Watermarking Scheme for Authenticity and Integrity Control of Digital Medical Image using Reed-Muller Codes and Hash Block Chaining

D. S. Prathiwi, W. Astuti, Adiwijaya, T.A.B. Wirayuda
School of Computing, Telkom University
Bandung 40257, Indonesia
d.prathiwi@gmail.com, {astutiwidi, adiwijaya, cokagung}@telkomuniversity.ac.id

***Abstract*—Watermarking scheme can be a solution to embed more than one type of watermark that has a different purpose. Signature watermark can be used to embed some ownership information that is resistant to attack (authenticity control). Meanwhile, reference watermark can be used to detect modifications on digital medical images (integrity control). In this paper, signature watermark is embedded in frequency domain (Integer Wavelet Transform). To increase the robustness, Reed-Muller Codes can be applied in order to detect and improve if there is any modification attack. Authenticity control in this system is using Hash Block Chaining with the function of hash MD5. Based on the testing results, the combination of Reed-Muller on embedding process can improve the robustness of signature watermark from attack. Moreover, the fragility of reference watermark can be used to detect any attack of gaussian noise, sharpening, blurring, and JPEG compression. It means the proposed scheme has a good performance for authenticity and integrity control of digital medical images.**

Keywords—watermarking; Reed-Muller; Spread Spectrum; HBC

I. INTRODUCTION

Nowadays, data are represented in digital form, including medical images. This phenomenon will lead to a condition that medical images have high mobility, through either the Internet or exchange data between gadgets [1]. Medical image is an image that represents part of human body, which produced through medical technology to be used in specific purposes, such as diagnosis purpose. There are some protection demands against digital medical images, such as [10]: (1) Origin authentication, represents authentication of image ownership. A medical image should be identified correctly, so the ownership of information should be kept well, means it has to be robust to any kind of image modification. (2) Integrity Control, represents checking the authentication of an image; whether a medical image has been modified or not. It is required an information related image integrity and must be vulnerable to possible modification in image distribution process.

The technique of data embedding is required to protect the image ownership. Since the technique should resistant to any attack, signature watermark can be used to solve it. Meanwhile, to detect any modification to the image, a technique of data embedding which is vulnerable to attack is required. Reference watermark can be used to solve it. Multiple watermarking can be a solution to embed more than one type of watermark with different characteristics.

In this paper, the embedding of signature watermark in frequency domain will use Integer Wavelet Transform [7,8]. The technique to embed signature watermark will use spread spectrum. Spread spectrum has superiority especially in information embedding to the sensitive data like medical images [6]. The robustness of signature watermark is increased by Reed-Muller Codes if there is any modification in the image. Meanwhile, the technique to embed reference watermark is using Hash Block Chaining with the function of hash MD5 [9]. Some tests will be conducted to find out robustness and vulnerability of watermark embedding result using objective assessment parameter, namely PSNR (Peak Signal-to-Noise Ratio) and BER (Bit Error Rate).

II. THE MULTIPLE WATERMARKING EMBEDDING PROCESS

The process of *multiple watermark* embedding is started by selecting images which will be the host or media to embed the *watermark,* then it separated into ROI and RONI parts [2,5].

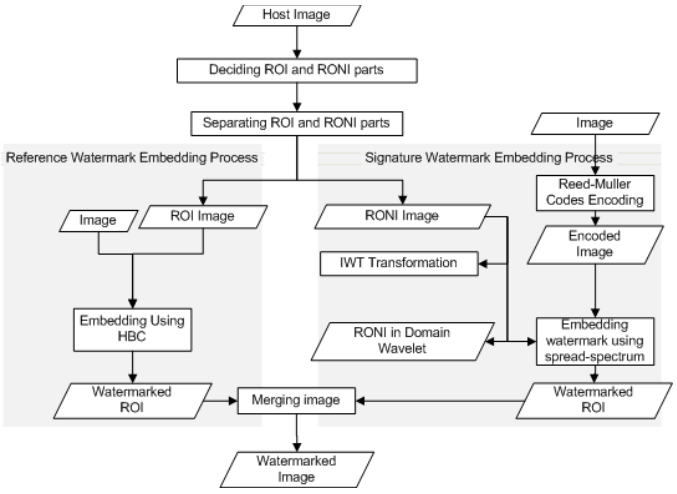

Figure 1. Watermark Embedding Process

A. Watermark Embedding

In RONI area, it is embedded to the frequency domain. While in the area of ROI image, it is embedded to spatial domain. Before embedding the *signature watermark,* firstly, it is conduct the *encoding* process using RM(1,3) [4].

