**PCT** WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

| | | |
|---|---|---|
| **(51) International Patent Classification** 6 : **G06K 19/06, B07C 3/18** | **A1** | **(11) International Publication Number:** **WO 96/13803**<br>**(43) International Publication Date:** 9 May 1996 (09.05.96) |

**(21) International Application Number:** PCT/CA95/00606

**(22) International Filing Date:** 27 October 1995 (27.10.95)

**(30) Priority Data:**
331,955 31 October 1994 (31.10.94) US

**(71) Applicant:** CANADA POST CORPORATION [CA/CA]; Suite N1110, 2701 Riverside Drive, Ottawa, Ontario K1A 0B1 (CA).

**(72) Inventors:** ULVR, Joseph; 106 Glenview Place, Carp, Ontario K0A 1L0 (CA). KHO, Adrian, Thong, Sun, Chai-Yu; 26 Westlock Way, Kanata, Ontario K2K 2K3 (CA).

**(74) Agents:** MCGRAW, James, M. et al.; Smart and Biggar, 900-55 Metcalfe Street, P.O. Box 2999, Station D, Ottawa, Ontario K1P 5Y6 (CA).

**(81) Designated States:** AU, CA, DE, GB, JP, SE, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

**Published**
*With international search report.*

**(54) Title:** BAR CODE FOR MAIL PROCESSING


**(57) Abstract**

A bar code for mail pieces uses bars each of which has four possible states. Two different bars indicate the start of the code and the same two bars in the same order indicate the end of the code. A data content identifier (DCI) follows the start bars and this indicates the structure and length of the following data field (K1A0B1) so that when the code is read it will be recognized and read properly. The use of the data content identifier (DCI) allows the code to be used for different customer and Post Office applied applications in which the code structure, length and content varies. The data field may contain a postal code with or without an address locator, a machine ID, customer information and service information. The code may include a country code field for mail pieces that are being mailed to a different country. The code may also include a field indicating whether the codeword is complete or whether it has to be concatenated with a preceding or subsequent codeword. Error protection in all cases is provided by a Reed-Solomon parity field (RS parity) following the data field. For customer applied codes this parity field may be made shorter than for Post Office applied codes because the potential for error in printing the code by the customer is less in view of the fact that he has more control over the paper quality, colour, extraneous markings, etc.

*FOR THE PURPOSES OF INFORMATION ONLY*

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| | | | | | |
|---|---|---|---|---|---|
| **AT** | Austria | **GB** | United Kingdom | **MR** | Mauritania |
| **AU** | Australia | **GE** | Georgia | **MW** | Malawi |
| **BB** | Barbados | **GN** | Guinea | **NE** | Niger |
| **BE** | Belgium | **GR** | Greece | **NL** | Netherlands |
| **BF** | Burkina Faso | **HU** | Hungary | **NO** | Norway |
| **BG** | Bulgaria | **IE** | Ireland | **NZ** | New Zealand |
| **BJ** | Benin | **IT** | Italy | **PL** | Poland |
| **BR** | Brazil | **JP** | Japan | **PT** | Portugal |
| **BY** | Belarus | **KE** | Kenya | **RO** | Romania |
| **CA** | Canada | **KG** | Kyrgystan | **RU** | Russian Federation |
| **CF** | Central African Republic | **KP** | Democratic People's Republic of Korea | **SD** | Sudan |
| **CG** | Congo | | | **SE** | Sweden |
| **CH** | Switzerland | **KR** | Republic of Korea | **SI** | Slovenia |
| **CI** | Côte d'Ivoire | **KZ** | Kazakhstan | **SK** | Slovakia |
| **CM** | Cameroon | **LI** | Liechtenstein | **SN** | Senegal |
| **CN** | China | **LK** | Sri Lanka | **TD** | Chad |
| **CS** | Czechoslovakia | **LU** | Luxembourg | **TG** | Togo |
| **CZ** | Czech Republic | **LV** | Latvia | **TJ** | Tajikistan |
| **DE** | Germany | **MC** | Monaco | **TT** | Trinidad and Tobago |
| **DK** | Denmark | **MD** | Republic of Moldova | **UA** | Ukraine |
| **ES** | Spain | **MG** | Madagascar | **US** | United States of America |
| **FI** | Finland | **ML** | Mali | **UZ** | Uzbekistan |
| **FR** | France | **MN** | Mongolia | **VN** | Viet Nam |
| **GA** | Gabon | | | | |

# BAR CODE FOR MAIL PROCESSING

## Background of the Invention

This invention relates to bar codes used in the processing of mail pieces.

In many countries a postal code is used to facilitate automation in sorting. In Canada, the postal code contains three alphabetic characters (letters) interleaved with three numeric characters (numbers) while in the U.S.A. the zip code consists of five numbers. If a customer has applied the postal code to an envelope this is converted by an optical character reader (O.C.R.) and computer in the Post Office to a bar code which is then printed on the envelope. If the customer has not applied the postal code, this will be generated in the Post Office and the bar code will be printed on the envelope as before.

It is also becoming more usual for the large corporate customer to apply the postal code in bar code format. A bar code is used because it is easier to read automatically than alphanumeric characters.

The British Post Office (BPO) has developed a 4 state bar code. The four possible states are one which comprises only a tracker element, one which comprises a tracker element and an ascender element, one which comprises a tracker element and a descender element and one which comprises a tracker element, an ascender element and a descender element. These elements will be described in detail hereinbelow.

The BPO code uses four bars to represent each alphanumeric character but for error protection each character must have two ascenders and two descenders which limits the number of possible combinations to 36. Error detection is dealt with by including a check sum. The BPO code is intended to be printed by customers (mailers) to encode sortation information.

The BPO code uses a single bar to indicate the start of the code and a different single bar to indicate

the end of the code. The start/stop bars in the BPO code can easily be confused and can result in decoding of an upside down bar code.

A paper entitled "A Damage Resistant Bar Code for the Royal Mail" by Blahut et al, 1992 discusses an improvement of the BPO bar code which essentially makes the BPO code more robust by adding Reed-Solomon error correction to handle missing bars (erasures) and bar print errors. The encoding is based on codewords of two bars each - a left and right codeword - and the interleaving of the codewords to form a bar code. Blahut has also identified some decoding logic to recover from missing or incorrect bars and overcome the weakness of the start/stop pattern with decoding logic. The Blahut Code is intended to be printed by the Post Office as an internal code and only postal sortation data is encoded.

It is an object of the present invention to improve on the BPO and Blahut codes by offering a robust code which is sufficiently flexible that it may be applied by the Post Office or the customer and which can be used not only for the postal code but for route sequencing, track and tracing, revenue accounting and other customer information as well as customer service information such as return mail management, automated data entry of customer information and the like.

## Summary of the Invention

According to the invention, vertical bars each of which has a state selected from a plurality of possible states, are used as the basis of the code.

