5. 文字・図形・付加音符号

正会員 磯 部 忠[†], 正会員 杉 森 吉 夫^{††}
 正会員 荒 木 洋 哉^{††}

1. ま え が き

番組の表現に必要な提示機能を実現し、番組データを効率的に伝送するため、文字、図形および付加音の各機能に則して独自の符号化を行っている。ここではおのこの符号化の概要を紹介する。

2. 文字の符号化

文字の符号化では漢字、カナ、英数および記号で綴られる文章、モザイク素片の組合せで描かれるモザイク図形、また、文字発生の ROM に含まれていない外字、外字伝送を応用した DRCS 図形などを取扱う。符号体系は 8 単位の基本符号体系であり、C0 および C1 の領域に制御符号集合を設定し、GL および GR の領域に指示、呼出しの制御により文字符号集合のいずれかを設定して、符号化する。

2.1 文字符号集合

文字符号集合の種類と概要を表 1 に示す。文字符号は番組の表現に必要な文字種を充分カバーすること、基本的な符号体系に準拠すること、伝送効率を高くすることなどの条件により、8 単位の 2 バイト符号と 1 バイト符号とを混用する。

一般の漢字カナ混じり文における漢字の比率は 25~35% であり、カナ、英数などを 1 バイト符号で伝送することにより効率を高めることができ、1 文字あたりの平均符号長が 13 ビット程度となる。我が国のビデオテックスであるキャプテンシステム¹⁾のように、2 バイト符号のみを使用する場合に比較して伝送効率が 20% ほど高い²⁾。

文字発生器に記憶されていない文字（外字）を伝送

表 1 文字符号集合

文字符号集合	符号長 (バイト)	字種	概 要
漢 字 系 集 合	2	3662	第 1 水準漢字 ⁽¹⁾ (2965) 非漢字 ⁽¹⁾ (524) 文字放送用追加記号 (173)
		3388 ⁽²⁾	DJCS (第 2 水準漢字 ⁽¹⁾)
英 数 集 合	1	94	ローマ字(大, 小) (52) 数 字 (10) 記 号 (32)
カタカナ集合	1	94	カタカナ (86) 記 号 (8)
ひらがな集合	1	91	ひらがな (83) 記 号 (8)
モザイク集合	A	94	矩形モザイク (65) 三角形, 台形モザイク (29)
	B	59	特殊な形のモザイク
	C	63	ノンスペーシング分離モザイク
	D	64	ノンスペーシング三角形モザイク
1バイト外字集合	1	94 ⁽²⁾ ×15	DRCS 1~DRCS 15
2バイト外字集合	2	8836 ⁽²⁾	DRCS 0
マ ク ロ 集 合	1	94	マクロ定義によりマクロ文 定義 6/0~6/15 にデフォルトマ クロ文

⁽¹⁾ JIS C 6226 情報交換用漢字符号系。

⁽²⁾ 実際の字数は DRCS, DJCS データユニットによりパターンデータを送った文字である。

する場合は、その文字符号を外文字符号集合 (DRCS) の中に一時的に定めて、また、外字のパターンデータを事前に DRCS データユニットで伝送する。外文字符号集合の字種の数 DRCS 図形により表示領域の大部分を表現できる量が確保されている。

漢字系集合のうち、第 2 水準漢字の符号領域は 1 つの外文字符号の領域 (DJCS と呼ぶ) とし、漢字符号 (第

[†] NHK 放送技術研究所

^{††} 日本テレビ放送網株式会社 技術局

"5. Coding of Characters, Pictures and Sound" by Tadashi Isobe (NHK Science and Technical Research Laboratories, Tokyo), Yoshio Sugimori and Yosai Araki (Nippon Television Network Co., Ltd., Tokyo)

2水準)を外字符号として使い、そのパターンデータはDJCSデータユニットにより伝送する。将来、第2水準の文字種が文字発生器に収容される場合を配慮した結果である。

マクロ符号集合はマクロ定義により一時的に定義される符号列(マクロ文)を1つの文字符号により代表して使う符号集合であり、効率的な符号化に用いる。マクロ符号集合の一部には(6/0~6/15)、デフォルトマクロ文が定義されており、文字符号集合の指示、呼出しを効率的に行うことができる。