1) *Encoding Watermark using RM(1,3)*

The process of encoding signature watermark using RM(1,3) as follows. Making generator matrix of $G_{RM(r,m)}$ using equations (1) and (2) that used recursively.

$$RM(0,m) = \{00\ldots0,11\ldots1\}, RM(m,m) = K^{2^m} \quad (1)$$

$$RM(r,m) = \{(x, x+y) | x \in RM(r,m-1),\}, \quad y \in RM(r-1,m-1) 0 < r < m \quad (2)$$

Encoding the message using equation (3).

$$c = m * G_{RM(r,m)} \quad (3)$$

where:

c = message which has been *encoded*

m = message which has been not *encoded* yet

$G_{RM(r,m)}$ = generator matrix

2) *Embedding Signature Watermark*

The embedding process of signature watermark to the medical images is conducted in domain frequency namely IWT. The RONI Image of I host is transformed by using IWT so it is formed the sub-bands of LL, LH, HL, and HH.

3) *Embedding Reference Watermark*

The steps in *reference watermark* embedding proces as follows. Divide host ROI image with MxN size into n Z_t block, where 0<t<n, each one has size of 8x16 pixels. Each of Z_t block will be embedded separately. Replicate the A reference watermark image which has the similar size with Host image of ROI. Divide A image which is now the size is MxN into n block of A_t, where o<t<n. Set the LSB bit into Z_t to be 0. Block Z_t with LSB value 0 is block Z_t *. Hash function which is used in proposed scheme is MD5.

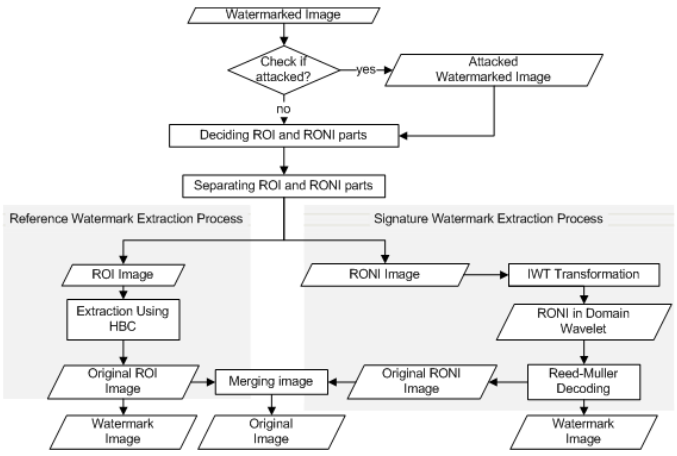

Figure 2. Watermark Extraction Process

III. THE MULTIPLE WATERMARKING EXTRACTION PROCESS

The extraction process of *multiple watermark* is not rather different with its embedding process, but the steps are done in reserve. The first step is similar to the embedding that is the separation of RONI and ROI areas. After the extraction process, it is required *decoding* process to the *signature watermark* using RM(1,3) to return the watermark image to the former condition.

1) *Extracting Reference Watermark:*

The steps in extracting reference watermark are follows.Divide the image ROI with watermark with the size of MxN into *n* of block X_t where 0 <*t*<*n*, each size is 8x16 pixels. Extract the LSB value in X_t block into D_t variable. Set the LSB bit into X_t to be 0. Block X_t with LSB value 0 is block X_t *. Do the MD5 hash function calculation. Finally, do the XOR Operation.

