CHARACTERISTIC ANALYSIS OF CARRIER BASED ON THE FILTERING AND A MULTI-WAVELET METHOD FOR THE INFORMATION HIDING

SHUAI REN*
School of Information Engineering
Chang’an University
Xi’an, Shaanxi 710064, China

TAO ZHANG
School of Electronic and Control Engineering
Chang’an University
Xi’an, Shaanxi 710064, China

FANGXIA SHI
School of Information Engineering
Xizang Minzu University
Xianyang, Shaanxi 712082, China

ABSTRACT. The characteristic consistency between the original and the stego carriers is an important indicator to evaluate an information hiding algorithm. Different from the traditional carrier pre-processing methods which are based on the operation domains, we propose a characteristics analysis-based pre-processing scheme. We use the Gaussian pyramid filtering and CARDBAL2 multi-wavelet transform to analyze the energy characteristics of the carrier, so the original carrier can be decomposed into several sub-regions with different energy level. And at the same time, the processed carrier shows us the redundancy space structurally through the combination bit plane method, which can provide some invisible hiding positions. Obviously, the energy and structure characteristics are at least related with the robustness and invisibility of the hiding result respectively. So we can improve these performances compared with the traditional methods. At the same time, some optimization theories like the Chebyshev map are used to improve other performances. At last, the experimental shows the achievements of this scheme in the form of data.

1. Introduction. Information hiding is an important way to secure the confidential communication. Improvement of the performance is always a research focus in information hiding scheme based on digital image. It has been believed that the most important purpose for the cover image pre-treatment method is to improve
the performance of hiding [5]. Currently, the space domain- and frequency domain-based are the main methods of the cover image pre-treatment for information hiding [12]. LSB is the basic algorithm of space domain methods, but the schemes based on LSB almost have drawbacks under embedding great amount of information [1]. Methods in frequency domain, such as DCT and DWT, have robustness against certain attacks, but are costly and of low robustness against rotating attacks [10]. Using spatial- or transform-based pre-treatment method rather than analysing the energy and structure of carrier, the existing information hiding approaches similar with the above methods can achieve neither a good feature matching between the carrier and the secret image nor the minimum distortion to the initial data or coefficient [13].

New technologies about image processing emerged to enrich the pre-processing methods of information hiding carrier. At the same time, more and more researchers want to combine methods of different operand to improve the performance of information hiding. In our previous research, we have studied about several energy and structure analyse methods. By energy analyse, we generally obtain the energy differential regions to embed the data with different robustness, and this procedure can improve the robustness and imperceptibility directly, even other performances. By structure analyse, we can obtain the color and shape parameter of the carriers which can be taken as the modifiable parameters for the embedding. After the pre-treatment based on feature analyse, the embedding procedure can modify the original information of carrier as less as possible.

In this paper, we use Gaussian pyramid and CARDBAL2 multi-wavelet transform theory as the analyse methods. Abed et al. [4] introduced Gaussian pyramid, in which energy of images can be gradually concentrated layer by layer through low-pass filtering and sub-sampling. CARDBAL2 multi-wavelet transform, which can satisfy compact support and symmetry of the image processing, was proposed to be a new frequency domain method recently [2]. Combining Gaussian pyramid and CARDBAL2 multi-wavelet transform theory, we proposed an information hiding scheme. These two methods can analyse the energy and structure characteristics of the carriers into some operable parameters, and then embed the secret information without going against the properties of the carriers. Put static original image $G$ into Gaussian pyramid to set up pyramid model of five layers, which are, comprised of five images namely $G_0, G_1, G_2, G_3, G_4$. Transform $G_0, G_2, G_4$ with CARDBAL2 multi-wavelet first-order transform separately namely $G'_0, G'_2, G'_4$. Then to get four components themselves from $G'_0, G'_2$ and $G'_4$, denoted by $LL_{i1}, LH_{i1}, HL_{i1}$ and $HH_{i1}$, $i = 0, 2, 4$ which stand for $G_0, G_2, G_4$ and $i = 1, 2$ which stand for the lowest resolution sub-image and the four sub-image of the lowest resolution sub-image. $LL_{21}, LH_{21}, HL_{21}$ and $HH_{21}$ are decomposed into RGB color components, and B component is picked up into bit plane decomposition. Bit Plane 0 is used as embedding region. In addition, this paper applies Chebyshev map to scramble cover image. Use genetic algorithm to find the optimal scrambling parameters to improve the consistency of embedded information bits with cover images [8]. Embed main information with RAID4 method. Data of Bit Plane 0 can be judged and restored by setting Bit Plane 3 and Bit Plane 7.

