Soft-decoding of the (23, 12, 7) Binary Golay Code

Wen-Ku Su, Pei-Yu Shih, Tsung-Ching Lin, and Trieu-Kien Truong

***Abstract*—Since the (23, 12, 7) binary Golay code is a perfect code, a weight-4 error occurred is always decoded as a weight-3 error pattern by a hard decoding method. In this paper, an efficient soft-decision decoder of the (23, 12, 7) binary Golay code up to the four errors is proposed. All probable patterns of occurred weight-4 error, which are always decoded to the same weight-3 error pattern, are determined from the look-up table of weight-7 codewords. And the most possible error pattern of weight-4 or weight-3 will be obtained by estimating the emblematic probability values of all probable patterns. The simulation result of this decoder in additive white Gaussian noise (AWGN) shows that at least 93% and 99% of weight-4 error patterns occurred are corrected if a bit-energy to noise-spectral-density ratios (E_b/N_0) are greater than 3 dB and 6 dB, respectively, and at least 96% of weight-3 error patterns occurred are corrected for any dB.**

I. Introduction

The (23, 12, 7) Golay code, which was found by Prange [1], is one of the most important binary quadratic residue (QR) codes. A *t*-error-correcting code can correct a maximum of *t* errors. A perfect *t*-error-correcting code has the property that every word lies within a distance of *t* to exactly one code word. The (23, 12, 7) binary Golay code is particular significance since it is a perfect 3-error-correcting code. That is, the (23, 12, 7) Golay code allows for the correction of up to three errors and each four error pattern occurred is always decoded error as a three error pattern. There are some efficient hard decoding methods for the (23, 12, 7) binary Golay code: the algebraic decoding algorithm proposed by Elia [2], the reliability-search decoding algorithm proposed by Dubney *et al.* [3]. In this paper, a new efficient soft decoding technique is presented to correct the four errors.

In [3], the reliability-search algorithm was developed to facilitate further decoding of the (23, 12, 7) Golay code. In that algorithm, using real channel data, the method developed by Reed [4] can be used to estimate the individual bit-error probabilities in a received word. In the soft decoding for more than two errors occured, one utilizes the same method to estimate the individual bit-error probabilities in a received word. And the emblematic probability values of possible error patterns occurred of five weight-4 and one weight-3 are defined. Finally, according to the greatest embalming probability value, the most possible error pattern occurred is obtained.

The structure of this paper is as follows: The background of the (23, 12, 7) binary Golay code is given in Section II. Section III presents the algebraic decoding algorithm for the binary Golay code. Section IV describes the soft decoding of the (23, 12, 7) binary Golay code for three or four errors occurred. A few short remarks and conclusions are given in the final section.

II. (23, 12, 7) Binary Golay Code

It is not difficult to show that $g(x)=x^{11}+x^9+x^7+x^6+x^5+x+1$ is an irreducible polynomial over $GF(2)$. Thus there exists an element $\alpha \in GF(2^{11})$ such that $g(\alpha)=0$. Hence the elements of $GF(2^{11})$ are found in the following set.

$$GF(2^{11}) = \{\, a_0 + a_1\alpha + a_2\alpha^2 + \cdots + a_{10}\alpha^{10} \mid a_0, a_1, \ldots, a_{10} \in GF(2)\}. \quad (1)$$

Note also that α is a primitive 23rd root of unity in $GF(2^{11})$.

The codewords of Golay code over $GF(2)$ are expressed first as the coefficients of a polynomial. In such a representation, a codeword is represented by

$$C(x) = \sum_{i=0}^{22} c_i x^i \quad (2)$$

where $c_i \in GF(2)$ and x is an indeterminate. Written as a vector, the codeword is $\boldsymbol{C} = (c_0, c_1, \ldots, c_{22})$.

The generator polynomial of a Golay code as discussed above in an irreducible polynomial given by

$$g(x) = \prod_{i=0}^{10}(x-\alpha^{2^i}) = x^{11}+x^9+x^7+x^6+x^5+x+1. \quad (3)$$

Now let polynomials

$$I(x) = c_{22}\,x^{22} + c_{21}\,x^{21} + \cdots + c_{11}\,x^{11} \quad (4)$$

The codeword $C(x)$ must also be a multiple of the generating polynomial $g(x)$. That is

$$C(x) = I(x)g(x). \quad (5)$$

Manuscript received Jannuary 8, 2008. The work was supported by National Science Council, R.O.C., under Grants NSC95-2221-E-214-042.