2.2 制御符号集合

提示に関連する基本制御符号集合C0と追加制御符号集合C1の制御機能の概要を表2に示す。

C0の符号割付けは、JISおよびISOの情報交換用符号のための制御符号に準拠している。また、C1ではキャプテンプロトコルと共通する機能については同一の符号を割付けている。

動作位置制御は表示領域における表示区画の位置を制御する。指定動作位置前進(PAPF)ではパラメータ N_1 で前進回数を指定し、動作位置指定(APS)では N_1 で行方向の位置、 N_2 で字方向の位置を指定する。

文字サイズは9種類のうち使用頻度の高い標準、中形、小形サイズについてはおのおのに符号を割付け、その他は指定サイズ(SZX)のパラメータ N_1 により指定する。

書式選択(SWF)はパラメータ N_1 により5種類の書式を指定する。ページデータヘッダで指定される初期表示書式により、その後、選択する書式が制約される。

表示画の合成の制御は表示領域へ文字、図形を書き込む動作を制御する書込みモード変更(WMM)と、表示区画ごとのパターンデータを合成する合成制御(CCC)とがある。

着色制御は表示区画ごとの前景色または背景色を全輝度あるいは半輝度の7色と黒または透明に指定することができる。使用頻度の高い前景色全輝度の着色はおのおのを符号割付けし、その他の前景色半輝度、背景色全輝度、背景色半輝度は色指定(COL)のパラメータ N_1 により指定する。これらの符号によりカラーマップアドレスも指定する。

表示属性のフラッシング、コンシール、パターン極性および囲み制御はおのおの1つの制御符号を割付け、パラメータにより制御内容を指定する。アンダラインについては、ヨーロッパのテレテキスト、ビデオテキストの方式規格に合わせ、開始と終了の2符号を

表2 制御符号

分類	制 御 符 号				
	名 称	略称	パラメータ	集合	符号 ^{注)}
書式	動作位置制御	動作位置 後退	APB	—	C0 00/8
		動作位置 前進	APF	—	C0 00/9
		動作位置行前進	APD	—	C0 00/10
		動作位置行後退	APU	—	C0 00/11
		動作位置復帰改行	APR	—	C0 00/13
		動作位置指定	APS	N_1, N_2	C0 01/12
		指定動作位置前進	PAPF	N_1	C0 01/6
	文字サイズ	小形サイズ	SSZ	—	C1 08/8
		中形サイズ	MSZ	—	C1 08/9
		標準サイズ	NSZ	—	C1 08/10
		指定サイズ	SZX	N_1	C1 08/11
表示属性	書式選択		SWF	N_1	CSI N_1 02/0 05/3
	合成	書込みモード変更	WMM	N_1	C1 09/4
		合成制御	CCC	N_1	CSI N_1 02/0 05/4
	着色	前景色 黒	BKF	—	C1 08/0
		前景色 赤 全輝度	RDF	—	C1 08/1
		前景色 緑 全輝度	GDF	—	C1 08/2
		前景色 黄 全輝度	YLF	—	C1 08/3
		前景色 青 全輝度	BLF	—	C1 08/4
		前景色 マゼンタ 全輝度	MGF	—	C1 08/5
		前景色 シアン 全輝度	CNF	—	C1 08/6
		前景色 白 全輝度	WHF	—	C1 08/7
		色指定	COL	N_1	C1 09/0
		パレット指定	COL	02/0 N_2	C1 09/0
		縮小着色 開始	MCS	—	C1 08/12
		縮小着色 終了	MCT	—	C1 08/13
	属 性	フラッシング	FLC	N_1	C1 09/1
		コンシール	CDC	N_1, N_2	C1 09/1
		パターン極性	POL	N_1	C1 09/3
		囲み	HLC	N_1	C1 09/7
		アンダライン終了	SPL	—	C1 09/9
		アンダライン開始	STL	—	C1 09/10
特殊制御	拡張	制御シーケンス導入	CSI		C1 09/11
		空 白	NUL	—	C0 00/0
	その他	取消し	CAN	—	C0 01/8
		ベ ル	BEL	—	C0 00/7
		繰返し	RPC	N_1	C1 09/8
		マクロ指定	MACRO	N_1	C1 09/5
		時間制御	TIME	N_1, N_2	C1 09/13

