Integrity Preservation and Privacy Protection for Digital Medical Images

M.Krishna Rani Dr.S.Bhargavi
IV Semester M.Tech (DCN) SJCIT Chickballapur Karnataka India

Abstract- In medical treatments, the integrity of the medical images, the authenticity of corresponding medical records, and the privacy of the patients are crucial requirements for the diagnosis of patient's disease. Therefore, how to retain the validity between images and medical records is an important topic for researches and real applications. In this project, we apply the concept and implementation in digital rights management (DRM) systems, and propose a functional scheme for the above-mentioned goals. We employ the reversible data hiding scheme, a newly developed branch in DRM researches, for combating the goals. With the term of reversibility, it means that data, including patients' private information and the diagnosis data, can be hidden into the medical image by some means. Later on, the medical image containing data might be retrieved while necessary, and both the original image and the hidden data can be perfectly recovered. Data can be authenticated for enhanced privacy protection.

***Index Terms* – Reversibilty,Histogram, DRM,Quad Tree Decomposition**

I. INTRODUCTION

The development of medical instruments, it gets much easier for the medical doctors to make diagnoses of the patients' diseases by using medical images. Due to the digital nature and enormous amount of medical images, several issues may arise. First of all, the patients' privacies need to be preserved. Therefore, embedding secret data into the medical images would be one of the useful methods for protecting the privacies. Next, because external data are hidden into the original image, some alterations are supposed to be induced. After data embedding, the output image should be as similar as its original counterpart, and medical doctors may lead to proper treatment by using the images with hidden data when necessary. Thus, how to preserve the integrity of the medical image is another important issue. And thirdly, due to the vast amount of images, when the medical doctor needs to retrieve Patient #1's images along with his/her medical records, Patient #2's images may unexpectedly be obtained by the doctor even though this kind of probability is minute. And this means that the correlation between the medical images and medical records of the same patient should be authenticated, and then the medical doctor can proceed with the diagnosis procedures.

With the term of "reversibility," it means that data, including patients' private information and the diagnosis data, can be hidden into the medical image by some means developed by ourselves. Later on, the medical image containing data might be retrieved by medical doctors while necessary, and both the original image and the hidden data can be perfectly recovered with the algorithm corresponding to the embedding scheme. For developing a proper algorithm for reversible data hiding, the output image and its corresponding original should be as similar as possible, or we may refer to the imperceptibility of the image. In addition, the size of the data, or the capacity of the algorithm, is supposed to be as large as possible in order to reside the medical records. Finally, the most important thing is that the data hiding algorithm should be reversible.

II. OVERVIEW

Digital watermarking or steganography, a hardly noticeable noise-like signal can be embedded into a digital medium, such as an image, audio, or video data, to protect it from illicit use and alteration, to authenticate its content and origin, or to enhance its value and enrich its information content. Unlike metadata, which is often appended to the digital file, a watermark is bound into the fabric of the media and cannot be removed or destroyed easily. The watermarking process usually introduces irreversible degradation of the original medium. Although this degradation is slight, it may not be acceptable to some applications, such as military uses, medical uses, and multimedia archiving of valuable original works. However, these applications may tolerate the addition of such noise if the watermark can be removed after decoding. This removal restores the original medium without any reference to information beyond what is available in the watermarked medium itself. The amount of tolerance to the watermark varies from one application to another. Although some applications desire high signal-to-noise ratio (SNR), many others accept low SNR. For example, Mintzer et al. used completely visible, but removable, patterns (low SNR) with their digital library application to promote image sale on the Internet. The user can download a free marked image as a "teaser," but he or she must purchase a "vaccine" program to restore a high quality image from the teaser image.

- TECHNIQUES USED

There are two major techniques for making reversible data hiding possible. These are: (1) the histogram-based scheme and (2) the difference expansion (DE) scheme . We will briefly describe the two techniques, and make initial comparisons between the two.