2) *Extracting Signature Watermark*

Process of embedding signature watermark as follows. RONI Image with I_w watermark is transformed into frequency domain by using IWT so it is formed the sub-bands of LL, LH, HL, and HH. Next, Generate the random number of R by using key. The matrix of R random number has similar size to the watermark image, that is M x N, where the value is between 0 until 1 and has N(0,1) distribution. Finally, calculate the correlation value and determine the watermark bit value from correlation value.

3) *Encoding Watermark using RM(1,3)*

After the extraction, then it is done the *decoding process* to the *signature watermark* to get the origin image. The steps of *decoding* process by using RM(1,3) with *majority logic* technique. *The majority logic* technique of the algorithm as follows. Next, the process in step *a* and *b* below to each row in the generator matrix, from the lowest row to upwards.

a) Select row in generator matrix of $G_{RM(1,3)}$. Find 2^{m-r} of characteristic vector for the row and do the dot product in each row by the encoded message.

$$G_{RM(1,3)} = \begin{bmatrix} \psi(1) \\ \psi(x_1) \\ \psi(x_2) \\ \psi(x_3) \end{bmatrix} = \begin{bmatrix} 11111111 \\ 11110000 \\ 11001100 \\ 10101010 \end{bmatrix} \quad \begin{matrix} \overline{x_1} = 00001111 \\ \overline{x_2} = 00110011 \\ \overline{x_3} = 01010101 \end{matrix}$$

To find the characteristic vector in each row, follow the steps as follows.

i. Determine the r monomial associated to the row in generator matrix.

ii. Determine the E which is the set of all x_i which are not in the r monomial, but are in the generator matrix. For example E if its *r* monomial of x_3 is $\{x_1, x_2, \overline{x_1}, \overline{x_2}\}$.

iii. The characteristic vector is the vector related to the combination x_i and $\overline{x_i}$. For example, the combination from $\{x_1, x_2, \overline{x_1}, \overline{x_2}\}$ will result characteristic vector as $\{x_1 x_2, x_1 \overline{x_2}, \overline{x_1} x_2, \overline{x_1}\,\overline{x_2}\}$. After that, do the dot product between characteristic vector and message.

b) Determine the coefficient value for selected row taken from the result majority value from dot product. For example for the row of $\psi(x_3)$, the coefficient value is 1 because 1 is the value of majority from the dot product result between characteristic vector with the message.

c) After repeating step 1 and 2 for each row except the first row, do the steps as follows. Do the multiplication between coefficients and its corresponding row. Sum the vectors resulted from the multiplication to build M_3. Add M_y with the accepted codeword. Determine the coefficient value for 1st row, it is taken from the summation above. If the number of 1 digit is bigger than 0 digit, so the coefficient is 1. If it is in contrast, so the coefficient is 0. Add the highest row which has been multiplied by its coefficient with M_y to get the original codeword. Arrange all coefficient in each row, start from the highest row until the lowest row to get the original message.

IV. Testing Results

A. Analysis of IWT Utilization

Based on the Fig. 3, it is showed that in LL and LH *sub-band,* the IWT performance is a little bit better than DWT, but in HL and HH sub-band, it is in contrast. But there is less different PSNR value of image with watermark between the ones using IWT and DWT. This is due to even the transformation process is different, the coefficient value of IWT and DWT transformation results are not much different.

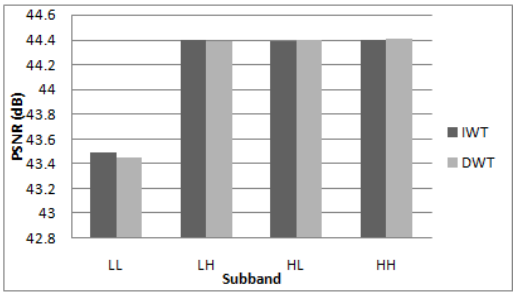

Figure 3. The comparison of IWT and DWT to the PSNR

Meanwhile, if compared in a whole, LH, HL and HH sub-bands have better quality if compared to LL. This is due to the LL sub-band has highest energy compared to other three sub-bands. So, if LL sub-band is modified, the quality will decrease.