注) 8単位符号の $b_8 \sim b_5$ を列番号、 $b_4 \sim b_1$ を行番号とし、列番号/行番号と符号を表わしている。

割付ける。アンダライン制御はモザイク図形の場合、 2×3 の格子でモザイク素片を分ける分離化処理を指定する。

特殊符号のスペース(SP)は表示区画を背景色で塗りつぶし、デリート(DEL)は前景色で塗りつぶすことを指定する。

2.3 文字の符号化における図形表現

(1) モザイク図形

モザイク図形素片を組合せて構成するモザイク図形は、図形素片の文字符号で伝送し、受信機では文字発生器を使って再構成する。モザイク図形*集合は表1に示す4種類があり、モザイク集合CおよびDはノンスペーシング文字で次に続くモザイク集合A、BまたはスペースとOR合成して使用する。アンダライン制御の符号で指定される分離化処理は合成したモザイク図形に対して適用される。

(2) DRCS 図形

DRCS 図形は外字の伝送の図形への応用である。復号が容易であり、1画面中で繰返して用いる記号的な図形に有効である。表示区画を単位としてパターンデータを伝送し、受信機のRAMに記憶させ、その後、外文字を伝送してそのパターンデータを組合せてDRCS 図形を再構成する。パターンデータはDRCS データユニットにより、外字の符号列に先立って伝送する。複数の外字を使用する場合、先頭の外文字符号から続いた符号を用い、パターンデータを符号の順に続けて伝送する。

(磯部)

3. 図形の符号化

図形伝送の方式にはフォトグラフィック伝送、ジオメトリック伝送の2種がある³⁾。

前者は我が国の図形伝送方式の特色ともいえ、表示画面の書き込み領域(フィールド)内にドットパターンデータを書込むことによって表示を行うもので、画面での図形の現われ方は上下方向にワイプ状に表示される。この方式はパターン伝送方式の表示を機能的に包含しており、後述の1層フォトグラフィック方式における非圧縮モードがこれに相当する。

後者は外国(北米圏)の図形伝送の主体機能である

ジオメトリック符号セットのサブセットを採用したもので、画面の表示の現われ方が独特で、図形記述符号の送出順に図形が生成されるので、演出面の各種の可能性をもっている。

我が国の方式における情報の伝送速度⁴⁾は実質的には1H 当たり約10 kb/s 強であるので、受信画面の図形生成速度や伝送番組量の制約から、一般的には図形画面1枚当たりのデータ量は2 kByte 程度が運用上の目安になる。この点から図形伝送の運用においては図形の精細さの選択や表示画面上の図形領域の設定、あるいは伝送方式の選択について工夫が凝らされるものと考えられる。

3.1 フォトグラフィック伝送

表示画面内に図形を表示する領域(フィールド)を定め、その中を走査して得られる情報を圧縮(モデファイドハフマン符号(MH 符号)によるランレングス表現)または非圧縮で符号化する。この符号化は1層のパターンデータと小区画(ブロック)ごとの属性データを伝送するブロック着色(2色塗り分け)式の1

表3 フォトグラフィック伝送とジオメトリック伝送

	フォトグラフィック		ジオメトリック
	1 層	多 層	
着色の種類	レベル A: 8 色 2 階調 レベル B: 4096 色 中 32 色	カラーマッピングによる 4096 色 中 32 色	同左
着色の単位	カラーブロックによる最小 4×4 ドット	画素単位	同左
着色の方法	2 色塗り分け 着色	ドット単位着色	同左
主体となる情報	パターンデータ	同左	図形記述命令 (描画コマンド)

注) ①～④はフィールド内のユニットの走査手順を示す。

図1 ブロック、ユニットおよびフィールドの関係

* 粗い図形であるが伝送の効率がよく、ヨーロッパのテレテキストで実用されている。

層フォトグラフィックと複数層のパターンデータを伝送することでドット着色を可能にした多層フォトグラフィックがある。図形の書き込み領域をフィールド、図形を走査し伝送する単位領域をユニット、1層フォトグラフィックで属性を制御する小区画をブロックと呼ぶ。