- ALGORITHEM USED

The histogram-based scheme considers the global characteristics of original image. Part of the histogram is intentionally altered to perform data hiding. The luminance with the maximal and no occurrences in histogram are labeled as max point and zero point, respectively. The luminance values of "max" and "zero" points, each is represented by 8 bits, are treated as side information. Hence, a total of 16 bits should be transmitted to the receiver for data extraction. The range of luminance values between max and zero points would be altered for data hiding. In the region between max and zero points recorded, luminance values are all increased by 1. For the embedding of binary watermark, if the watermark bit is '1' the luminance value is increased by 1; if the watermark bit is '0,' it is decreased by 1.

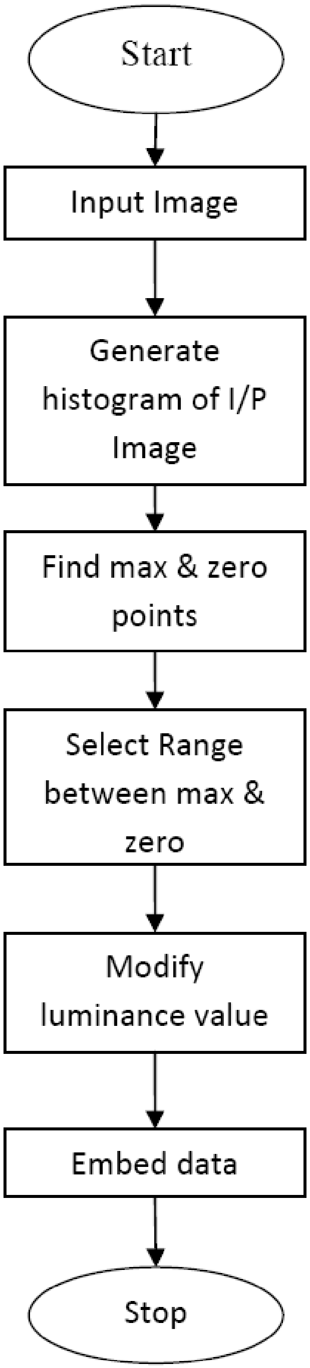


Fig 1.Histogram based Data Hiding

Image imperceptibility can be guaranteed to be more than 48.13 dB in PSNR. The only drawback is that the capacity is constrained by the occurrences of max point, and how to get the increased capacity is another important issue for researches.

On the other hand, for the difference expansion scheme, only the local characteristics are considered. The difference of luminance values between neighboring pixels are intentionally altered for the data hiding purposes. Suppose that the luminance values of a pair of pixels can be represented by (x, y). With the concepts in wavelet transform, the average value l, denoting the low frequency, and difference value h, denoting the high frequency, between the two can be represented by $l = \left\lfloor \frac{x+y}{2} \right\rfloor$ and $h = x - y$. The notation $\lfloor . \rfloor$ denotes the floor function. Suppose that the data bit for embedding is b. We keep the l value the same, and alter the h value by $h' = 2. h + b$. For the output pixel pair (x',y'), the two pixel values can be calculated by $x' = l + \left\lfloor \frac{h'+1}{2} \right\rfloor$ and $y' = l - \left\lfloor \frac{h'}{2} \right\rfloor$. If we consider practical cases that x' or y' may lie outside the range between 0 and 255, such locations are unsuitable for making data hiding possible, and should be recorded into the side information for data extraction.

We take conventional test image, for instance, grey-level image with the size of 512X512 and 8 bit/pixel, for making preliminary comparisons. Both the output image quality and the embedding capacity are compared.

Due to the limited embedding capacity, in addition to applying the conventional schemes, we take the inherent characteristics into account to increase the embedding capacity, and to reduce the side information produced. We simply divide the original image into smaller blocks, or 256X256, 128X128, 64X 64, and 32X32 blocks, and treat each of the blocks independently as a small image. By doing so, the embedding capacity for both techniques can be greatly enhances, and we can clearly see that the histogram-based scheme performs better than its DE counterpart.