The average embedding time of IWT is longer than DWT. This is due to the transformation process of RONI image into frequency domain in IWT, is through 3 steps after *filtering, namely Split, Predict and Update step.* While, if using DWT, the process done after filtering is *downsampling* with scale factor 2. Because the step with IWT is more if compared to DWT so, logically, if the watermark embedding time using IWT needs longer time if compared to DWT.

B. Analysis of Spread-Spectrum

Based on Fig. 4, it can be seen that the *spread spectrum* technique can improve the quality of image with watermark. The spread spectrum can improve the PSNR value of image with watermark up to 1,6 dB in each sub-band due to its characteristic. In spread spectrum, watermark is placed in many frequency locations so the energy in each location will be less, even can not be detected. Furthermore, the use of random number matrix having N(0,1) distribution can improve the robustness of *signature watermark [6]*.

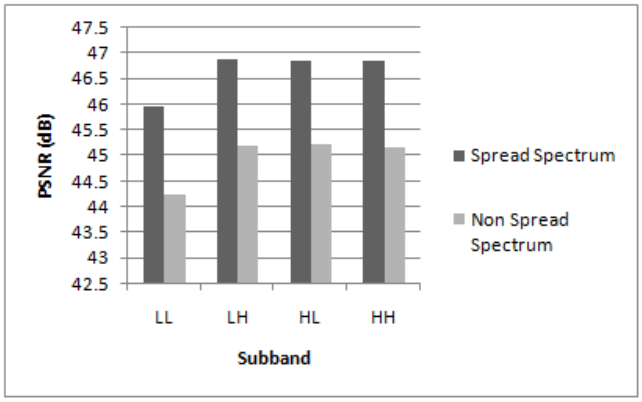

Figure 4. The comparison of Spread Spectrum and Non-SS to PSNR

In Fig. 5, it is seen that the average value of PSNR of image extraction result using *spread spectrum* is bigger than ones which are not using *spread spectrum.* As shown in figure 6, the use of watermark spread spectrum is spread in many frequency locations so the energy in each location is smaller, even it is not detected, so it will improve the robustness of *signature watermark.*

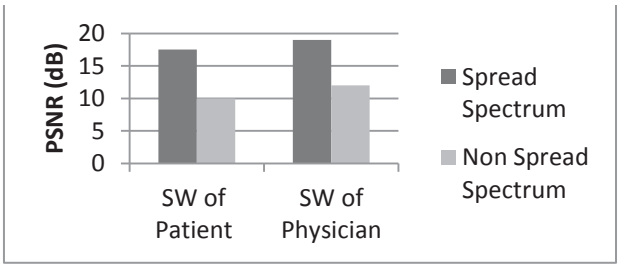

Figure 5. The effects of Spread Spectrum to the PSNR value of signature watermark image

C. Analysis of Embedding Scale Factor

In Fig. 6, it can be seen that a higher alpha level will lead to a quality of image with watermark is getting to decrease.

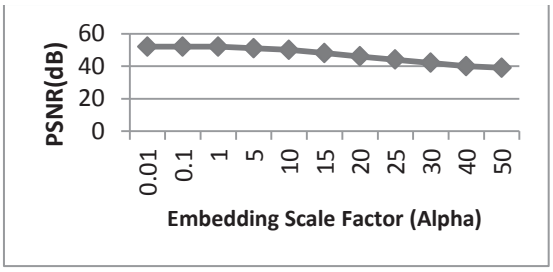

Figure 6. The alpha level effect to the quality of image with watermark

This can be seen from the PSNR decrease by the increase of alpha level. This shows that the higher alpha level, so the origin image pixel will get higher modification, resulting the decrease quality of the image. But, the contrary happens to the image of extraction result. The higher level of alpha makes watermark be increasingly robust. This is due to the alpha value shows the embedding scale strength. The bigger alpha

value, so the embedding strength of *signature watermark* will be stronger so it results to the more robustness *signature watermark.*

D. Analysis of Reed-Muller Codes Utilization

The use of Reed-Muller increases the number of bit which will be embedded, depending on its parameter. This is because of the effects to the *encoding* process, where in RM(1,3) will add the number of bit in every 4 bits to be 8 bits of codeword. It is similar to RM (1,4) that will be add the number of bit in every 5 bit to be 16 bits of codeword.

