QR コードについて

２年Ｄ組　小野　悠花　佐藤　悠都
佐藤　真生　住吉　慧斗
齋藤龍之介　須田　　怜

要約

QR コードは私たちの身の周りにたくさんある。例えば、LINE や店の商品、インターネット上の情報など、私たちの生活に欠くことのできないものである。また、バーコードよりも圧倒的に情報量が多く、ある程度の汚れならば復元可能である。その理論を支えているのは、代数学の符号理論や暗号理論で、2 進法が大きく関わっていることを知った。そこで、私たちは自分たちで情報を符号化し、誤り訂正符号を求め、QR コードの作成を試みた。

１．研究の動機と目的

膨大な情報の中から自分が必要としている情報を見つけ出すのは困難である。そんなとき私たちの生活を支えているものとしてバーコードや QR コードがあげられる。私たちは特に QR コードに興味をもった。QR コードは、スマートフォンで簡単に読み込めるためバーコードよりも日常生活において必要不可欠であり、数学的要素が含まれているため自分たちでも作成できないか考えたことが本研究の動機である。

また、私たちは QR コードの特徴として次の 3 点に着目した。第 1 に QR コードを作成する過程において 2 進数や 16 進数の計算が活用されていることがある。第 2 に折り曲げたり、汚れ、擦れおよび破れなどによってコードの一部が読めなくなった場合でもある程度なら自動的に誤りを訂正して、読むことができることがある。第 3 に QR コードはバーコードと比較すると 10 倍から 100 倍の情報量を保存することができ、スマートフォンでも読み取れるため、バーコードのように読み取り機器にコストをかけなくても済むという以上の 3 点に着目した。これらを踏まえて自分たちで QR コードを作成し、符号化した情報を読み取ることが本研究の目的である。

２．QR コードの作成に必要な知識

QR コードを作成するにはさまざまな知識が必要となる。その中でも QR コードを塗る際に必要となる知識について説明する。

第 1 に、「誤り訂正コード」について説明する。「誤り訂正コード」とは、誤り、つまり汚れや損傷などを検出し、訂正を行う符号のことである。QR コードを作る際に「真実為原」を符号化したデータのコードの後に並べるコードである。

第 2 に、「RS 符号」について説明する。「RS 符号」は、複雑な計算手順が必要であるが、誤り訂正能力が高く、連続した複数ビットの訂正に強い符号である。本研究では、「誤り訂正コード」を求めるために、総コード 26、データコード数 13、誤り訂正数 6 のときの「RS 符号」をいう。

第 3 に、「BCH 符号」について説明する。「BCH 符号」は、複雑な計算手順が必要なく、符号化、暗号化されたデータを元のデータに復元することができる符号である。主に QR コードの形式情報に用いられる。

第 4 に、「ビット」について説明する。「ビット」とは QR コードの作成において、最小の情報量で表されるものである。主に（01101100）などの「0」や「1」のような単位である。

第 5 に、「シフト JIS 漢字コード」について説明する。日本で決められた漢字 1 つ 1 つに割り振られた 16 進数の符号のことであり、「10」は「A」、「16」は「G」などのように対応している。例えば、「真」は(905E)で表され、$9\times16^3+0\times16^2+5\times16^1+14\times1$ の意味である。

第 6 に、「漢字コード」について説明する。「漢字コード」は（1000）で定められており、QR コードに入力される情報に漢字が含まれている場合に活用されるものである。

３．QR コードの作成の基本設定

QR コードを作成する際に、どのような情報を用いて、どのような誤り訂正レベルにするかを考えた。まず、QR コードで表す情報として由利高校の校訓である「真実為原」を採用した。理由としては、情報量を多くすると、その分自分たちで行う計算量が増えるため、できるだけ最小の情報量で自分たちに関わりの深い 4 文字に設定した。次に、誤り訂正レベルについて考えた。今回は、参考文献の誤り訂正レベルが Q（全体の 25％復元可能）だったため作成する QR コードも誤り訂正レベル Q にした。理由としては、完成させせることを第一の目的としたためである。