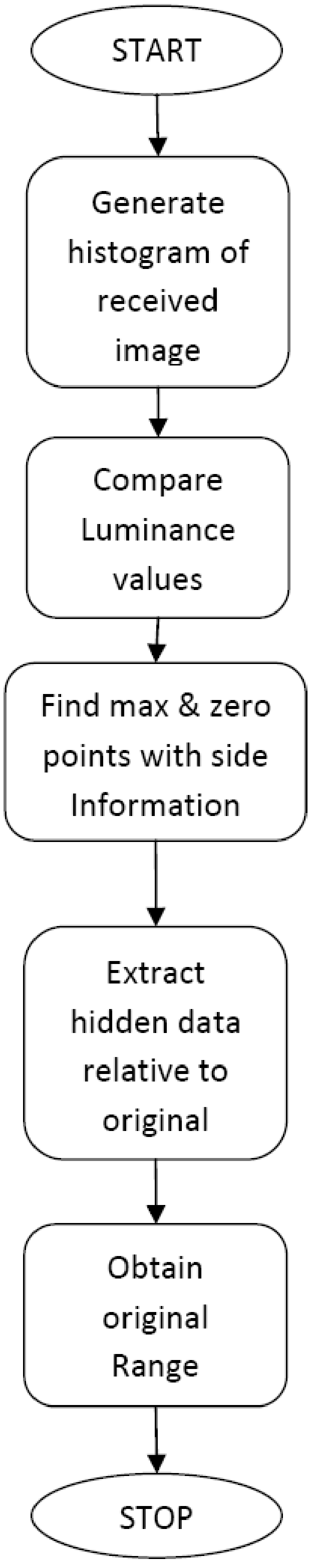

START
Generate histogram of received image
Compare Luminance values
Find max & zero points with side Information
Extract hidden data relative to original
Obtain original Range
STOP

Fig 2.Extraction of Data From Histogram Data Hiding

- COMPARISION

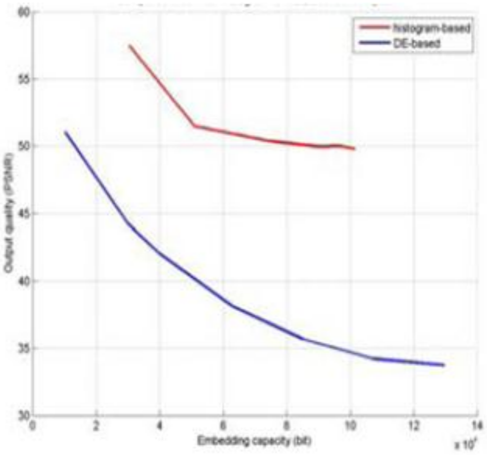
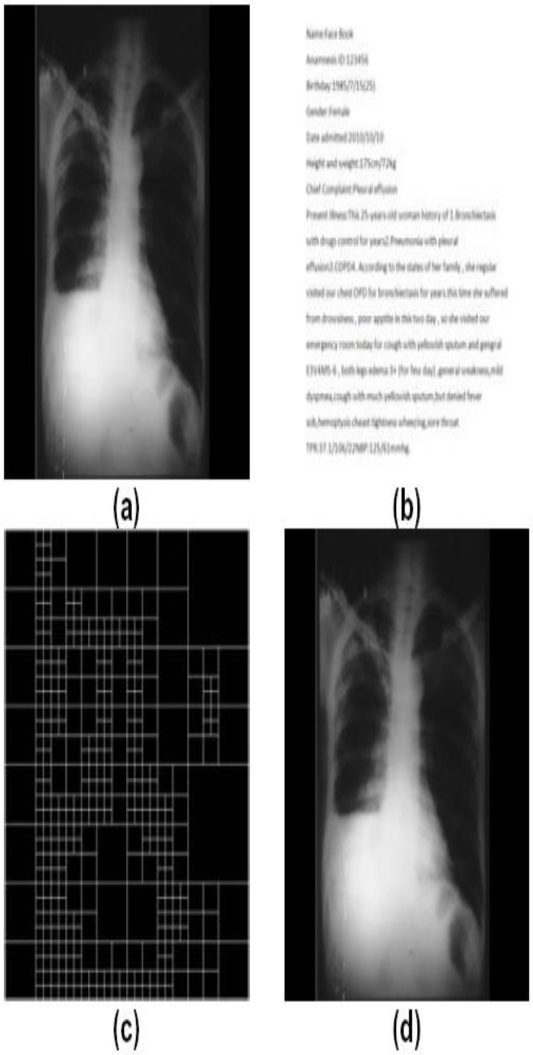