In Fig. 7, it is seen that the PSNR value decreases with the change of parameter in *Reed-Muller Codes,* so it shows that the quality of image with watermark is getting to decrease.

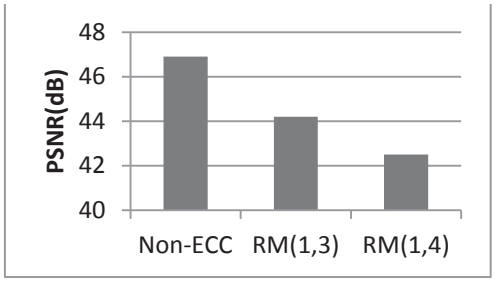

Figure 7. The effects of the use of Reed-Muller to PSNR

This is due to the number of embedded bit in the image is getting bigger by the change of parameter in Reed-Muller. If the number of embedded bit is bigger, the image quality will decrease, because there are more pixels modified for images.

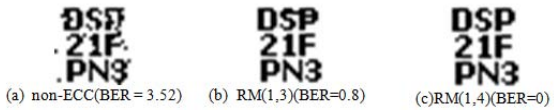

Figure 8. The extraction result of signature watermark

But, the use of Reed-Muller Codes can improve the robustness of *signature watermark*. Based on figure 7, the value of BER in watermark image which is not by the *encoding-decoding Reed-Muller* process has higher value compared to the ones using *Reed-Muller.* This is because *Reed-Muller as* an error correcting codes has an ability to correct error bits. Rm(1,3) is able to correct 1 bit error from the existence 4 bits, while RM(1,4) is able to correct 3 bits error from the exist 5 bits.

E. Impact Analysis of Block Size and Hash Function in HBC

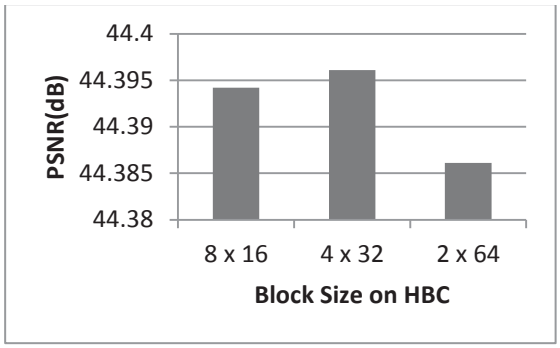

Figure 9. The effects of block size to the quality of image with watermark

Fig.9. provide the block with size of 4x32 pixels has the highest average PSNR compared to the other blocks. But, the change of PSNR value is not too significant, that is 0.01 dB. Although the block sizes are different, they all have the similar number of pixel, that is 128 pixels. The number of block is 128 pixels, because the output from the used MD5 hash function that is 128 bits is mapped to each pixel to the image.

Besides the block size, the one effecting the embedding of *reference watermark* by using HBC is the used hash function. The used hash function in this system is MD5. To know the effect of hash function at the embedding process, the used hash function is different. The used hash function in the examination are MD5, SHA-256 and SHA-512. Each hash function has different number of output bit. In MD5 output hash function results 128 bits. The SHA-256 output hash function results 256 bits. While SHA-512 output hash function results 512 bits. This will influence the used block size at embedding process. In MD5, the used block size is 8 x 16 pixels. In SHA-256, the used block size is 16x16 pixels. While, in SHA-512, the used block size is 32 x 16 pixels.