In one embodiment of the invention, as in Blahut, bars which have four possible states are used but the invention is not limited to the use of a four state code. The codeword is defined by two start bars and two stop bars which are identical to the start bars. Sandwiched between these two bars are a data content identifier field and a data field. The data content identifier field in a

preferred embodiment is formed of three bars and indicates the structure and number of data elements following in the data field.

The data field may contain different types of information depending on the application. For example, it may contain a postal code in the ANANAN format where A is an alphabetical character and N is a numerical character. The A's are each represented by three bars and the N's are each represented by two bars.

The use of the start and stop bars provides an indication of correct orientation and direction of reading in a single efficient manner.

The use of the data content identifier permits a flexible format because the data content of any bar code can be identified by reading the first character. This first character, the DCI (Data Content Identifier) maps to a table which identifies the data content and therefore the application.

Not only postal (sortation and sequencing) data can be encoded but also customer data to support the creation of value added products and services selectable by the customer at the time of printing the mail pieces.

The new coding is space efficient because in a preferred embodiment it encodes characters in 3 bars with the numerics in the CPC postal code using only 2 bars each.

The code is highly damage resistant and provides varying levels of error correction appropriate to the application. For example, more error correction is provided in the internal applied codes than the customer applied codes.

The invention also enables in a specific embodiment the concatenation of multiple bar codes.

## Brief Description of the Drawings

Figure 1a illustrates the basic elements of the four state code used in a preferred embodiment of the invention;

Figure 1b illustrates the minimum and maximum dimensions of the bars used in the four state code;

Figures 2a to 2d illustrate how the code pattern or the individual bars within the code may be skewed;

Figure 3 illustrates a typical bar code employing the invention;

Figure 4a illustrates another example of a bar code employing the invention;

Figure 4b illustrates a bar code similar to that of Figure 4a in which the customer field has been broken down into sub-fields; and

Figures 5 through 11 illustrate further examples of bar codes employing the invention.

Description of the Preferred Embodiments

Referring to Figure 1, the basic elements of the printed four state code are four vertical bars which differ in length and/ or starting point with respect to the horizontal.

More particularly, there is a full height bar H which is the longest of the four bars and extends between lower and upper horizontal references $R_1$ and $R_2$, respectively. The bar immediately to the right of bar H is a bar D which is greater than half of the height of bar H and extends from above the mid-point of bar H down to the level of reference $R_1$. Immediately to the right of bar D is a bar A which has the same height as bar D but extends from below the mid-point of bar H up to the level of reference $R_2$. The final element is a bar T which is centred about the mid-point of bar H and which has a height represented by the overlap of bars A and D.

Another way of defining bars H, D, A and T is in terms of the three basic elements, Tracker, Ascender and Descender shown in Figure 1a. The Tracker element is a short element centred exactly between the lower and upper references $R_1$ and $R_2$. The Ascender and Descender elements are identical in length, the Ascender extending upwardly

from the upper limit of the Tracker to reference $R_2$ and the Descender extending downwardly from the lower limit of the Tracker to reference $R_1$.

The Tracker is present in all of the four bars. In the T bar the Tracker is the only element, the D bar consists of Tracker and Descender, the A bar consists of Tracker and Ascender and the H bar consists of the Tracker, Ascender and Descender. The four possible bars and their assigned numerical values can be summarized as follows:

| BAR | ELEMENTS | VALUE |
|---|---|---|
| T | Tracker | 3 |
| D | Tracker and Descender | 2 |
| A | Tracker and Ascender | 1 |
| H | Tracker, Ascender and Descender | 0 |

The maximum and minimum permissible dimensions of elements A, D, H, T and bar width X are indicated in Figure 1b where the maximum bar outlines are the shaded portions, and are summarized below:

| Element | Minimum | | Maximum | |
|---|---|---|---|---|
| | mm | in. | mm | in. |
| T | 1.0 | 0.04 | 1.6 | 0.06 |
| A | 2.6 | 0.10 | 3.7 | 0.145 |
| D | 2.6 | 0.10 | 3.7 | 0.145 |
| H | 4.2 | 0.165 | 5.8 | 0.23 |
| X (bar width) | 0.4 | 0.015 | 0.6 | 0.025 |
| bar gap | 0.4 | 0.015 | | |

Note: The minimum gap between bars takes priority over all other dimensions.

The bar density should be around 20-24 bars per 25.4 mm. Figure 2 illustrates two types of skew that can occur when printing bar codes. Figures 2a and 2b illustrate code skew $\alpha$ in which the entire code pattern is skewed with respect to the bottom edge of the mail piece and Figures 2c and 2d illustrate bar skew $\beta$ in which individual bars are skewed with respect to the centreline of the code pattern.

For code skew the acceptable limit is less than

+/- 5o from the horizontal and for bar skew the limit is less than +/- $5^{o}$ from the vertical. It is possible for both types of skew to occur on a single item and in that case the total skew $|\alpha|+|\beta|$ should be less than $5^{o}$.

Canada Post Corporation (CPC) proposes using the basic four state bar code in different applications. The term PostBar has been coined by CPC to refer to the basic four state bar code and the letters xyz appear after the designation PostBar where

"x" is "D" for domestic (Canada) applications
"G" for global (international) applications
"C" for CPC internal applications and
"S" for service applications.

"yz" specifies the number of characters in the bar code.

The PostBar applications are as follows:

| Domestic | Global | Service | Internal |
|---|---|---|---|
| PostBar.D07 | PostBar.G12 | PostBar.S06 | PostBar.C10 |
| PostBar.D12 | PostBar.G22 | PostBar.S11 | |
| PostBar.D22 | | PostBar.S21 | |

Let us consider in detail PostBar.C10 to illustrate how the four state code is applied. The format of this code is illustrated in Figure 3. This may be summarized as follows:

| DATA FIELD | BARS | DATA CHARACTERS |
|---|---|---|
| Start/synchronization | 2 | |
| Data Content Identifier (DCI) | 3 | Z |
| Postal code (FSA LDU) | 15 | ANANAN |
| RS Parity Check | 30 | |
| Machine ID | 4 | BBBB |
| Stop/synchronization | 2 | |
| Total | 56 | |

The data characters are denoted A, N, Z and B. A is an alphabetic character and is denoted by 3 bars, N is a numeric character and is denoted by 2 bars, Z is an alphanumeric (i.e., either alphabetic or numeric) character

and is denoted by 3 bars and B is 1 bar.

Table 1 shows the encoding for the "A" and "N" characters and Table 2 shows the encoding for the "Z" characters.