Fig 3: Comparision between DE based and Histogram based Data hiding

In Figure 3, we can clearly see that with the histogram based technique, data capacity can reach 101475 bits for the image with the size of 512X512, meaning that 0.3871 bit/pixel can be obtained with somewhat degraded output quality. Obviously, because the inherent characteristics between natural images (the test images for instance) and the medical images are quite different, we may apply the histogram-based reversible data hiding scheme into medical images due to the abundant amount of data capacity. We note that in order to protect the privacy of the real patient, the patient' name and ID are intentionally changed, while all the other data are real information relating to the medical records and diagnoses. The file size and the capacity should be abundant to reside such the medical records. Another reason for choosing the histogram-based reversible data hiding scheme is that the side information has only two bytes. In comparison to the DE-based scheme, for which the positions that are unsuitable for data embedding should be recorded, we can lead to the result that the histogram-based scheme would be a better choice for practical implementations..

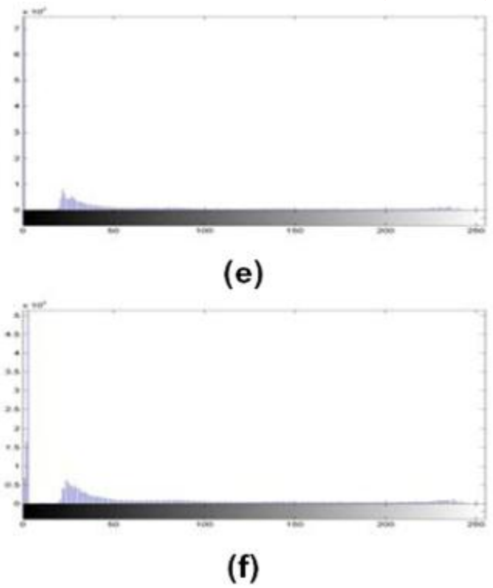
III RESULT AND DISCUSSION

The image in Figure 4(a), with the size of 512×512 , is employed, and binary data in Figure 4(b) are embedded.We apply quadtree decomposition in Figure 4(c), and the output image is depicted in Figure 4(d) with the image quality of 50.36 dB. Also, Figure 4(e) and (f) denote the original and output histograms, respectively. We can see that both the histograms are similar



Name:Face Book
Anamnesis ID:123456
Birthday:1985/7/15(25)
Gender:Female
Date admitted:2010/10/10
Height and weight:175cm/72kg
Chief Complaint:Pleural effusion
Present Illness:This 25 years old woman history of 1.Bronchiectasis
with drugs control for years2.Pneumonia with pleural
effusion3.COPD4. According to the states of her family , she regular
visited our chest OPD for bronchiectasis for years.this time she suffered
from drowsiness , poor appetite in this two day , so she visited our
emergency room today for cough with yellowish sputum and general
E3V4M5-6 , both legs edema 3+ (for few day) ,general weakness,mild
dyspnea,cough with much yellowish sputum,but denied fever
sob,hemoptysis chest tightness wheezing,sore throat
TPR:37.1/106/22NBP:125/61mmHg