Application of CARDBAL2 multi-wavelet transform in information hiding algorithm can sufficiently take advantage of energy distributing characteristics of CARDBAL2 multi-wavelet sub-images [7]. In terms of image scrambling, many
scrambling methods proposed and often applied in design of information hiding algorithms. But in most cases, scrambling algorithms are used to scramble embedded information. Considering about extraction complexity, the strategy should greatly lower the security of algorithms. Chebyshev map with three parameters provides more key parameters selections which are controlled freely by legitimate users. And Chebyshev map also has lower time complexity of calculation [9].

2. Gaussian pyramid and CARDBAL2 multi-wavelet theory. Gaussian Pyramid (GP) is an effective and simple structure with multi-resolution to explain images. Compared with other pyramid algorithms, GP is of good visual effect and less computation. In GP layers, the bottom is the original image. All the layers can be obtained by the following equation iteratively.

\[ G_l = \sum_m \sum_n w(m, n)G_{l-1}(2i + m, 2j + n) \]  

Eq.(1) can be abbreviated as \( G_l = \text{REDUCE} \ [G_{l-1}] \), in which \( l \) represents level of pyramid layers and \( 0 < l < N \), \( w(m, n) \) is the Gaussian kernel function.

The Figure 1 (a) shows the structure map of GP with five layers, and GP Transform to the Lena image. The original image \( G_0 \) is repeatedly operated to generate the sequence of reduced resolution image \( G_1, G_2, G_3, G_4 \).

Multi-wavelet transformation is a new concept of wavelet transformation architecture, which has more than one scaling function \( \phi(t) \) and wavelet function \( \varphi(t) \). Multi-wavelet is usually indicated by multi-dimensional vector function. CARDBAL2 multi-wavelet transformation is earliest constructed and the most widely used multi-wavelet, it has notable features such as compact support, second-order approximation, scalar function with integer translation and orthogonality, higher-order vanishing moments and symmetry [3].

Figure 2 is the first-order decomposition example of the original image \( G_0 \) with CARDBAL2 multi-wavelet transformation. It can be seen from the above figure, after the CARDBAL2 multi-wavelet transform \( LL_0^0, LH_1^0, HL_2^0, HH_2^0 \) are clearly visible.

3. Information hiding algorithm. According to study by Run et al. [11], we can learn the energy distribution of first-order CARDBAL2 multi-wavelet shown in Table 1. The energy distribution of four components is similar to 1.2:1.1:1.1:1.0. Uses components characteristics after CARDBAL2 decomposition, and uses weights
of different components in RGB mode and weights of all Bit Planets to propose regions of GP-CDB. Specifically, first, $LL_2^2$, $LH_2^2$, $HL_2^2$ and $HH_2^2$ are decomposed into RGB color components, and B color component is picked up into bit plane decomposition. Bit Plane 0 is used as embedding region. Information embedding rules are as follows:

Rule1. Embeds main information in the medium energy area $LH_2^2$ and $HL_2^2$. Uses high energy component $LL_4^2$ as information recovery and $HH_0^2$ as distortion judgement unit;

Rule2. Bit Plane 0 of four components is information embedding plane;

Rule3. Embed main hiding information in $LH_2^2$, $HL_2^2$ components with RAID4;