TABLE 1

'A' CHARACTERS

| Letter | Bars |
|---|---|
| A | HHD |
| B | HAH |
| C | HAA |
| D | HAD |
| E | HDH |
| F | HDA |
| G | HDD |
| H | HHA |
| I | AHA |
| J | AHD |
| K | AAH |
| L | AAA |
| M | HHH |
| N | ADH |
| O | ADA |
| P | ADD |
| Q | DHH |
| R | DHA |
| S | DHD |
| T | DAH |
| U | DAA |
| V | DAD |
| W | DDH |
| X | DDA |
| Y | DDD |
| Z | AHH |

'N' CHARACTERS

| Number | Bars |
|---|---|
| 0 | HH |
| 1 | HA |
| 2 | HD |
| 3 | AH |
| 4 | AA |
| 5 | AD |
| 6 | DH |
| 7 | DA |
| 8 | DD |
| 9 | TH |

TABLE 2

'Z' CHARACTERS

| Symbol | Bars |
|---|---|
| Space | HHT |
| A | HHH |
| B | HHA |
| C | HHD |
| D | HAH |
| E | HAA |
| F | HAD |
| G | HDH |
| H | HDA |
| I | HDD |
| J | AHH |
| K | AHA |
| L | AHD |
| M | AAH |
| N | AAA |
| O | AAD |
| P | ADH |
| Q | ADA |
| R | ADD |
| S | DHH |
| T | DHA |
| U | DHD |
| V | DAH |
| W | DAA |
| X | DAD |
| Y | DDH |
| Z | DDA |

| Symbol | Bars |
|---|---|
| 0 | DDD |
| 1 | THH |
| 2 | THA |
| 3 | THD |
| 4 | TAH |
| 5 | TAA |
| 6 | TAD |
| 7 | TDH |
| 8 | TDA |
| 9 | TDD |

Note that the encoding shown in Table 1 is used only for the postal code mapping. As seen in Table 1 only the 9 digit contains the tracking element T which because of its size is the most likely of the four elements to be obscured or missing. This minimization of the occurrence of the T bars provides extra security for the Postal Code.

With regard to the "B" bars, these are used only for the Machine ID and can be decoded using a quadral representation with 'n' as the number of bars in the data block and '$V_n$' as the value of each bar in the following equation:

$$B_nB_{n-1}\ldots B_1 = V_n x 4^{n-1} + V_{n-1} x 4^{n-2} + \ldots t\ V_1 x 4^0$$

The bar values ($V_n$) are assigned as follows: H=0; A=1; D=2; T=3

e.g. ADHA = $1 \times 4^3 + 2 \times 4^2 + 0 \times 4 + 1 \times 4^0$

= 64 + 32 + 0 + 1 = 097

For 4 bars, the maximum value is:

TTTT = $3 \times 64 + 3 \times 16 + 3 \times 4 + 3 \times 1$ = 225

Referring to the various data fields shown in Figure 3, the first field is START which comprises an A bar followed by a T bar. The last field is STOP which also comprises an A bar followed by a T bar. This sequence provides an orientation or direction of flow of the code so that an upside down label or letter inserted backwards can be identified immediately. The sequence also provides an additional marker for synchronization and a unique identifier so that the code can be recognized immediately.

The next field is the DCI (Data Content Identifier) which specifies the structure and the number of data elements within the bar code. When a bar code reader decodes a DCI it will know how to decode the remaining data elements. The DCI can be either an alphabetic or numeric character ("Z" character) encoded using three bars according to Table 2. Within CPC the DCI's are assigned in the following way:

1-9 Reserved for global (international) applications

A-L Reserved for domestic (Canada) applications

M-U Reserved for service applications

V-Z Reserved for internal applications.

The DCI illustrated in Figure 3 comprises two D bars followed by one A bar and from Table 2 this corresponds to the letter Z. When the DCI is a Z this specifies that there are 6 decodable characters in the form ANANAN for sortation and a binary machine ID in the form BBBB.

This does in fact correspond with Figure 3 where the next field is the postal code which in Canada is constructed as alternate alphabetic/numerical characters ANANAN with each letter being formed by 3 bars and each number formed of 2 bars as shown in Table 1. The postal code thus consists of 15 bars. By consulting Table 1, it can be seen the postal code in the example illustrated in Figure 3 is K1A 0B1.

The next field is the Reed-Solomon Parity Check consisting of 30 bars comprising 10 alphanumeric characters Z. The RS code chosen is a (16,6) Reed-Solomon code over GF(64) which can correct 5 symbol errors and up to 10 symbol erasures (30 bars). That is, more than half the bar code could be missing and the remaining bar code would still be successfully decoded. The error correcting capability of this RS code will be discussed in greater detail below.

The next field is the Machine ID field which identifies the particular machine which applied the bar code. The four bars shown in this example are:

$$\text{ADHA} = 1 \times 4^3 + 2 \times 4^2 + 0 \times 4^1 + 1 \times 4^0$$

$$= 64 + 32 + 0 + 1 = 097$$

Turning now to Figure 4a, this illustrates the format of PostBar.D22 which uses a (25,21) Reed-Solomon code over GF(64). PostBar.D22 is a customer applied bar code for domestic Canadian applications. The data structure may be summarized as follows:

| DATA FIELD | BARS | DATA CHARACTERS |
|---|---|---|
| Start/synchronization | 2 | |
| Data Content Identifier (DCI) | 3 | Z |
| Postal code (FSA LDU) | 15 | ANANAN |
| Address Locator (AL) | 12 | ZZZZ |
| Customer Information | 33 | ZZZZZZZZZZZ |
| RS Parity Check | 12 | |
| Stop/synchronization | 2 | |
| Total | 79 | |

As for the PostBar.C10 code discussed above, data character A is an alphabetic character denoted by 3 bars, N is a numeric character denoted by 2 bars and Z is an alphanumeric character denoted by 3 bars. Table 1 shows the encoding for the "A" and "N" characters and Table 2 shows the encoding for the "Z" characters.

Referring to the various data fields shown in Figure 4a, the start and stop fields, the DCI field and the Postal code fields are identical to the corresponding fields in the PostBar.C10 code. In the particular example shown, by consulting Table 2 it will be seen the DCI corresponds to the letter C and by consulting Table 1 it will be seen the Postal code corresponds to L3B 4T9.

When the DCI is a C this specifies that there are 21 decodable characters which follow in the form of the postal code (ANANAN), address locator (ZZZZ) and 11 customer data characters (ZZZZZZZZZZZ).

Unlike PostBar.C10, PostBar.D22 does not have a Machine ID field as the printer applying the code is not a CPC (Canada Post Corporation) machine and is, therefore, of no real interest.

PostBar.D22 has two fields, namely AL and Customer Information, not present in PostBar.C10. The field AL is an address locator field which consists of 12 bars and appears immediately to the right of the Postal Code. The Customer Information field follows and this has 33 bars. These fields are encoded according to alphanumeric Table 2 and so there are 4 characters for the AL field and 11 for the Customer Information field.

Reference should be made to U.S. patent application Serial No. 888,905 filed on May 26, 1992 and assigned to Canada Post Corporation, which application is incorporated hereby by reference, for a further explanation of the AL and Customer Information fields. More particularly and in brief, the AL field, referred to in the earlier application as PODI (Point of Delivery Indicator) is a suffix to the postal code which is determined from the address on the mail piece as well as the Postal code. The postal code together with the AL allows a mail piece to be sorted for delivery to the specific address. The term POCI (Point of Call Identifier) has been coined for the combination (Postal Code + the Address Locator).