Fig 4 (a) Original Image (b) Selected Record Information (c) Quad Tree Decomposition (d) Output Image



(e)

(f)

Fig 4 (e) Original Histogram (f) Output Histogram

- TABLE COMPARISION

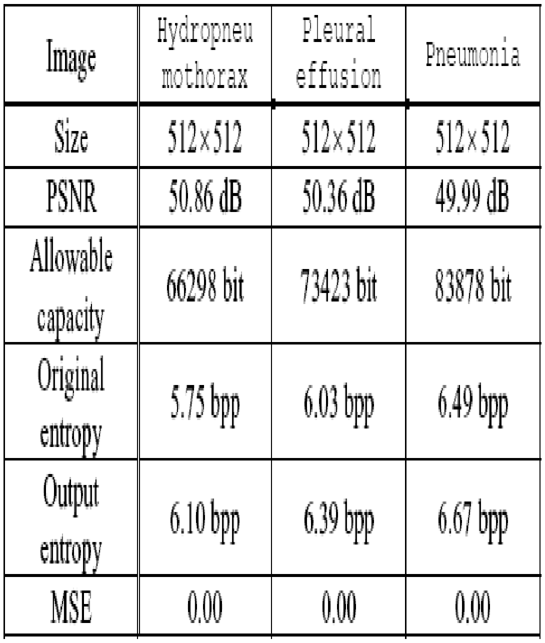


Image	Hydropneumothorax	Pleural effusion	Pneumonia
Size	512×512	512×512	512×512
PSNR	50.86 dB	50.36 dB	49.99 dB
Allowable capacity	66298 bit	73423 bit	83878 bit
Original entropy	5.75 bpp	6.03 bpp	6.49 bpp
Output entropy	6.10 bpp	6.39 bpp	6.67 bpp
MSE	0.00	0.00	0.00

The comparisons of performances among three different medical images with the size of 512×512 the PSNR values are high enough to make the hidden data imperceptible. And the allowable capacities are much larger than conventional values shown in literature. After data embedding, the entropies of the output image get somewhat increased to compare with their original counterpart due to data embedding. Finally, for validating the reversibility of our algorithm, we calculate the mean squared error (MSE) between the original image and the image after data extraction. We can see that all the MSE values are 0.0, meaning that both the input at the encoder and the output at the decoder are identical. This proves the applicability of our implementation.

IV CONCLUSION

In this paper, we discussed the integrity preservation of medical image, and the data protection and authentication for the medical data. With reversible data hiding, the medical images and medical records of the same patient can be authenticated, and proper treatment can be performed by medical doctors. The proposed scheme is suitable for X-ray or CT medical images, and it has the potential to be integrated into the databases for managing the medical images in the hospital,

- SOFTWARE REQUIREMENT

This is implement using matlab and modeling is done using simulink.

V REFERENCES

[1] H. C. Huang, and W. C. Fang, “Metadata-based image watermarking for copyright protection,” Simulation Modelling Practice and Theory, vol. 18,pp. 436–445, 2010.

[2] H. Castro, A.P. Alves, C. Serrão, and B. Caraway, “A new paradigm for content producers,” IEEE Multimedia, vol. 17, pp. 90–93, 2010.

[3] H. C. Huang and Y. H. Chen, “Genetic fingerprinting for copyright protection of multicast media” *Soft Computing*, vol. 13, no. 4, pp. 383–391, Feb. 2009.

[4] Z. Ni, Y.-Q. Shi, N. Ansari, and W. Su, “Reversible data hiding,” IEEE Trans. Circuits Syst, Video Technol., vol. 16, pp. 354–362, 2006.

[5] A. M. Alattar, “Reversible watermark using the difference expansion of a generalized integer transform,” IEEE Trans.Image Process., vol. 13 pp. 1147–1156, 2004.