フィールド内の各ユニットの走査順序とユニット内の各ブロックまたは各ドット（論理画素）の走査順序は、走査方向の指定により定められる。主走査は左から右へ走査されるが、副走査は後述のフィールド制御部の指定によって表示効果を高めるための上下方向が選択できる。

（1） 1層フォトグラフィック

データユニットの構成を図2に示す。ユニットデータ部は、バイト単位の変長の図形データで、フィールドを構成するユニットの数だけ送出される。図中のUC1によってパターンデータの圧縮・非圧縮、属性データ、パターンデータの有無などが指定できるから、図形によってデータ量が最少になるように運用することができる。属性データは着色、コンシール、フラッシングを指定するもので、前述のように送出を省略することもできる。パターンデータは圧縮のモードを選んだ場合は、前述のように各ドットの1または0の連続長（ラン長）に符号化される。符号化はMH符号であり、31以下のラン長はターミネイト符号のみで、32以上のラン長はメークアップ符号+ターミネイト符号で表現される。このようにラン長で表現すれば、一般的な図形では非圧縮モードに比してデータ量は3分の1ないし4分の1程度に減少する⁵⁾。

（2） 多層フォトグラフィック

フィールド制御部では多層書き込みモード種別、層数などを指定する。

ユニットデータ部はバイト単位の変長の図形データで、各ユニットごとに各層のパターンデータを層数の小さい順に配列したものである。画面上の各画素は、カラーlookupアップテーブルのマッピングアドレ

スを層別伝送によって指定されることになる。第1層がLSB、最終層がMSBの値を示す。

3.2 ジオメトリック伝送

ジオメトリック図形とは、図形記述命令(PDI)の組合せにより描画を行うもので、レベルBの機能である。使用する符号集合は基本制御符号集合、追加制御

表 4 ランレングス符号

ターミネイト符号			メークアップ符号		
ラン長	0 ラン	1 ラン	ラン長	0 ラン	1 ラン
0	00001	00000100	32	1110	101100
1	0100	01	64	11011	10101010
2	0111	11	96	111100	10101011
3	1010	100	128	111110	10101100
4	1100	0001	160	1111110	10101101
5	00100	00100	192	00111110	10101110
6	00110	10100	224	00111111	10101111
7	01100	001010	256	01010110	10110100
8	01101	001100	288	01010111	10110101
9	000100	0000100	320	01011110	10110110
10	000110	0011111	352	01011111	10110111
11	000111	1010100	384	10111110	10111100
12	001010	00000101	416	10111111	10111101
13	001110	00000110	448	11111110	10111110
14	010100	00000111	480	11111111	10111111
15	10110	00001010			
16	11010	001110			
17	0000010	00001011			
18	0000011	00001100			
19	0001011	00001101			
20	0010110	00001110			
21	0010111	00001111			
22	0011110	00101100			
23	010110	00101101			
24	101110	00101110			
25	0101010	00101111			
26	0101110	00110100			
27	00010100	00110101			
28	1011110	00110110			
29	00010101	00110111			
30	1111010	00111100	EOL	100	101110
31	1111011	00111101	EOU	00000000

図 2 1層フォトグラフィックの構成

符号集合, PDI 符号集合, ジオメトリックマクロ符号集合である. 表 5 に描画コマンドを示す.

PDI のコマンドは描画命令 (オペコード) とその内容, 量などを示すオペランドで構成される.

表 5 PDI 描画コマンド

列 行	2	3
0	RESET	RECT OUT LINED
1	DOMAIN	RECT FILLED
2		SET & RECT OUT- LINED
3	TEXTURE	SET & RECT FILLED
4	POINT SET ABS	POLY OUT- LINED
5	POINT SET REL	POLY FILLED
6	POINT ABS	SET & POLY OUT- LINED
7	POINT REL	SET & POLY FILLED
8	LINE ABS	
9	LINE REL	
10	SET & LINE ABS	
11	SET & LINE REL	
12	ARC OUT- LINED	SET COLOR
13	ARC FILLED	
14	SET & ARC OUT- LINED	SELECT COLOR
15	SET & ARC ILLED	BLINK

この伝送方式は, 画面に幾何学的な図形を表示する際には少量のデータで伝送できる特長があるが, 複雑な図形を伝送する場合はドットパターンを伝送するフォトグラフィック伝送方式より効率が悪くなる.