４．QR コードの作成

QR コードの作成手順として、幸山(2008)を参考にして、本研究でも次のように 3 つの段階に分けて QR コードを作成した。

STEP1　符号化

STEP2　マスク処理

STEP3　モジュールの配置

STEP1 の「符号化」とは、「真実為原」をシフト JIS 漢字コードに読み替える段階である。「真」は(905E)、「実」は(8EC0)、「為」は(88D7)、「原」は（8CB7）となる。次に、それぞれから 16 進数のもとで（8140）を引き、その上位バイトに(C0)を掛ける。そして、その結果と下位バイトを足す（図 1)。最後に、その値を 13 ビットの 2 進数に変換すると以下のように圧縮変換される。

図 1　符号化

真	実	為	原
905E	8EC0	88D7	8CB4
−8140	−8140	−8140	−8140
F1E	D80	797	B74
F	D	7	B
×C0	×C0	×C0	×C0
B40	9C0	540	840
B40	9C0	540	840
+　1E	+　80	+　97	+　74
B5E	A40	5D7	8B4

「真」は(0101101011110)、「実」は(0101001000000)、「為」は(0010111010111)、「原」は(0100010110100)

STEP2 の「マスク処理」について述べる。「マスク処理」とは、データを配置したときに一方の色が多すぎたり、位置検出パターン(QR コードシンボルの右上、左上、左下に配置されている模様)に似た模様があると読み込みにくく

次に「漢字コード」(1000)、「文字数コード」(000001

の順に並べる。最後に、「埋め草コード」（11101100）と（00010001）を交互に並べる。実際に並べた

図 2　圧縮変換

[10000000][01000101][10101111][00101001]

漢字コード　文字数コード　真

[00000000][10111010][11101000][10110100]

実　為　原

[00000000][11101100][00010001][11101100]

終端パターン　埋め草ビット　埋め草コード語

[00010001]

コードは図 2 である。

ここで言う「文字数コード」とは 1 バイトに収納する文字数を指示するものであり、漢字の場合は 1 バイト 8 ビットになるようにする。「終端パターン」とは、符号化したデータの終わりにつけて、データの終わりを指示するものである。「埋め草コード」とは（11101100）と（00010001）を交互に 8 ビットの 13 バイトになるように並べるコードであり、空いた部分を埋める空コードでもある。

次に、誤り訂正コードを求めるために図 2 のデータをαのべき指数の形に変換する。そのためにガロア体の考え方を用いる。

ガロア体とは、整数を素数で割ったときの余りの集合である。QR コードは原始多項式

$F(x)=x^8+x^4+x^3+x^2+1$ の係数を示す 100011101 を持つ $GF(2^8)$ を用いて計算されている。まず、x をαとする。$F(\alpha)=\alpha^8+\alpha^4+\alpha^3+\alpha^2+1=0$ と決められており、それを変換して $-\alpha^8=\alpha^4+\alpha^3+\alpha^2+1$ となる。GF(2) 上の数であるから 1+1=0 は -1=1 で表され、$-\alpha^8=\alpha^8$ となる。よって $\alpha^8=\alpha^4+\alpha^3+\alpha^2+1$ となる。α^8 以上のべきはすべて α^0～α^7 の多項式として表現できる。この多項式を並べて書くと(10100110)のような「0」と「1」を使った 8 ビットの「ベクトル表現」ができる。そして、$GF(2^8)$ の 0 以外の元(0 と 1)は全てαのべき指数として表現できたため、$GF(2^8)$ のすべての元と α^0 から α^{225} （2^8=256 を 0 から数えるものとする）の整数を対応させることができる。これにより、1 コード語と $GF(2^8)$ の α^0～α^{255} の「べき表現」を α^7＝10000000 のように対応させることができる。

次に図 2 のデータを多項式 F(x)として表現する。誤り訂正レベルを Q に設定したため総コード語数が 26 となり、x の 25 次から 13 次の和で表される。その際の係数が図 2 のコードとなる。

データの 8 ビットを 1 コードとしたものを係数にして、F(x)を求めると次のようになる。

$$F(x)=10000000x^{25}+01000100x^{24}+10101111x^{23}+00101001x^{22}+00000000x^{21}+10111010x^{20}+11101000x^{19}$$
$$+10110100x^{18}+00000000x^{17}+11101100x^{16}+00010001x^{15}+11101100x^{14}+00010001x^{13}$$