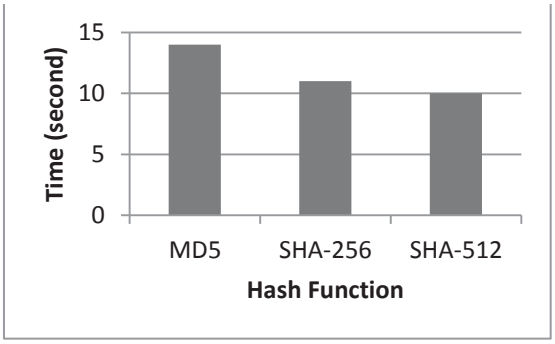

Figure 1. The hash function effects to the embedding time of reference watermark

Based on the Fig. 10, it can be seen that the MD5 hash function has the longest embedding time if compared to the two other hash functions. This is because the MD5 hash function has bigger number of block compared to the 2 other hash functions. The number of block processed by MD5 is 1152 blocks, while in SHA-256, there are only 576 blocks and in SHA-512, there are 2888 blocks.

F. The SW Robustness and RW vulnerability

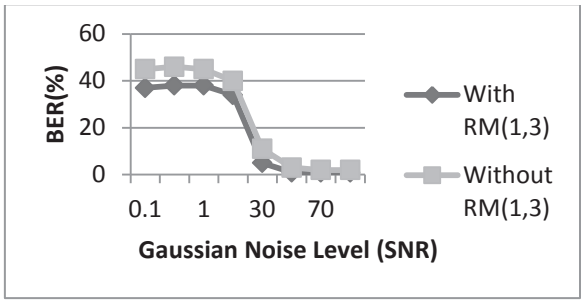

Figure 21. The effects of Reed-Muller to the attack of Gaussian Noise

The testing (Signature Watermark) robustness and RW (Reference Watermark) vulnerability is provided. Kind of attacking are gaussian noise, sharpening, blurring, and JPEG compression. In Fig. 11, it can be seen that the BER in the image of extraction result using Reed-Muller is smaller than the ones without Reed-Muller code.

This is because the Reed-Muller ability as the error correcting codes enabling to reduce the error bits while the image is attacked by Gaussian Noise. Reed-Muller has the role to improve the robustness of *signature watermark* to the attack of Gaussian Noise until reaching 4,94%.

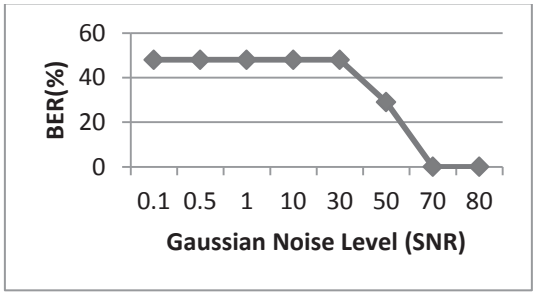

Figure 32. The effect of SNR level to the RW vulnerability

Based on Fig. 12, it is seen that the bigger SNR level given makes the smaller BER value. This is because the bigger SNR level makes the smaller noise strength given. At SNR 70 level, the BER value in *reference watermark* image has the average value that is 0, so it can be concluded that the *reference watermark* is able to detect the Gaussian Noise attack in SNR level < 70.

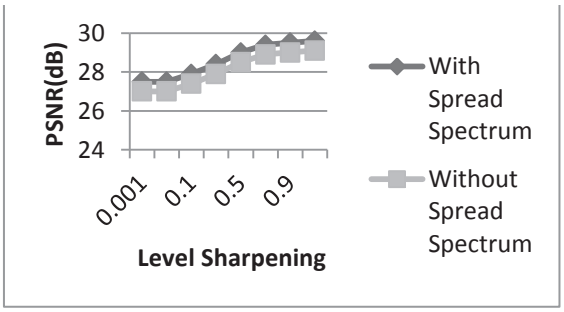

Figure 43. The effects of Spread spectrum to the sharpening attack

Based on Fig. 13, it is seen that the spread spectrum can improve the image quality with watermark until 0.5 dB constantly. This shows that the spread spectrum works well at all levels of sharpening attacks, namely from 0.001 level to 1. Also, it is seen the PSNR value increasingly bigger by the increase of sharpening level value. Even though the bigger level value of sharpening attacks, so the given attack is increasingly harder, but the image quality with watermark does not decrease. This occurs because the bigger sharpening level makes the clearer image, and meaning the more changes in the image and bigger interference with watermark bits, but the changes to the image does not significantly affect the quality of the image with watermark.