It is noted that the RS Parity field in PostBar.D22 consists of only 12 bars in contrast to the 30 bars of PostBar.C10. This is because more protection is needed for Post Bar.C10 than for PostBar.D22. This results from the fact that the PostBar.C10 is a code printed by CPC on mail pieces which have a great variety of surfaces and background and so there is a likelihood of background noise from extraneous printing or marking. On the other hand PostBar.D22 is applied by the customer who has greater control over the printing surface and so there is less potential for background noise.

From a consideration of Table 2 it can be seen that, for the specific example shown for PostBar.D22, the AL is 1420, and the Customer Information is CFFMIPLXF6V. From a consideration of Table 1 the postal code converts to L3B 4T9.

Figure 4b shows how the Customer Information field of Figure 4a may be broken down into sub-fields. In Figure 4b the bars are not shown. The customer data can be broken into sub-fields such as those shown. The Product Code identifies the specific customer, the Sequence Number identifies the batch mailed on a particular day and the Month and Day of Month are self-explanatory. This

information uniquely identifies a mail piece and allows track and trace of the mail piece as well as revenue accounting for example.

The numbers in the brackets represent the number of combinations possible for each sub-field.

Figure 5 illustrates an example of an International or Global code, PostBar.G12 which would be used when the mail piece is addressed to another country.. This is a (15,11) Reed-Solomon code over GF(64). The format may be summarized as follows:

| DATA FIELD | BARS | DATA CHARACTERS |
|---|---|---|
| Start/synchronization | 2 | |
| Data Content Identifier (DCI) | 3 | Z |
| Country Code (CC) | 6 | NNN |
| Postal code* | 24 | ZZZZZZZZ |
| RS Parity Check | 12 | |
| Stop/synchronization | 2 | |
| Total | 49 | |

* Unused characters in the postal code field will be filled with space characters.

The DCI is determined from Table 2 to be 1 and this specifies that there are 11 decodable characters that follow in the form of a three numeric character country code (NNN) and an 8 character postal code (ZZZZZZZZ).

It is noted that in this code no A or N characters of Table 1 are used for the Postal Code. Also, a new field, namely Country Code, is present and as can be seen by consulting Table 1, for the specific example shown this translates to 180 which may, for example, identify the USA. The postal code or zip code is determined from Table 2 to be 91266 followed by three spaces. For the U.S.A. only 15 bars are needed for the ZIP code but the field is provided with 24 bars because other countries require more than 5 characters for the postal code.

Figure 6 illustrates an example of a customer applied service code, PostBar.S21 which is a (25, 21) Reed-Solomon code over GF(64) having a data structure summarized as follows:

| DATA FIELD | BARS | DATA CHARACTERS |
|---|---|---|
| Start/synchronization | 2 | |
| Data Content Identifier (DCI) | 3 | Z |
| Bar Code Sequencer (BCS) | 3 | Z |
| Service Information (SI) | 57 | ZZZZZZZZZZZZZZZZZZZ |
| RS Parity Check | 12 | |
| Stop/synchronization | 2 | |
| Total | 79 | |

As for PostBar.G12 discussed above all the data characters are Z characters obtained from Table 2. The DCI is determined from Table 2 to be S and this specifies that there follows as data a 19 character (57 bars) Service Information field which in this case translates to ABCDEFGHIJ123456789 from Table 2. There is no routing data field because this code is used for special services required by a customer. For example the Service Information field could be used for return mail management or to provide other information useful to the customer.

Between the DCI field and the Service Information field is a Bar Code Sequencer (BCS) field which is a single character consisting of 3 bars that allows the concatenation of two bar codes to encode data longer than 19 characters. When a single 19 character code is used the BCS is DDD which from Table 2 is 0. To indicate the first of two concatenated bar codes the BCS would be chosen to be 1 and to indicate the second of two concatenated bar codes the BCS would be chosen to equal 2.

The operation of the error correcting code (ECC) for PostBar.C10 (Figure 3) and for PostBar.D22 (Figure 4a) will now be discussed in greater detail.

The error correcting code (ECC) for PostBar.C10 protects the postal code and DCI but not the machine ID. The ECC is a (16, 6) Reed-Solomon code defined over the field GF(64) with 64 elements. Precisely, the defining roots of the code are $\alpha^i$ for $i = 1, 2, \ldots, 10$ where $\alpha$ is a root of $X^6+X+1$. Each codeword consists of 6 message (or information) symbols and 10 check symbols. Each symbol is an element of GF(64) and is represented by 3 bars. The ECC

can correct up to t symbols with errors and e erased symbols as long as $2t+e \leq 10$. For example, since each symbol corresponds to 3 bars, an erasure of 30 consecutive bars is correctable. One should be aware that if there are e=10 erased symbols then there is no way for the code to detect any additional errors. Any other error in addition to these 10 will force a decoding error. (The constant erase max in the software can be set less than 10 to stay away from this possibility.)

The error correcting code for PostBar.D22 covers all of the fields except the start and stop bars. The code is a (25, 21) Reed-Solomon code defined over GF(64). It can correct t errors and e erasures in the symbols of the code as long as $2t+e \leq 4$. As before, each code symbol corresponds to 3 bars. The code uses 4 check symbols which contribute 12 bars to the code.

The error correcting code specified for each of the bar codes is a Reed-Solomon code over the field GF(64) with 64 elements. We use a primitive element $\alpha$ of the field GF(64) which is a root of $X^6+X+1$ (over GF(2)). This means that the 63 powers $\alpha^i$, for i=0, 1, ..., 62 are the 63 distinct non-zero elements of the field. Also the 6 elements $\alpha^0$, $\alpha^1$, $\alpha^2$, $\alpha^3$, $\alpha^4$, $\alpha^5$ form a linear basis for the field GF(64). Thus each element w of GF(64) has a unique expression as a sum,

(*) $$w = w_5\alpha^5 + w_4\alpha^4 + w_3\alpha^3 + w_2\alpha^2 + w_1\alpha^1 + w_0\alpha^0$$

where each $w_i = 0$ or 1. By using the identity $\alpha^6 = \alpha + 1$ any power of $\alpha$ can be expressed in the form (*).

The translation from bar to field elements is accomplished by grouping the bars in sets of three. Each bar corresponds to a pattern of 2 bits:

H = 00, A = 01, D = 10, T = 11

A set of three bars corresponds to 6 bits which we take as the values of $w_i$ in the expression (*). So for example,

TDT $\longrightarrow$ 11 10 11 $\longrightarrow$ $1\alpha^5 + 1\alpha^4 + 1\alpha^3 + 0\alpha^2 + 1\alpha^1 + 1\alpha^0 = \alpha^{21}$

Table 3 displays the correspondence between bar patterns and field elements. In the encoding and decoding software the field elements are represented as integers in the range 0 to 63, (thus as 6 bits). For example, the element $\alpha^{21}$, above, which has bit pattern 111011 corresponds to the integer 59 which is 111011 in binary (59=32+16+8+2+1). In Table 3, the column int gives the integer corresponding to each field element in this way.