そして、先名(2011)の「$GF(2^8)$ のべき表現とベクトル表現」の表をもとに「べき表現」α^x（ただし x は 0 から 255 の数)に直すと α^7＝10000000、α^{221}＝01000100 となり他も同様に変換できる。F(x)を変換した形にすると以下のようになる。

$$F(x)=\alpha^7x^{25}+\alpha^{221}x^{24}+\alpha^{97}x^{23}+\alpha^{147}x^{22}+\alpha^0x^{21}+\alpha^{57}x^{20}+\alpha^{11}x^{19}+\alpha^{20}x^{18}+\alpha^0x^{17}+\alpha^{122}x^{16}+\alpha^{100}x^{15}+\alpha^{122}x^{14}$$
$$+\alpha^{100}x^{13}$$

次に誤り訂正コードを作る。誤り訂正レベルによって生成多項式が定められている。本研究では訂正コードの語数が 13 なので決められた生成多項式は以下のようになる。

$$G(x)=(x-1)(x-\alpha)(x-\alpha^2)(x-\alpha^3)(x-\alpha^4)(x-\alpha^5)(x-\alpha^6)(x-\alpha^7)(x-\alpha^8)(x-\alpha^9)(x-\alpha^{10})(x-\alpha^{11})$$
$$(x-\alpha^{12})$$
$$=x^{13}+\alpha^{74}x^{12}+\alpha^{152}x^{11}+\alpha^{176}x^{10}+\alpha^{100}x^9+\alpha^{86}x^8+\alpha^{100}x^7+\alpha^{106}x^6+\alpha^{104}x^5+\alpha^{130}x^4\alpha^{218}x^3+\alpha^{206}x^2+\alpha^{140}x$$
$$+\alpha^{78}$$

F(x)÷G(x)の割り算を行い、余りを求める。初めの項が α^7x^{25} なので α^7x^{12} と F(x)を掛けると $G_1(x)$ が求められる。

$$G_1(x)=\alpha^{7}x^{25}+\alpha^{81}x^{24}+\alpha^{159}x^{23}+\alpha^{183}x^{22}+\alpha^{107}x^{21}+\alpha^{93}x^{20}+\alpha^{107}x^{19}+\alpha^{113}x^{18}+\alpha^{111}x^{17}+\alpha^{137}x^{16}+\alpha^{225}x^{15}$$
$$+\alpha^{213}x^{14}+\alpha^{147}x^{13}+\alpha^{85}x^{12}$$

次に $G_1(x)$のべき表現をベクトル表現に直すと以下のようになる。

$$G_1(x)=10000000x^{25}+11100111x^{24}+01110011x^{23}+11000100x^{22}+01101000x^{21}+10110110x^{20}+01101000x^{19}$$
$$+00011111x^{18}+11001110x^{17}+10011110x^{16}+00100100x^{15}+11110010x^{14}+00101001x^{13}+11010110x^{12}$$

これを F(x)に足すと最高次の項が消えて、次数が 1 つ減る。

$$F(x)+G_1(x)=10100010x^{24}+11011100x^{23}+11101101x^{22}+01101000x^{21}+00001100x^{20}+10000000x^{19}$$
$$+10101011x^{18}+11001110x^{17}+01110010x^{16}+00110101x^{15}+00011110x^{14}+00111000x^{13}$$
$$+11010110x^{12}$$

同様な計算をくり返し行い、次数を下げていくと、最終的に以下の式が求められる。

$$00000100x^{12}+11011100x^{11}+11001000x^{10}+01011110x^{9}+11100101x^{8}+01000100x^{7}+10001000x^{6}$$
$$+01010001x^{5}+01001111x^{4}+11011000x^{3}+00101110x^{2}+00100111x+00001101$$

最高次数が 12 次のため、これ以上割れない。つまり誤り訂正コードは以下のようになる。

00000100 11011100 11001000 01011110 11100101 01000100 10001000 01010001
01001111 11011000 00101110 00100111 00001101

次に、QR コードの読み込みを確実にし、シンボルにおける白黒をバランス良く配置する操作を行うため「マスク処理」を行う。この処理を行うためにマスクパターンを用いる。マスクパターンは、条件によって 8 種類の形がある。この 8 種類のマスクのどれを用いるかはマスク処理後の状態を評価することによって行われるが、参考にしたものが参照子 011 だったため Mask011 を用いた。これを次に