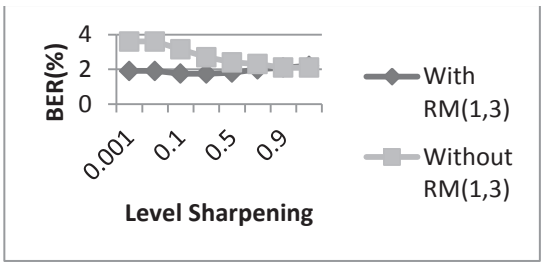

Figure 14. The effect of Reed-Muller to the Sharpening attack (BER)

While the role of Reed-Muller codes, can be seen in Fig. 14. BER value of extraction image result by using the Reed-Muller at the level of the same attack is smaller until 0.75 level than the ones which are without using the Reed-Muller. This is because the ability of Reed-Muller as error correcting codes can reduce the error bits at sharpening level of 0.001 to 0.9. But, at the attack level of > 0.9, Reed-Muller is not able to reduce the error bits. Reed-Muller has the role to improve watermark robustness against sharpening attacks up to 1.8% in the sharpening attack level <0.9. Meanwhile, the effect of sharpening to the vulnerability of reference watermark produce a quite big BER namely, 30% to 48.1%. This shows a sharpening attack is a harsh attack.

The influence of Reed-Muller codes, it can be seen in Fig.15. At the level of 0.001 to 0.05, the given attack blur is not too strong so the average BER value is small. The average BER value to the system using RM (1.3) is smaller than the ones which are without the use of RM (1.3). This is because the Reed-Muller as error correcting codes can reduce the error bits after the images are being attacked.

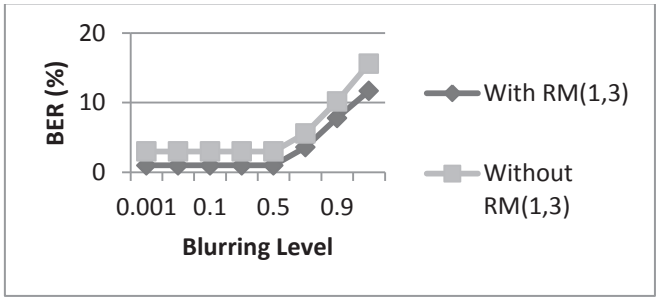

Figure 5. The effect of Reed-Muller to the Blur attack (BER)

The average BER value in the blur attack level which is above 0.5 increases by the given increase blur levels. This is due to the stronger given blur attack, but Reed-Muller is still able to reduce the BER value up to 4% in blur attack level with 1 value. The Reed-Muller has the role to improve watermark robustness against the blur attack by lowering the average BER value up to 2.07% on all levels of tested blur attack. That means that the given blur attack is very small so it is not detected by damaged reference watermark. But at the blur attack level above 0.5, attack power is high enough so it can increase the average BER value in the reference watermark. This is because the higher the blur attack level means the stronger attack which is given to the image.

Based on Fig. 16, it shows that the BER value on the extraction result image with 0.1 to 25 compression ratio is quite high until reaching the average of 18%. This shows very big compression scale factor which is increasing the number of error bits. But the strength compression decreases in compression above 25 by the addition of compression ratio. Reed-Muller can work in a small value of compression ratio so, it can lower the average BER value up to 2% in 25 of ratio value. This is because the Reed-Muller as error correcting codes can reduce the error bits after the image is attacked. Even though the performance of Reed-Muller is less optimal in the 75 ratio, but the change in BER value is not so significant so in the ratio of 90 to 99, Reed-Muller is able to reduce BER value up to 2.3%.