TABLE 3

| bars | i | $\alpha^i$ | int | bars | i | $\alpha^i$ | int |
|---|---|---|---|---|---|---|---|
| HHH | ** | 000000 | 0 | DAA | 31 | 100101 | 37 |
| HHA | 0 | 000001 | 1 | HDA | 32 | 001001 | 9 |
| HHD | 1 | 000010 | 2 | AHD | 33 | 010010 | 18 |
| HAH | 2 | 000100 | 4 | DAH | 34 | 100100 | 36 |
| HDH | 3 | 001000 | 8 | HDT | 35 | 001011 | 11 |
| AHH | 4 | 010000 | 16 | AAD | 36 | 010110 | 22 |
| DHH | 5 | 100000 | 32 | DTH | 37 | 101100 | 44 |
| HHT | 6 | 000011 | 3 | ADT | 38 | 011011 | 27 |
| HAD | 7 | 000110 | 6 | TAD | 39 | 110110 | 54 |
| HTH | 8 | 001100 | 12 | DTT | 40 | 101111 | 47 |
| ADH | 9 | 011000 | 24 | ATA | 41 | 011101 | 29 |
| THH | 10 | 110000 | 48 | TDD | 42 | 111010 | 58 |
| DHT | 11 | 100011 | 35 | TAT | 43 | 110111 | 55 |
| HAA | 12 | 000101 | 5 | DTA | 44 | 101101 | 45 |
| HDD | 13 | 001010 | 10 | ADA | 45 | 011001 | 25 |
| AAH | 14 | 010100 | 20 | THD | 46 | 110010 | 50 |
| DDH | 15 | 101000 | 40 | DAT | 47 | 100111 | 39 |
| AHT | 16 | 010011 | 19 | HTA | 48 | 001101 | 13 |
| DAD | 17 | 100110 | 38 | ADD | 49 | 011010 | 26 |
| HTT | 18 | 001111 | 15 | TAH | 50 | 110100 | 52 |
| ATD | 19 | 011110 | 30 | DDT | 51 | 101011 | 43 |
| TTH | 20 | 111100 | 60 | AAA | 52 | 010101 | 21 |
| TDT | 21 | 111011 | 59 | DDD | 53 | 101010 | 42 |
| TAA | 22 | 110101 | 53 | AAT | 54 | 010111 | 23 |
| DDA | 23 | 101001 | 41 | DTD | 55 | 101110 | 46 |
| AHA | 24 | 010001 | 17 | ATT | 56 | 011111 | 31 |
| DHD | 25 | 100010 | 34 | TTD | 57 | 111110 | 62 |
| HAT | 26 | 000111 | 7 | TTT | 58 | 111111 | 63 |
| HTD | 27 | 001110 | 14 | TTA | 59 | 111101 | 61 |
| ATH | 28 | 011100 | 28 | TDA | 60 | 111001 | 57 |
| TDH | 29 | 111000 | 56 | THA | 61 | 110001 | 49 |
| THT | 30 | 110011 | 51 | DHA | 62 | 100001 | 33 |

For the CPC internal bar code, PostBar.C10, the 56 bars are arranged as follows:

$\underbrace{b_0b_1}_{\text{start}}$ $\underbrace{b_2b_3\ldots b_{49}}_{\text{coded portion}}$ $\underbrace{b_{50}b_{51}b_{52}b_{53}}_{\text{machine ID}}$ $\underbrace{b_{54}b_{55}}_{\text{stop}}$

The coded portion is divided into blocks of three bars. Each block represents one element of the field. The ECC runs from right to left across the bars. We set $c_1=b_{47-31}b_{48-31}b_{49-31}$.

Thus we have:

$\underbrace{b_0b_1}_{\text{start}}$ $\underbrace{b_2b_3b_4}_{c_{15}}$ $\underbrace{b_5b_6b_7}_{c_{14}}\ldots\ldots\underbrace{b_{44}b_{45}b_{46}}_{c_1}$ $\underbrace{b_{47}b_{48}b_{49}}_{c_0}$ $\underbrace{b_{50}b_{51}b_{52}b_{53}}_{\text{machine ID}}$ $\underbrace{b_{54}b_{55}}_{\text{stop}}$

Using these 16 elements $c_0, \ldots, c_{15}$ of GF(64) the ECC is defined by specifying that a codeword must satisfy:

$$(**) \qquad \sum_{i=0}^{15} c_i\alpha^{ij} = 0 \text{ for } j = 1, 2, \ldots, 10$$

The equation represents 10 equations with 10 unknown coefficients which is solved by the computer bar code generating software.

Using the symbology defined above, we record the postal code (PC) in bars $b_5 \ldots b_{19}$ hence in elements $c_{14}, \ldots, c_{10}$. The DCI goes in bars $b_2b_3b_4$ hence in $c_{15}$. The check symbols $c_0, \ldots, c_9$ are now uniquely determined by the parity checks (**).

For example, DCI=Z, postal code K1S 5B6, mach ID - DHAH encodes to the following codeword:

18 29 12 5 14 27 52 54 4 6 8 33 9 6 20 41

$c_0$ ... ... $c_{15}$

Table 2 then produces the bar code

AT DDA AAH HAD HDA DHA HDH HAD HAH TAD
TAH ADT HTD HAA HTH ATA AHD DHAH AT

Reorganized into A- and N- fields this is

AT DDA AAH HA DHD AD HAH DH

Start DCI Postal Code -ANANAN

HAD HAH TAD TAH ADT HTD HAA HTH ATA AHD DHAH AT

checks mach ID stop

The Domestic Bar Code, PostBar.D22 is similar in definition to the CPC Internal code. All of the 79 bars of the Customer Code, except for the 4 start/stop bars are covered by the ECC. Again the bars correspond to field elements in sets of 3. Thus field element $c_i = b_{74-3i} b_{75-3i} b_{76-3i}$. We have:

$b_0b_1$ $b_2b_3b_4$ $b_5b_6b_7$ ..... $b_{71}b_{72}b_{73}$ $b_{74}b_{75}b_{76}$ $b_{77}b_{78}$

start $c_{24}$ $c_{23}$ ..... $c_1$ $c_0$ stop

Using these 25 elements $c_0, \ldots, c_{24}$ of GF(64) the ECC for the Customer Code is defined by specifying that a codeword must satisfy:

$$(***) \qquad \sum_{i=0}^{24} c_i \alpha^{ij} = 0 \text{ for } j = 1,2,3,4$$

This equation represents 4 equations with 4 unknown coefficients which is solved by the computer bar code generating software.