Rule4. Bit Plane 3 and Bit Plane 7 are respectively Datum Plane and Auxiliary Plane. $C_7$ and $C_3$ are respectively binary data of Bit Plane 7 and Bit Plane 3. When $C_0$ is the embedding information, $C_3$ is modified data according to Eq. (2). Expanding the embedded rule to all bit planes can overcome defect of the former least significant bit.

$$C_3 = C_7 \oplus C_0$$

Information hiding algorithm based on Gaussian Pyramid and CARDBAL2 multi-wavelet is divided into nine steps:

**Step 1.** The carrier image is decomposed into five layers by Gaussian Pyramid algorithm. $G_0$, $G_2$, $G_4$ are divided respectively through CARDBAL2 multi-wavelet to receive $HH_0^2$, $LH_2^2$, $HL_2^2$ and $LL_4^2$ components. Choose $HH_0^2$ as the fragile sign embedding unit; choose $LH_2^2$ and $HL_2^2$ as hiding data unit; choose $LL_4^2$ as robustness information embedding unit;

**Step 2.** Scramble $HH_0^2$, $LH_2^2$, $HL_2^2$ and $LL_4^2$ images with Chebyshev map Eq.(3);

$$\begin{align*}
1, & \quad -1 \leq x_{k+1} < \eta \\
0, & \quad \eta \leq x_{k+1} \leq 1
\end{align*}$$

$$x_{k+1} = \cos(\mu \arccos(x_k)) , x_n \in [-1, 1]$$
Step 3. Draw the Bit Plane 0, Bit Plane 3 and Bit Plane 7 of our embedding unit after the scrambling;

Step 4. After the scrambling, traverse the Bit Plane 0 of $HH^0_2$, $LH^2_2$, $HL^2_2$ and $LL^4_2$ line by line separately to get a one-dimensional sequence, and denoted separately $C_1$, $C_2$, $C_3$ and $C_4$. Denote $LH^2_2$ and $HL^2_2$ separately by $C_{LH} = t'_1, t'_2, \ldots, t'_k, t'_i$ and $C_{HL} = t''_1, t''_2, \ldots, t''_k, t''_i$. Get the final coding of units $LH^2_2$ and $HL^2_2$ from $C_{LH}$ and $C_{HL}$ and denoted by $C = t'_1, t''_1, t'_2, t''_2, \ldots, t'_k, t''_k, t'_i, t''_i = t_1, t_2, \ldots, t_n$, $n = 2i$;

Step 5. Coding the hiding data with traversing line by line, get a one-dimensional sequence namely $L_1$ size of $m \times n$. Calculate $R = \left\lceil \frac{N \times N}{m \times n} \right\rceil$. Repeat the elements of $L_1$ successively for $R$ times to obtain a one-dimensional sequence which is $R$ times length of the original one. If the length of the new sequence is less than $N \times N$ then fill zero in the following space to make the length of the sequence achieve $N \times N$. The new sequence size of $N \times N$ denoted by $L_2 = b_1, b_2, \ldots, b_n$. Secret information $L_2$ is cross hidden into hiding area generated from $LH^2_2$ and $HL^2_2$ by RAID4 (eight bits is a basic data block unit);

Step 6. In order to optimize the sequence of embedded bits with genetics algorithm, suppose $F$ as the amount of the same bit value in matched positions of $L$ and $C$. Use genetic algorithm to change parameters $x_n, \mu, \eta$, in order to maximize $F$. The optimization model based on GP-CDB is Eq.(4):

$$F(y) = \text{Max} F(x_k) = \text{Max} \sum_m (t_n \oplus b^n_k) \quad (4)$$

Step 7. Embed RAID4 check data and Hash value (denoted as $R^L$) of information in $LL^4_2$, $HH^0_2$ is the most vulnerable part. This algorithm embeds, Hash value (denoted as $R^H$) of information and scrambling times in $HH^0_2$ same as the above step. Then receivers can quickly determine whether the image has been distorted by checking $R^L$ and $R^H$;