Wen-Ku Su is with the Department of Information Engineering, I-Shou University, Tahsu, 84001, Kaohshiung, Taiwan (886-939913808; fax: 886-7-6577293; e-mail: d9403003@stmail.isu.edu.tw).

Pei-Yu Shih is with the Department of Information Engineering, I-Shou University, Tahsu, 84001, Kaohshiung, Taiwan (e-mail: d9203005@stmail.isu.edu.tw).

Tsung-Ching Lin is with the Department of Information Engineering, I-Shou University, Tahsu, 84001, Kaohshiung, Taiwan (e-mail: joe@isu.edu.tw).

Trieu-Kien Truong is with the Department of Information Engineering, I-Shou University, Tahsu, 84001, Kaohshiung, Taiwan (e-mail: truong@isu.edu.tw).

To illustrate algebraic decoding algorithm, define

$$E(x) = e_{22}\, x^{22} + e_{21}\, x^{21} + \cdots + e_1\, x + e_0 \tag{6}$$

to be the error polynomial. Written as a vector, the error vector is $\boldsymbol{E} = (e_0, e_1, \ldots, e_{22})$. Then the received codeword has the form

$$R(x) = C(x) + E(x). \tag{7}$$

Suppose that e errors occur in the received codeword $R(x)$, and assume that $2t \le d - 1$. The decoder begins by dividing the received codeword $R(x)$ by the generator polynomial $g(x)$, i.e.

$$R(x) = q(x)g(x) + E(x). \tag{8}$$

Since α and α^3 are both roots of $g(x)$, one has

$$\begin{aligned} S_1 &\overset{\Delta}{=} E(\alpha) \\ S_3 &\overset{\Delta}{=} E(\alpha^3) \end{aligned} \tag{9}$$

where S_1 and S_3 are called the syndromes of the code.

The error-locator polynomial is defined by

$$L(z) = \prod_{i=1}^{v} (z - Z_i) = z^v + \sum_{j=1}^{v} \sigma_j z^{v-j}\ . \tag{10}$$

Here, Z_j for $1 \le j \le v$ are the locations of the v errors, i.e. $Z_j = \alpha^{r_j}$, where r_j locates the position of the error to be corrected and $v < t$.

III. Algebraic Decoding Algorithm

The (23, 12, 7) Golay code has a cyclotomic set $Q = \{1, 2, 3, 4, 6, 8, 9, 12, 13, 16, 18\}$. Therefore $g(x)$ has roots α^1, α^3, α^9. Then the error-locator polynomial in (11) can be found from the Newton identities. Consider now the following equations from the Newton identities with S_i^2 substituted for S_{2i}:

$$S_1 + \sigma_1 = 0 \tag{11}$$

$$S_3 + \sigma_1 S_1^2 + \sigma_2 S_1 + \sigma_3 = 0 \tag{12}$$

$$S_5 + \sigma_1 S_1^4 + \sigma_2 S_3 + \sigma_3 S_1^2 = 0 \tag{13}$$

$$S_7 + \sigma_1 S_3^2 + \sigma_2 S_5 + \sigma_3 S_1^4 = 0 \tag{14}$$

$$S_9 + \sigma_1 S_1^8 + \sigma_2 S_7 + \sigma_3 S_3^2 = 0 \tag{15}$$

In these five equations, only the syndrome S_1, S_3, and S_9 are known, since their subscripts belong to the cyclotomic set Q of the code. The unknown syndromes S_5 and S_7, which also occur in these equations, must be eliminated.