(杉森)

4. 付加音の符号化

付加音の符号化では 2 章で述べられた付加音発生機能の機能为了满足するための音高, 音長, 音色, 音の強さその他の制御符号を取扱う. 符号体系は 8 単位の基本符号体系で, 2 バイトセットの音高音長符号集合と, 1 バイトセットの制御符号集合, 音強修飾符号集合より成り, 3 章に述べられた指示呼出しによって, おおの GL, C1, GR 集合に呼出されて符号化が行われる. 音色と音の強さは制御符号のパラメータとして指示される. 付加音制御符号のパラメータは 10 進数表現の可変長符号で, 数値を扱う場合は上位桁から順に並べ, かつ異なる意味を持つパラメータとパラメータの最後には区切り, 終端符号を置く.

付加音の符号化はビデオテックスのキャプテン PLPS のメロディー符号化¹⁾と完全整合 (レベル B を除く) がとられているが, 今後, 他の PLPS への移植等の可能性も考え, ISO に準拠したパラメータ方式の採用や, 7 ビット系符号への対応を考慮して定められたものである.

4.1 音高音長符号集合

通常の楽譜表現では音高と音長はセットで取扱われていること, および前述の 7 ビット系への配慮から, 音高と音長を組合せた 2 バイトセットとなった. 当初は 1 バイトでおおの GL, GR に割当て, 連続する場合の省略によるデータの制限なども検討されたが, 前述の理由の他, 大幅なデータ制限には至らないこと, 処理を複雑化することなどで, 2 バイトセットとして取扱うこととなった.

音高は最も標準的な 12 音平均律による音階で表すものとし, 代表的なピアノの音域 88 音と, リズムパートで用いる不定ピッチ音を表す符号 (IP), および休符を表す符号 (RST) の合計 90 個が割当てられた. 音の高さの周波数は国際的にも標準的な A₄ を 440 Hz と定められた.

音長は楽譜の場合, 通常 1/2" で分割した値で表現し, かつ, その単位の絶対値を規定するために楽譜の冒頭に指定した長さの音符の 1 分間における拍数を指示する方法がとられている. 付加音符号化では, このような楽譜における表現方法の長所と, デコードにお

ける処理をできるだけ単純にするため、音長の最小単位（基本時間） T_u を定め、 $1 \sim 64 T_u$ の64種が音長符号として割当てられている。64種としたのは、64分音符までの分解能があれば一般の音楽では充分であること、かつ楽譜ではタイなどで結合されて行われている $1/2^n$ 系列以外の長さの表現を直接表現できるため、符号化効率が良いことによるものである。64 T_u 以上の長い時間の表現は後述の長音長符号により可能である。 T_u の絶対時間は付加音開始符号のパラメータであらかじめ指定される。

4.2 付加音制御符号集合

付加音制御符号集合には表6に示す16種の制御符号が割当てられている。*印のついたものはレベルBでのみ用いられるものである。以下に主な制御符号についての概略を示す。

(1) 付加音開始・終了符号（SMC, EMC）

SMCは付加音データの開始を示す符号で、パラメータとして符号化レベル識別と、音長の基本単位 T_u の時間を示す数値をとまう。EMCは付加音データの終了を示す符号で、両者に囲まれた符号列が1つの付加音（一般には1曲分）として取扱われる。音長の基本単位の時間を示す第2パラメータは1～3バイトで表す。単位はmsである。

(2) パート開始・終了符号（SMP, SRP, EPT）

SMPはメロディーパートの、SRPはリズムパート

表6 付加音制御符号

列 行	08	09
0	SMC*	SLV*
1	SMP*	
2	SRP*	
3	SCP**	
4	EMC	
5	EPT	
6		
7		
8	LBL*	LRT*
9	JMP*	
10	RPT*	TIE**
11	BRA*	
12	CTM*	MNB**
13		
14		
15		TEP