Step 8. operation to extend size of $G_1$, $G'_2$, $G_3$ and $G'_4$ to the same as the original $G_0$, shown as the Eq.(4) in which $G_{l,k}$ is the image extended $k$ times of $G_l$. Five layer images after EXPAND are $G_{00}$, $G_{11}$, $G_{22}$, $G_{33}$ and $G_{44}$. Denoted as $G_{l,k} = \text{EXPAND}[GG_{l,k-1}]$, or $G_{l,0} = G_l$, in which $k > 0$; $G_l = 4 \sum_m \sum_n G_{i,k-1}(2^\frac{m+n}{2}, 2^\frac{3+n}{2})$;

Step 9. Get reconstruction image $G'$ (Stego image) by superposition of $G_{00}$, $G_{11}$, $G_{22}$, $G_{33}$ and $G_{44}$ shown as the following Eq. (5).

$$g'_{ij} = f(g^0_{ij}, g^{11}_{ij}, g^{22}_{ij}, g^{33}_{ij})_{ij} = \begin{cases} g^x_{ij}, & g^x_{ij} \neq g^0_{ij} \\ g^0_{ij}, & x = 1, 2, 3, 4. \end{cases} \quad (5)$$

The process of information extraction is divided into four steps:

Step 1. The stego image is decomposed into five layers by Gaussian Pyramid algorithm. $G'_0$, $G'_2$ and $G'_4$ are divided respectively through CDBAL2 multi-wavelet to receive four first-order component sub-images;

Step 2. Draw the information from Bit Plane 0 of $LL^4_2$ and $HH^0_2$. Extract $R^L$ from $LL^4_2$, and $R^H$ from $HH^0_2$. Of course, need to using Bit Plane 0 and Bit Plane 7 to verify and adjust the $R^L$ and $R^H$;
Table 2. Comparison of hiding algorithm imperceptibility based on PSNR.

| Hiding Algorithm | GP-CDB | GHM-CT | LSB | DWT |
|------------------|--------|--------|-----|-----|
| PSNR             | 37.92  | 34.2   | 25.13 | 30.1 |

Step 3. If \( R_L = R_H \), it means the stego image has not been attacked. If not, it means the stego image has been attacked or distorted, then continue to determine;

Step 4. Extract RAID4 check data and Chebyshev map times, and using them to extract hiding information from \( LH_2^2 \) and \( HL_2^2 \) with strategies of BP shown in Fig. 3.

4. Simulation experiment. Simulation environment is Matlab7.0.0.19920. Cover image is Lena (256 × 256) (Figure 3 (a)). Stego image is binary image Baboon (64 × 64) (Figure 3 (b)).

Get the stego image shown in Fig. (4c) which is based on the Gaussian Pyramid and CARDBAL2 when \( l = 5 \). PSNR of the stego image to the carrier image is 37.92, this means that the proposed algorithm is of good imperceptibility. Compared to the current hiding algorithms GHM-CT and DWT algorithms, this scheme has its superiority through imperceptibility contrast according to PSNR. When embedding rate is 25\%, the imperceptibility can be averagely improved by 27.21\%, shown in Table 2.

Define texture evaluation and modification rate of binary image \((n \times n)\) separately in Eq. (6) and (7).

\[
w = \frac{\sum_{i=0}^{n-1} \sum_{j=0}^{n-1} f(i,j) \oplus f(i+\mu,j+\eta)}{2n^2} \tag{6}
\]

\[
p = \frac{\sum_{i=0}^{n-1} \sum_{j=0}^{n-1} f(i,j) \oplus f'(i,j)}{n^2} \tag{7}
\]

Where \( n = N/2^k \), \( k \in \{1,2,\ldots,\log_2(N-1)\} \). \( f(i,j) \) and \( f'(i,j) \) are separately for the pixel at \((i,j)\) of normal and extraction image with \( n \times n \) pixels.