First, it is clear that $S_1 = \sigma_1$. Then S_5 and S_7 can be eliminated by substituting S_5 from (14) into (15) and S_7 from (15) to (16). These two substitutions yield the relation

$$\begin{aligned} S_9 + S_1^9 + \sigma_2[S_1 S_3^2 + \sigma_2(S_1^5 + \sigma_2 S_3 \\ + \sigma_3 S_1^2) + \sigma_3 S_1^4] + \sigma_3 S_3^2 = 0. \end{aligned} \tag{16}$$

Simplifying this equation and substituting σ_3 from (13) into (17), one finally obtains the result

$$\begin{aligned} S_9 + S_1^9 &= (S_3 + S_1^3)(\sigma_2^3 + \sigma_2^2 S_1^2 + \sigma_2 S_1^4 + S_3^2) \\ &= (S_3 + S_1^3)[(\sigma_2 + S_1^2)^3 + S_1^6 + S_3^2]. \end{aligned} \tag{17}$$

First, if $S_3 = S_1^3$, then from (18) one has $S_9 = S_1^9$. This happens only if there are one or no errors. Thus if $S_3 + S_1^3 \neq 0$, there is more than one error. In this case, the coefficient σ_2 can be found from (18). Simultaneously, σ_3 is found from (13) as follows:

$$\sigma_2 = S_1^2 + D^{1/3} \text{ and } \sigma_3 = S_3 + S_1 D^{1/3} \tag{18}$$

where

$$D = S_1^6 + S_3^2 + (S_9 + S_1^9) / (S_3 + S_1^3). \tag{19}$$

Since $2^{11} - 1 = 0$ *modulo* 23 and $GCD(3,\ 2^{11} - 1) = 1$, the cube root is unique.

When there are exactly two errors, $\sigma_3 = Z_1 Z_2 Z_3 = 0$, so that from (19) one has $D^{1/3} = (S_3/S_1)$. Therefore one has the decoding scheme for the Golay code given below. This is the same result given by Elia [2].

$$L(\mathrm{z}) = \begin{cases} 0 \\ \quad \text{no error if } s_i = 0 \text{ where } i \in Q. \\ z + s_1 \\ \quad \text{one error if } s_1^{\ 3} = s_3 \text{ and } s_1^{\ 9} = s_9. \\ z^2 + s_1 z + (s_1^{\ 2} + \dfrac{s_3}{s_1}) \\ \quad \text{two errors if } s_1 D^{1/3} = s_3 \\ \quad \text{where D is computed from (20).} \\ z^3 + s_1 z^2 + (s_1^{\ 2} + D^{1/3}) z + (s_3 + s_1 D^{1/3}) \\ \quad \text{otherwise three errors.} \end{cases} \tag{20}$$

IV. Soft Decoding of Four Errors for (23, 12, 7) Golay Code

It is convenient also to let the symbol P denote the decoding procedure of the (23, 12, 7) code. The look-up table of weight-7 codewords of the (23, 12, 7) QR codes is generated first. Let the set of the indices of nonzero components be denoted as $Ind(\cdot)$. Assume E_P obtained from P, be the error vector. If the weight of E_P is less than three, the error pattern can be corrected by the algebraic decoding algorithm. Otherwise, the weight of E_P is 3 and let $Ind(E_P) = \{i_1, i_2, i_3\}$. One exactly determines five codewords C_1, C_2, C_3, C_4, and C_5 with three indices of nonzero components i_1, i_2, and i_3 from the look-up table of weight-7 codewords. And let C_j'' be as follows:

$$C_i'' = \begin{cases} (C_i - E_p), & i = 1, 2, 3, 4, 5 \\ (E_p), & i = 6 \end{cases} \tag{21}$$

Consider the six vectors, i.e. C_1'', C_2'', C_3'', C_4'', C_5'', and C_6'', the “actually” occurred weight-4 or weight-3 error pattern is one of them. In the following, the error pattern, which is the most possible one among them, will be obtained by estimating the emblematic probability values of the six vectors.

In the soft decoding, one first estimates the individual bit-error probabilities p_0, p_1,…, p_{23} in a received word [3]. Then the emblematic probability values, denoted as $\hat{p}$, of the six possible patterns of occurred error are defined as follows:

$$\hat{p}_i = \begin{cases} \prod_{j \in Ind(C_i'')} p_j , & i = 1,2,3,4,5 \\ \prod_{j \in Ind(E_P)} p_j , & i = 6 \end{cases} . \quad (22)$$

Finally, the most possible error pattern, denoted as $C''_{\max}$, corresponding to the maximal $\hat{p}_i$, denoted as $\hat{p}_{\max}$, is obtained. The simulation result in additive white Gaussian noise (AWGN) shows that at least 93% and 99% of weight-4 error patterns occurred are corrected if a bit-energy to noise-spectral-density ratio E_b/N_0 are greater than 3 dB and 6 dB, respectively, and at least 96% and 99% of weight-3 error patterns occurred are corrected if a bit-energy to noise-spectral-density ratio E_b/N_0 are greater than 0 dB and 3 dB, respectively.