*はパラメータ付符号

の、共に開始を示す符号で、前者はメロディー楽器の、後者はリズム楽器の音色パラメータをとまう。EPTはパートの終了を示す符号で、この両者に囲まれた符号列が1つのパート内データである。音色パラメータは表8に（一部）を示す $P_1 \sim P_4$ の4バイトから成り、それぞれ楽器分類に従って上流（ P_1 ）から下流（ P_4 ）へと定められている。音色の指定を楽器名にしたことは、電子楽器で最も一般的に用いられていること、音色のイメージをとらえ易く、かつデコードの性能向上に対する制約を与えないこと、符号化効率が良いことなどの理由による。楽器分類に従った音色指定は、同族の異なる楽器の選択が可能で、機能差のあるデコードを駆動することが可能である。SCPは制御パートの開始符号で、画と音との同期など制御のためのデータを配列するときに用いられる。

表7 音色パラメータの一部

大分類 P_1	中分類 P_2	小分類 $P_3 P_4$	備考
弦楽器 1	弓 奏 0	バイオリン (ビオラ)	0 0
		チェロ (ダブルベース)	0 2
		ストリングス	8 0
			合奏音色
	弾 奏 1	ギター	0 0
		エレキギター	0 1
		ハーブシコード	0 2
		ハーブ	0 3
		マンドリン (バンジョー) (バラライカ) (リュート) (ウクレレ)	0 4
		ピッチカート	7 0
			合奏音色
	打 弦 2	ピアノ	0 0
		電気ピアノ	0 1
			フェンダ ローズ
吹奏楽器 2	エア リード 0	フルート (リコーダ)	0 0
		口笛 (ジャズフルート)	0 2
			奏法による 音色
	シングル リード 1	クラリネット	0 0
		サキソフォン (リードオルガン)	0 1
		アコーディオン (ハーモニカ) (バンドネオン)	0 3
	ダブル リード 2	オーボエ (バスーン) (イングリッシュホルン)	0 0

(3) くり返し、ジャンプの符号

楽譜で使われるくり返しやジャンプ、リピートなどに対応する符号で、LBL, JMP, RPT, BRAがあり、それぞれ回数や行先を示すパラメータをともらう。この4種の組合せですべての場合に対応することが可能である。

(4) その他の制御符号

SLVは16種のパラメータによって音の強さを指定、LRTは音長を倍増する数値をパラメータで指定、CTMはSMPのパラメータと同じ音色のパラメータによって変更後の音色を指定する。その他、1つの付加音データの番号を指定するMNB、パラメータの終端を示すTEP、音高音長符号を結合するTIEなどが用意されている。

4.3 音高修飾符号集合

強弱符号で指定された音強をさらに細かく指定するもので、1音に対して1dB単位に音強を修飾することが可能である。64種の符号が定められている。

4.4 付加音符号の配列

文字や図形の場合は書式に従って順に符号化されるように、付加音符号の場合はパート順（優先度の高い順）に、かつパート内は時間順（音の流れの順）に符

号化を行う。また、付加音符号ではレベルA、Bの違いの他、すべての符号を復号できない機能のデコーダでも、それなりの再生が可能なよう、ダウンワードコンパチビリティを有するのも特徴である。

5. む す び

符号化伝送方式文字放送の符号化の検討に際しては、将来の拡張にも充分対応できるよう常に考慮してきた。文字放送が普及、発展し、本格的に符号の拡張を論ずる時がいつの日か来ることを望みたい。（荒木）

（昭和60年9月5日受付）

【参 考 文 献】

- 1) 小林幸雄：“キャプテンシステムとキャプテンプロトコル”，テレビ誌，**39**，5，pp. 400-407（May，1985）
- 2) 柳町昭夫ほか：“文字コード放送の方式検討”，テレビ学技報，IT 52，5，pp. 75-80（Feb.，1982）
- 3) 杉森，木俣：“図形の符号化”，放送技術，**38**，5，p. 77（May，1985）
- 4) 柳町：“データの伝送および符号の体系”，放送技術，**38**，5，p. 65（May，1985）
- 5) 杉森ほか：“文字放送における図形情報ならびに音楽情報の圧縮伝送技術”，テレビ誌，**36**，10（Oct.，1982）
- 6) “電波技術審議会答申第4編”，電波振興会（1985）