Robustness test algorithm is defined in Eq. (7). \( Q \) is robustness test value (RTV) and \( Q \in [0,1] \). In the following experiments, \( \mu = \eta = 5 \), and will expands \( Q \) a hundred times to accommodate judgment habit. Information extraction is the most preserved when \( Q = 100 \).

\[
Q = w(1 - p) \tag{8}
\]
Table 3. RTV comparison results of filtering and noise.

| Attacks             | Each component energy distribution |
|---------------------|------------------------------------|
|                     | GP-CDB    | GHM-CT   | LSB     | DWT     |
| [3,3] mean filter   | 66.31     | 62.12    | 30.58   | 66.26   |
| Wiener2 filter      | 57.76     | 52.47    | 40.25   | 57.32   |
| Gaussian            | 73.25     | 74.46    | 20.15   | 66.22   |
| Salt& pepper noise  | 34.37     | 32.83    | 21.01   | 29.26   |

Figure 4(a) shows the extracted images corresponding to different Q while detecting extracted images in accordance with the robustness test formula. JPEG2000, cutting, rotation, are the common image processing operations. Figure 4(b) to Figure 4(d) shows the RTV of GP-CDB and other comparison algorithm under JPEG2000 compression, cutting processing from 0% to 100% and rotation processing from 0° to 180°.

![Extracted images corresponding to different Q values](image1)

![RTV results under JPEG2000 from 0% to 100%](image2)

![RTV results under Cutting from 0% to 100%](image3)

![RTV results under rotation from 0° to 180](image4)

Figure 4. Extracted images corresponding to different Q values and Comparison results of robustness experiment.

Table 3 lists the RTV of GP-CDB and other algorithm under the mean filter [3,3], the wiener2 filter [3,3], the Gaussian noise ($\mu = 0, \sigma^2 = 0.003$) and the ' salt & pepper ' ($d = 0.15$).

Figure 4(a) shows that the extracted information can be identified when $Q > 30$. Experiment show that GP-CDB algorithm is robust against JPEG2000 below 80%, cutting below 63%, rotation below 26°, common filtering and adding noise.

Sensitivity to image attacks is the peculiar characteristic in GP-CDB. Comparing Hash value of $LL_4$ and $HH_2$ indicates GP-CDB has excellent sensitivity to image processing. Table 4 lists the detectable rate when JPEG2000 compression ratio is 5%, random cutting ratio is 5%, rotation ratio is 1°, [3,3] median filter, Gaussian
Table 4. Detectable rate of attacks.

| Processing  | JPEG2000 | Cutting | Rotation | Filtering | Gaussian | ’Salt & Pepper’ |
|-------------|----------|---------|----------|-----------|----------|----------------|
| Detectable  rate | 98.08%   | 98.83%  | 96.70%   | 97.98%    | 95.56%   | 98.08%         |

($\mu = 0, \sigma^2 = 0.003$) and ’salt & pepper’ ($d = 0.15$). The average of detectable rate is 97.54%.

RS can detect whether or not having hiding information by comparing the difference value between $R_m$ and $R_m$, $S_m$ and $S_m$. Higher order statistics detection algorithm based on wavelet coefficients (HOSWC) is a general detection algorithm [6]. Use the algorithms above-mentioned to analyze the performance of GP-CDB. Experiment results are shown in Fig. 5.

**Figure 5.** Experiment result of RS and HOSWC to GP-CDB.

Fig. 5 (a) show that the maximum difference value of $R$ is 341 (except initial difference value about 121). The maximum difference value of $S$ is 153. It indicates the embedding rate cannot directly influence on difference value. Detect the 50 random pictures which hiding information based on HOSWC. Fig. 5 (b) show there is no one or more threshold value found to recognize which pictures embedded information. Thus show that GP-CDB resists such analysis.