Mask011={01100001}{10000110}{00011000}{10010010}{01001001}{00100100}{10000110}{00011000}
{01100001}{00100100}{10010010}{01001001}{00011000}{01100001}{10000110}{00010001}
{10000110}{01001001}{00100010}{01001001}{00100100}{10010010}{00011000}{00100100}
{10010010}{01001001}

形式情報コードは、誤り訂正レベルやマスクパターン、誤り訂正コードをもとに計算して 15 ビットで表される数である。縦方向と横方向の 2 回塗る重要なコードである。今回用いた形式情報コードは 110110100110111 である。誤り訂正レベルが Q であることから、最初の数は 11 となる。そして、マスクパターン Mask011 を用いたことから次の数は 011 となる。それを続けて 5 桁を表している。しかし、残りの 10 桁については、計算の仕方や本質的な意味を理解できなかったため、池田(2007)の形式情報コードを用いた。

最後に「マスク処理」を行う。圧縮変換で求めた図 2 の 13 のコードと STEP2 の段階で求めた誤り訂正コードの 13 のコードの合計 26 コードを続けて並べる。そして、上記の mask011 のコードを 2 進数のもとで足し算を行うと最終的なコードを求めることができる。実際には mathematica というソフトを使って計算した。以下に示したコードが QR コードに配置される。

11100001　11000011　10110111　10111011　01001001　10011110　01101110　10101100
01100001　11001000　10000011　10100101　00001001　01100101　01011010　11011001
11011000　10101100　01100110　11000001　01110101　11011101　11000000　00001010
10110100　01000100

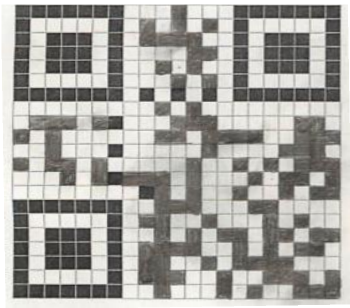
STEP3 の「モジュールの配置」とは、並べたコードをもとにモジュールと呼ばれるデータ容量に応じた型番に色を塗り、形式情報コードを配置するという段階である。データ領域にまず右下から、8 つのセルに 1 コード語を割り当てていく。STEP2 で求めたコードを右下から塗り始める。1 を黒、0 を白とする。それぞれ割り当てが決まっており、図 3 の左図の場合は下から上に、右図の場合は上から下に塗る時の順番である。

STEP2 でマスク処理されたコードと形式情報コードを塗ると図 4 のように「真実為原」の QR コードができる。

図 3　QR コードの塗り方

8	7		2	1
6	5		4	3
4	3		6	5
2	1		8	7

図 4　作成した QR コード



５．結果と今後の課題

1 回目に作成した QR コードは失敗したが、2 回目に作成したコードは成功した。

最初に作成したものは QR コードとして携帯に認識はされたが、符号化したはずの「真実為原」という文字は読み取ることができなかった。その原因として、「真実為原」の「真」の漢字コードが間違っており、また「真」の文字を含むほかの 3 文字の途中計算も正確ではなく間違っていたことにより、符号化したコードが本来導くはずだったコードと違ったため読み取りに失敗したということが分かった。

これらのことから「真」の漢字コードを訂正し、すべての文字の計算を最初からやり直した。QR コードを構成し直した結果、「真実為原」という文字を携帯で読み取ることができた。しかし、計算により本来のコードを導くことができたが、それらの計算過程で行われている数学の深い内容や使用されている単語の詳細な意味までは理解することができなかった。以上のことから、今後は単語の詳細な意味を理解した上で、「真実為原」よりも複雑な情報をコード化していきたい。また、今回は参考文献の誤り訂正レベルが Q であったため私たちもそれに従ったが、次は参考文献に載っていなかった誤り訂正レベルを使って QR コードの作成に努めたい。

６．引用・参考文献

池田和興(2007).「例題が語る符号理論」.共立出版.

竹生修己(2010).「よくわかるバーコード・二次元シンボル」.オーム.

Darel,W,Hardy　Carol,L,Walker.(2005).「応用代数学入門」.ピアソン・エデュケーション.

先名健一(2011).「符号理論入門」.森北出版.

幸山直人(2016).富山大学公開講座 2008「QR コードを作ろう」.
http://kouyama.sci.u-toyama.ac.jp/main/etc/2008/koukai2008/(閲覧日 6 月 30 日)