Using the symbology defined above, we record the DCI in bars $b_2b_3b_4$ hence in the element $c_{24}$. The postal code is placed in bars $b_5$ to $b_{19}$ hence in the elements $c_{23}, \ldots c_{19}$. The Address Locator (AL) goes in bars $b_{20}, \ldots b_{31}$ hence in $c_{18}, \ldots, c_{15}$. The customer field goes in bars $b_{32}, \ldots, b_{64}$ hence in $c_{14}, \ldots, c_4$. The check symbols $c_0, \ldots c_3$ are now uniquely determined by the parity checks (***),

For example, DCI=C, Postal Code=M4J 3W8, AL-1420, Cust field=ABCDEFGHIJK encodes the following codeword:

40 30 23 4 17 16 10 9 8 6 5 4 2 1 0 42 49 52 48 10 10 9 21 02

$c_0 \ldots$ $\ldots c_{24}$

Using Table 3 to change from integer representation to bars

produces the bar code:

AT HHD HHH AAA HDA HDD HDD THH TAH THA DDD HHA HHH
HHD HAH HAA HAD HDH HDA HDD AHH AHA HAH AAT ATD DDH AT

This breaks down into A-, N-, and Z-fields as:

AT HHD HHH AA AHD AH DDH DD THH TAH THA DDD

Start DCI Postal Code-ANANAN Address Locator

HHA HHH HHD HAH HAA HAD HDH HDA HDD AHH AHA HAH AAT ATD DDH AT

Customer Field checks stop

The checks do not follow any of the Table 1 or 2 symbologies but are simply field elements coded as in Table 3.

It is noted that in a (16,6) Reed-Solomon code the probability of a random pattern being a valid codeword is $p=3.78\times10^{-6}$. Moreover, for PostBar.C10 the letters D,F,I,O,Q, or U currently do not appear in the bar codes and W or Z do not appear as the first letters of the code. There are 64 possible 3 bar system patterns, and only 20 are used for postal code letters. The digits for the postal code are encoded as 2 bar patterns, and only 10 of 16 are used. This gives a probability of:

$$\frac{18}{64} \times \frac{10}{16} \times \frac{20}{64} \times \frac{10}{16} \times \frac{20}{64} \times \frac{10}{16} = .0067$$

or about ½ of 1% that a random sequence of bars is a valid postal code. Combining this with $p$ above gives a probability of $2.5\times10^{-8}$ or 1 chance in 40,000,000 that a random string of bars will be interpreted as a valid postal code. Other internal checks, such as the use of the code type, will reduce this probability even further. This information should be incorporated into the OCR software.

Figures 7 through 11 respectively illustrate the format of PostBar.D07, D12, S06, S11 and G22. These structures will not be described herein as they will be readily understood from the foregoing description.

It should be understood that the inherent flexibility of the new code permits of many more

applications than described and illustrated herein and the particular applications described are to be considered representative only.

Furthermore, it is envisaged that data compression techniques may be used to accommodate longer messages than currently provided for with PostBar.D07, D12 and D22. Data compression could also be used with the G and S codes.

Where data compression is used, it is likely the DCI will not be compressed so as to allow speedy determination of the make-up of the bar code. An example is the case of a service bar code present on a mail item that is being processed for sortation. As there is no sortation data in S codes, once the DCI is derived, the machine logic knows that no further decoding is required to process and decompress the data.

Suitable software listing for encoding of PostBar.C10 is provided on Appendix A attached hereto and the software listing for encoding of PostBar.D07, D12 and D22 is provided on Appendix B attached hereto. Appendix C attached hereto is suitable software listing for decoding PostBar.C10 and Appendix D attached hereto is the software listing for decoding PostBar.D07, D12 and D22.

CLAIMS:

1. A mail piece bearing a bar codeword containing information for the processing of the mail piece, the bar codeword having a plurality of parallel bars each of which has a state selected from a plurality of possible states, the bar codeword comprising a start field followed by a data content identifier (DCI) field specifying the structure and number of data elements within the codeword followed by a data field, followed by a Reed-Solomon parity field, followed by a stop field.

2. A mail piece according to claim 1 in which the start field and the stop field are identical, each consisting of two different state bars.

3. A mail piece according to claim 1 in which the Reed-Solomon field contains a plurality of Reed-Solomon characters each encoded in 3 bars.

4. A mail piece according to claim 1 in which the DCI is encoded in three bars.

5. A mail piece according to claim 1 in which the data field includes at least one of a postal code, with or without address locator, customer information and service information.

6. A mail piece according to claim 5 in which the data field contains characters each of which is encoded in three bars.

7. A mail piece according to claim 5 in which the postal code contains alphabetic characters and numeric characters, the alphabetic characters being encoded in three bars and the numeric characters being encoded in two bars, any remaining data being encoded in three bar

characters.

8. A mail piece according to claim 1 in which the data field includes a country code followed by at least one of a postal code, with or without an address locator, customer information and service information.

9. A mail piece according to claim 8 in which the data field contains characters each of which is encoded in three bars.

10. A mail piece according to claim 1 in which the data field includes a postal code and machine ID.

11. A mail piece according to claim 10 in which the machine ID is represented by four bars and each bar state has a unique numerical value.

12. A mail piece according to claim 1 in which a bar code sequencer field is located between the DCI field and the data field, the bar code sequencer field containing a code indicating one of the bar code is a single code, the bar code is the first of two concatenated bar codes and the bar code is the second of two concatenated bar codes.

13. A mail piece according to claim 1 in which the code may be applied by one of the Post Office and a customer, the Reed-Solomon parity field being longer for a Post Office applied code than for a customer applied code.

14. A mail piece bearing a bar codeword containing information for the processing of the mail piece, the bar codeword having a plurality of parallel bars each of which has one of four possible states selected from a full height bar (H), extending between an upper and lower level, a partial height bar (D) descending to the lower level, a

partial height bar (A) ascending to the upper level, and a bar (T) the height and position of which is determined by overlap of the descending and ascending bars, the bar codeword comprising a start field followed by a data content identifier (DCI) field specifying the structure and number of data elements within the codeword followed by a data field followed by a Reed-Solomon parity field followed by a stop field.

15. A mail piece according to claim 14 in which the start field and the stop field are identical, each consisting of two different state bars.

16. A mail piece according to claim 15 in which the start and stop bars are an A bar followed by a T bar.

17. A mail piece according to claim 14 in which the DCI is encoded in three bars.

18. A mail piece according to claim 14 in which the data field includes at least one of a postal code, with or without address locator, customer information and service information.

19. A mail piece according to claim 15 in which the data field contains characters each of which is encoded in three bars.

20. A mail piece according to claim 18 in which the postal code contains alphabetic characters and numeric characters, the alphabetic characters being encoded in three bars and the numeric characters being encoded in two bars, any remaining data being encoded in three bar characters.

21. A mail piece according to claim 14 in which the

data field includes a country code followed by at least one of a postal code, with or without an address locator, customer information and service information.

22. A mail piece according to claim 21 in which the data field contains characters each of which is encoded in three bars.

23. A mail piece according to claim 14 in which the data field includes a postal code and machine ID.

24. A mail piece according to claim 23 in which the machine ID is represented by four bars and each bar state has a unique numerical value.

25. A mail piece according to claim 14 in which a bar code sequencer field is located between the DCI field and the data field, the bar code sequencer field containing a code indicating one of the bar code is a single code, the bar code is the first of two concatenated bar codes and the bar code is the second of two concatenated bar codes.