5. **Conclusion.** This paper proposes a new reversible information hiding algorithm based on Gaussian pyramid and CARD BAL2 multi-wavelet transformation. The algorithm decomposes image into hierarchical structure of five images, in which energy of images gradually becomes concentrated with the raise in layers. Using CARD BAL2 multi-wavelet transformation decomposes images into small pieces according to energy distribution. Known from Gaussian pyramid characteristics, image frequency reduces and image energy concentration increases layer-by-layer. In this paper, the original image is decomposed into five layers. Energy of the zero layers or the original image is distributed uniformly, so puts the zero layers into CARD BAL2 multi-wavelet transformation to obtain second-order sub-images and chooses second-order high frequency sub-image as fragile sign embedding region. Therefore, the fragile sign is robust against regular use and sensitive to common changes. Energy of the second layer image becomes relatively concentrated. In the four second-order components, choose medium frequency sub-images as information embedding region. This design makes imperceptibility, robustness, anti-analysis and embedding capacity of the algorithm reached a more balanced level. The fourth layer
image has most energy of the original image, and put it into CARDBAL2 multi-wavelet transformation. Embed robustness sign into low frequency part of the four second-order sub-images. This part is the most robust region, so this algorithm is of good robustness. The algorithm is reversible shown in part 3. Experiments show that the algorithm is effective and feasible and it is of better robustness and imperceptibility.

REFERENCES

[1] A. Basu, A. Saha, J. Das, S. Roy, S. Mitra and I. Mal, On the implementation of a digital watermarking based on phase congruency, *Advances in Intelligent Systems & Computing*, 328 (2014), 113–120.

[2] L. Cui and W. Li, Adaptive multiwavelet-based watermarking through JPW masking, *IEEE Transactions on Image Processing*, 20 (2011), 1047–1060.

[3] A. Durdek, S. R. Jensen and J. Juselius, et al., Adaptive order polynomial algorithm in a multiwavelet representation scheme, *Applied Numerical Mathematics*, 92 (2015), 40–53.

[4] K. Jafari-Khouzani, Multiwavelet grading of pathological images of prostate, *IEEE Trans. Biomed. Eng.*, 50 (2003), 697–704.

[5] J. Kim, M. Yi and M. S. Obaidat, Advanced computer mathematics based cryptography and security technologies, *International Journal of Computer Mathematics*, 90 (2013), 2512–2514.

[6] Y. Kutlu, Feature extraction for ECG heartbeats using higher order statistics of WPD coefficients, *Comput. Methods Programs Biomed.*, 105 (2012), 257–267.

[7] Y. Li, H. L. Wei and S. A. Billings, Identification of time-varying systems using multi-wavelet basis functions, *Control Systems Technology IEEE Transactions on*, 19 (2011), 656–663.

[8] X. B. Ren, An optimal image thresholding using genetic algorithm, *Proceedings of the 2009 International Forum on Computer Science-Technology and Applications*, 1 (2009), 169–172.

[9] B. Stoyanov and K. Kordov, Novel image encryption scheme based on chebyshev polynomial and duffing map, *The Scientific World Journal*, 2014 (2014), Article ID 283639, 11 pages.

[10] M. Saini and R. Chhikara, Performance evaluation of DCT and DWT features for blind image steganalysis using neural networks, *International Journal of Computer Applications*, 114 (2015), 20–23.

[11] A. R. Tate, N. Beloff and A. R. Balques, et al., Exploiting the potential of large databases of electronic health records for research using rapid search algorithms and an intuitive query interface, *Applied Numerical Mathematics*, 21 (2014), 292–298.

[12] L. Yang, J. Lv and Y. Xiang, Underdetermined blind source separation by parallel factor analysis in time-frequency domain, *Cognitive Computation*, 5 (2013), 207–214.

[13] T. Zhang, D. J. Mu and S. Ren, A confidential communication-oriented information hiding algorithm based on GHM multi-wavelet and DCT, *Applied Mathematics & Information Sciences*, 7 (2013), 1803–1807.

Received June 2015; revised September 2015.