V. Conclusion

The proposed soft-decision decoder can be used to correct very large percent of patterns of quadruple errors, and almost all patterns of three errors, and all fewer random errors. Note also that as a bit-energy to noise-spectral-density ratio increases, the percentage of patterns of quadruple errors or three errors, which are decoded successfully, is improved. Furthermore, the method developed in this paper can be generalized to decode for more than four errors occurred.

References

[2] E. Prange, “Some Cyclic Error-Correcting Codes with Simple Decoding Algorithms,” Air Force Cambridge Research Center-TN-58-156, Cambridge, MA: 1958.

[2] M. Elia, “Algebraic decoding of the (23, 12, 7) Golay code,” *IEEE Trans. Inform. Theory*, vol. IT-33, no. 1, pp. 150-151, Jan. 1987.

[3] G. Dubney and I. S. Reed, “Decoding the (23, 12, 7) Golay code using bit-error probability estimates,” *Proceedings of IEEE Global Telecommunications Conference*, November 2005.

[4] I. S. Reed, “Statistical error control of a realizable binary symmetric channel,” Group Report 47.35, Lincoln Laboratory, Massachusetts Institute of Technology, Massachusetts, Nov. 1959.

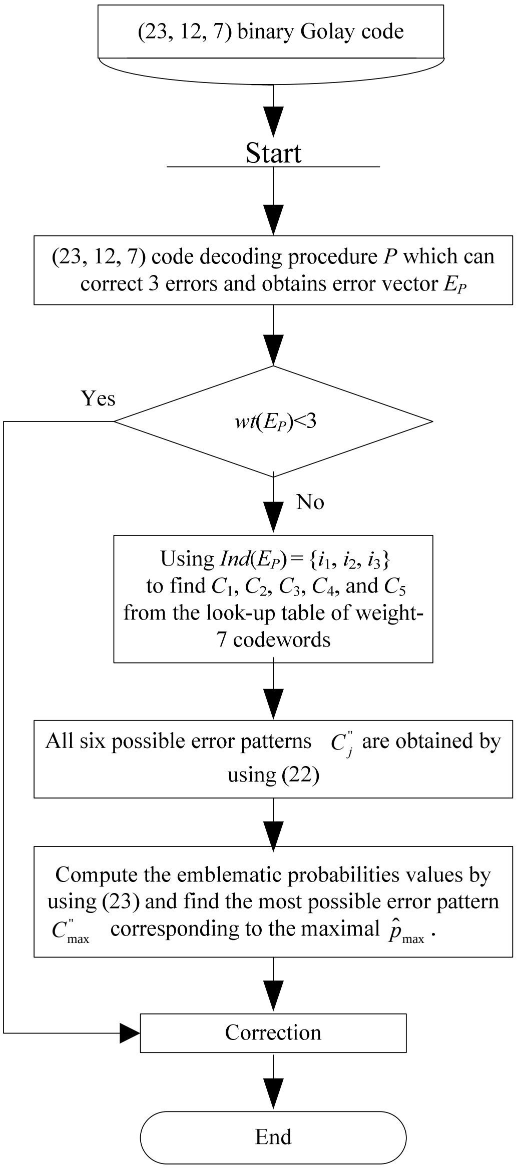
(23, 12, 7) binary Golay code
Start
(23, 12, 7) code decoding procedure P which can correct 3 errors and obtains error vector E_P
Yes
$wt(E_P)$<3
No
Using $Ind(E_P)=\{i_1, i_2, i_3\}$ to find C_1, C_2, C_3, C_4, and C_5 from the look-up table of weight-7 codewords
All six possible error patterns C_j'' are obtained by using (22)
Compute the emblematic probabilities values by using (23) and find the most possible error pattern $C''_{\max}$ corresponding to the maximal $\hat{p}_{\max}$.
Correction
End

Figure 1. Flowchart of the soft-decision decoder of (23, 12, 7) Golay code up to four errors.