26. A mail piece according to claim 14 in which the code may be applied by one of the Post Office and a customer, the Reed-Solomon parity field being longer for a Post Office applied code than for a customer applied code.

27. A mail piece according to claim 14 in which the Reed-Solomon field contain a plurality of Reed-Solomon characters each encoded in 3 bars.

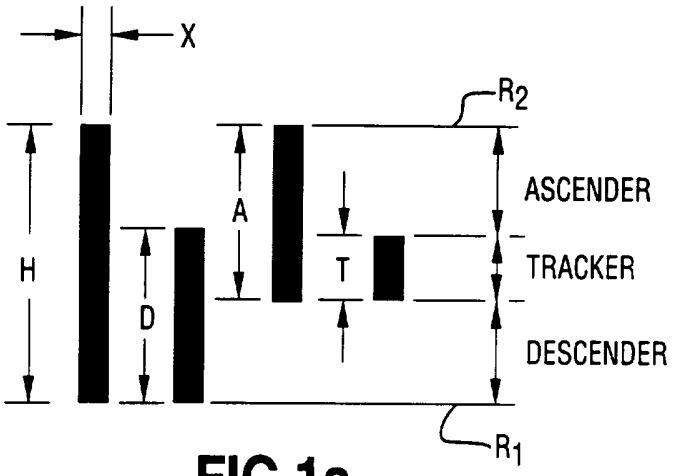
X
R2
ASCENDER
A
H
T
TRACKER
D
DESCENDER
R1

FIG.1a

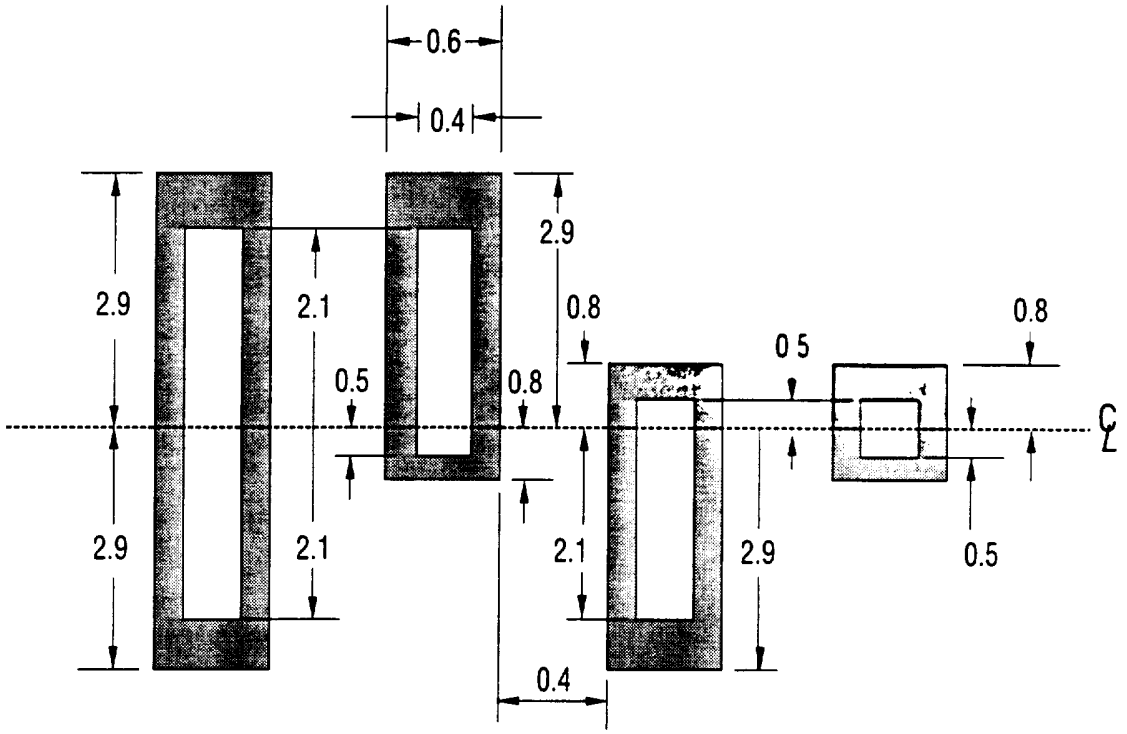
0.6
0.4
2.9
2.9
2.1
0.8
0.8
0.5
0.5
0.8
0.5
2.9
2.1
2.1
2.9
0.5
0.4