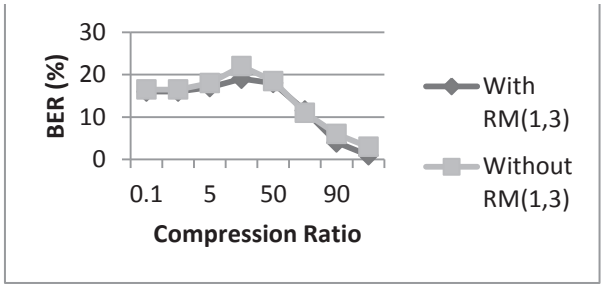

Figure 16. The effects of Reed-Muller to the JEPG compression

Meanwhile, the effect of blurring to the vulnerability of reference watermark produce a quite big BER namely, 30% to 50 %. The average BER which is high enough shows that the attack of JPEG compression is a strong enough attack resulting in severe damage to the image extraction result of reference watermark. on the other side, big enough BER value, it is can be said that the reference watermark is vulnerable to the attack of JPEG compression because in compression ratio of 99, the resulted BER value is quite high.

V. Conclusion

The use of Reed-Muller Codes on the embedding of signature watermark produce a robust signature watermark. Moreover, HBC on the embedding of reference watermark provide vulnerable of reference watermark to the attacks of gaussian noise, sharpening, blurring and JPEG compression. It means that the proposed scheme give a good performance for authenticity and integrity control of digital medical images.

References

[1] Adiwijaya, T.A.B. Wirayuda, S.D. Winanjuar, U. Muslimah, The Multiple Watermarking on Digital Medical Image for Mobility and Authenticity, Operations Research Proceedings 2012, 457-462. DOI: 10.1007/978-3-319-00795-3_68

[2] Adiwijaya, Faoziyah, P.N., Permana, F.P., Wirayuda, T.A.B., Wisesty, U.N., Tamper detection and recovery of medical image watermarking using modified LSB and Huffman compression, Second International Conference on Informatics and Applications (ICIA), 2013, pp.129 - 132. DOI: 10.1109/ICoIA.2013.6650242

[3] Agustina,R., Adiwijaya, and Barmawi, A.M., "*Pendeteksian dan Perbaikan Citra Termanipulasi yang Disisipi Watermark Menggunakan Block Truncation Coding (BTC) Berbasis Wavelet*," Jurnal PP Telekomunikasi, vol 15 No. 2, Juni 2011.

[4] Cooke, Ben, "*Reed-Muller Error Correcting Codes*," MIT Undergraduate Journal of Mathematics, 1999, Vol. 1, pp. 21-26.

[5] Kurniawan, M.T., Adiwijaya, Agung, W., Multiple watermarking on digital medical images for tamper detection and integrity control, International Conference on Uncertainty Reasoning and Knowledge Engineering, 2012, pp. 145 - 148. DOI: 10.1109/URKE.2012.6319530

[6] Kumar, B.,Anand, A., Singh, S.P.,Mohan, A., "*High Capacity Spread-Spectrum Watermarking for Telemedicine Applications*," World Academy of Science, Engineering and Technology 79 2011.

[7] Lee, Sunil,et al., "*Reversible Image Watermarking Based on Integer-to-Integer Wavelet Transform*". IEEE Transactions on

Information Forensics and Security Vol. 2 No. 3 September 2007.

[8] Mostafa, S.A.K., El-sheimy,N., Tolba, A.S., Abdelkader,A.S., Elhindy,H.M., "*Wavelet Packets-Based Blind Watermarking for Medical Image Management"*. Mill The Open Biomedical Engineering Journal Vol. 4 pages 93-98 2010.

[9] Rivest, R., " *The MD5 Message-Digest Algorithm*". MIT Laboratory for Computer Science and RSA Data Security, Inc

[10] Zain, J.M, Fauzi, A.R.M., Medical Image Watermarking with Tamper Detection and Recovery, International Conference of Engineering in Medicine and Biology Society, 2006.pp. 3270 - 327