existent content feature. This is handled in a manner similar to the one for the removal case. Thus, with the proposed robust digital signature scheme, the signature will match only when all (and no more) regions-of-interest can be detected.

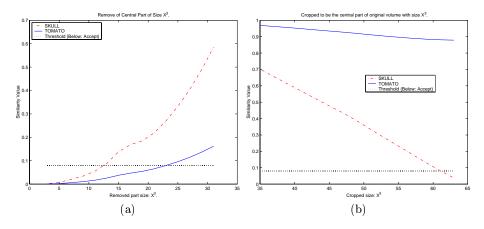


Fig. 6. Results for Removal (a) and Cropping (b)

5 Conclusion

We have described a novel robust content-based authentication technique for volume data. The technique uses segmentation followed by key voxel selection which are used as fovea for a wavelet-based foveation procedure to derive the content-based key for the volume data-set. This key is then encrypted using public-key cryptography and used as a robust digital signature. For authenticating a questionable volume data-set, the affine transformation parameters are first determined and feature extraction is done for the transformed volume. The feature for the transformed volume is then matched against the feature values in the original digital signature to determine whether the volume data is reliable or not. The method has been implemented and tested against various manipulations. The experimental results show that this is a very promising approach.

Our future work is to come up with a reliable volume authentication technique which can be incorporated into all types of scanners (like CT Scanners and MRI devices) in order to make them trustworthy [8].

The research was partly supported by NUS R-252-000-090-112. We also thank the EGMM 2001 reviewers for the helpful comments.

References

- J. Aloimonos, I. Weiss and A. Bandyopadhyay. Active Vision. st International Conference on Computer Vision London, England, 35-54, 1987.
- A. Basu and K.J. Wiebe. Videoconferencing using spatially varying sensing with multiple and moving fovea. *IEEE Trans. on Systems, Man and Cybernetics*, 28(2):137-148, 1998.
- 3. P.J. Burt. Smart sensing within a pyramid vision machine. *Proceedings of the IEEE*, 76(8):1006–1015, 1988.
- 4. E.-C. Chang, S. Mallat, and C. Yap. Wavelet foveation. *Journal of Applied and Computational Harmonic Analysis*, 9(3):312-335, 2000.
- P. Cignoni, C. Montani, E. Puppo, and R. Scopigno. Optimal Isosurface Extraction from Irregular Volume Data. Proceedings of IEEE Symposium on Volume Visualization. 1996, 31-38.
- 6. D.J. DeFatta, J.G. Lucas and W.S. Hodgkiss. Digital Signal Processing: A System Design Approach. *John Wiley & Sons*. New York, Second Edition, 1988.
- A. Eleftheriadis and A. Jacquin. Automatic face location detection and tracking for model-assisted coding of video teleconferencing sequences at low bit rates. Signal Processing: Image Communication. Vol. 7, No. 3, 231-248, 1995.
- 8. G.L. Friedman. The Trustworthy Digital Camera: Restoring Credibility to the Photographic Image. *IEEE Transactions on Consumer Electronics*. Vol. 39, No. 4, 905-910, November 1993.
- 9. X. Guan, Y. Wu, M.S. Kankanhalli, and Z. Huang. Invisible Watermarking of Volume Data Using Wavelet Transform. *International Conference on Multimedia Modeling MMM2000*. November 2000, Japan, 153-166.
- F. Hartung and M. Kutter. Multimedia Watermarking Techniques. Proceedings of the IEEE. vol. 87, no. 7, July 1999, 1079-1107.
- A. Kaufman, D. Cohen and R. Yagel. Volume Graphics. IEEE Computer. vol. 26, no. 7, July 1993, 51-64.
- 12. M. Levoy and R. Whitaker Gaze-Directed Volume Rendering. Computer Graphics. Vol. 24, No. 2, 217-223, 1990.
- E.T. Lin and E.J. Delp. A Review of Fragile Image Watermarks. Proc. Multimedia and Security Workshop (ACM MM'99). Orlando, 25-29, October 1999.
- 14. S. Mallat. A Wavelet Tour of Signal Processing. Academic Press. 1998.
- 15. E. Praun, H. Hoppe, A. Finkelstein. Robust Mesh Watermarking. Computer Graphics ACM SIGGRAPH '99 Proceedings. 69-76, September 1999.
- 16. M. Schneider and S.-F. Chang. A robust content based digital signature for image authentication. *Proceedings of ICIP*. Volume: 3, pages: 227-230, 1996.
- 17. E.L. Schwartz. Topographical mapping in primate visual cortex: history, anatomy, and computation. In D.H.Kelly, editor, *Visual Science and Engineering: models and applications*, pages 293–360. Marcell Dekker, Inc, New York, 1994.
- 18. E.L. Schwartz, D.N. Greve and G. Bonmassar. Space-variant active vision: Definition, Overview and Examples *Neural Networks*. Vol. 8, No. 7-8, 1297-1308, 1995.
- 19. J.M. Shapiro. Embedded image coding using zerotrees of wavelet coefficients. *IEEE Trans. on Signal Processing.* vol. 41, no. 12, December 1993, 3445-3462.
- W. Stallings. Cryptography and Network Security: Principles and Practice. Prentice-Hall Inc. Second Edition, 1998.