FIG.1b

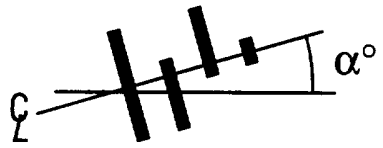


FIG. 2a

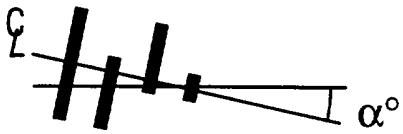


FIG. 2b

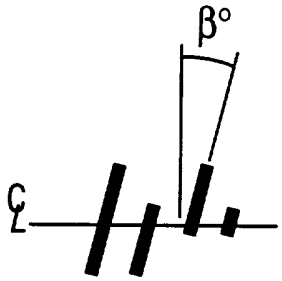


FIG. 2c

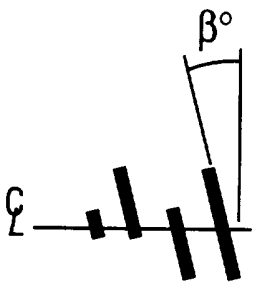


FIG. 2d

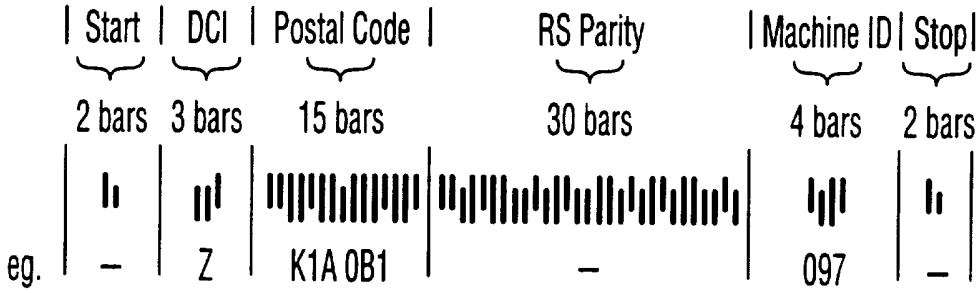


FIG. 3

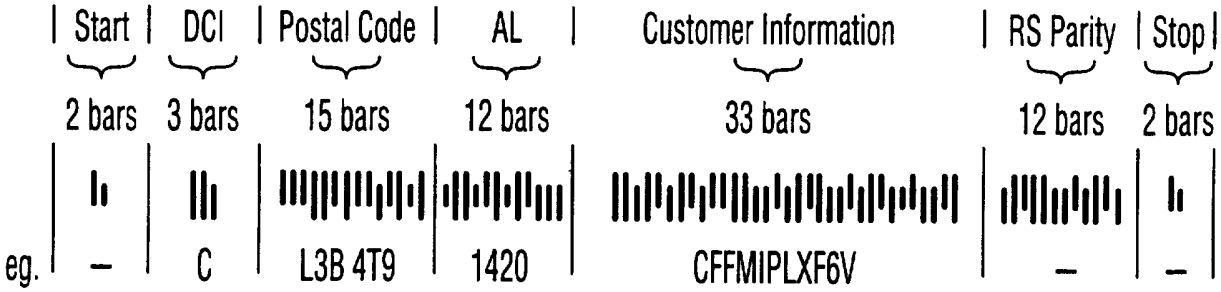


FIG. 4a

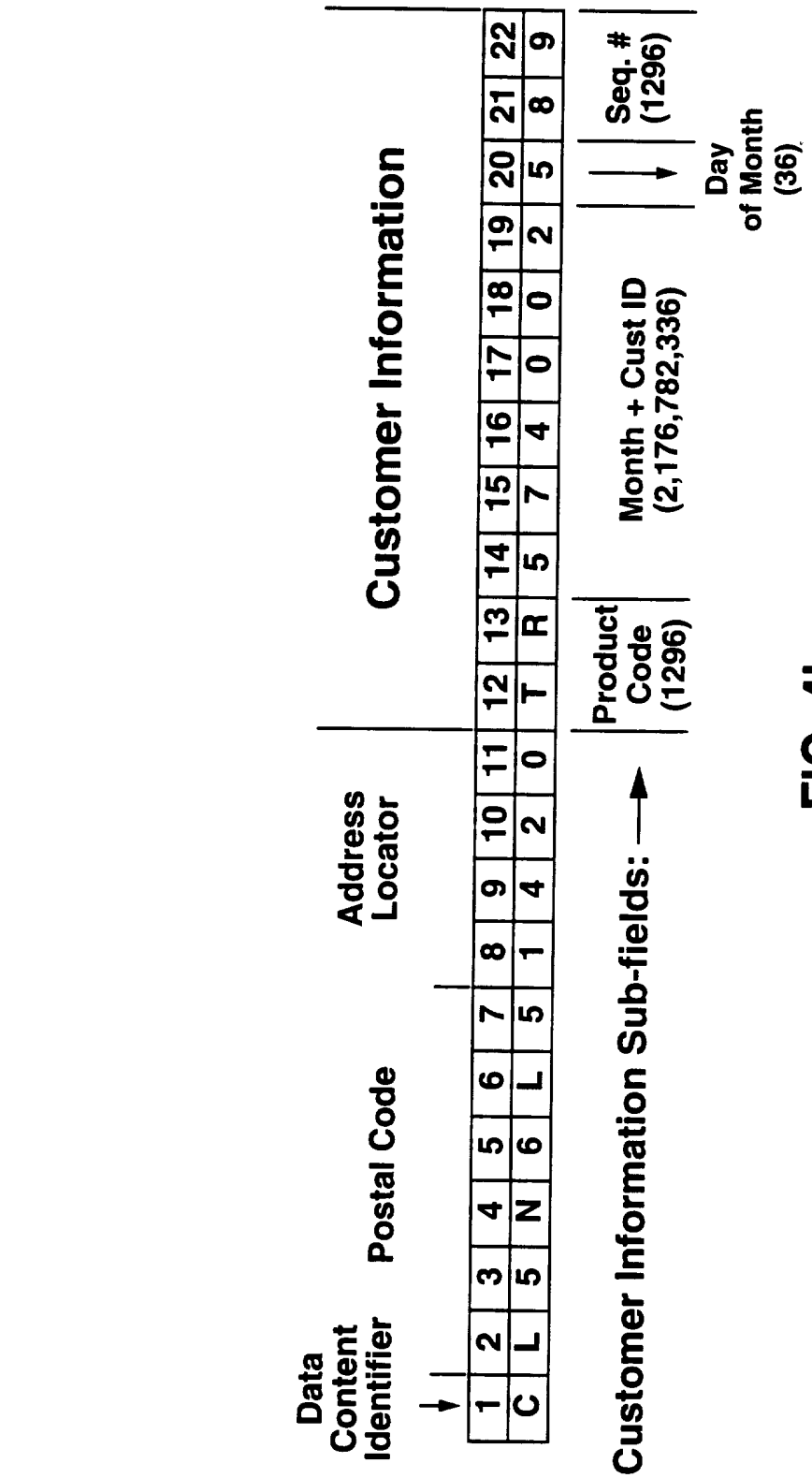
Data Content Identifier
Postal Code
Address Locator
Customer Information
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22
C L 5 N 6 L 5 1 4 2 0 T R 5 7 4 0 0 2 5 8 9
Customer Information Sub-fields:
Product Code (1296)
Month + Cust ID (2,176,782,336)
Day of Month (36)
Seq. # (1296)

FIG. 4b

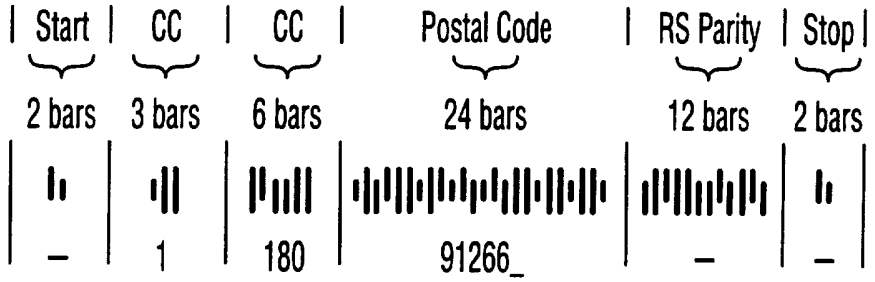


FIG. 5

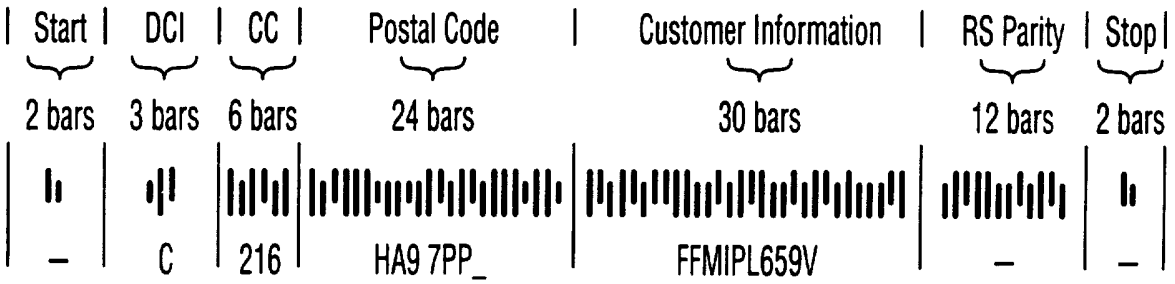


FIG. 11

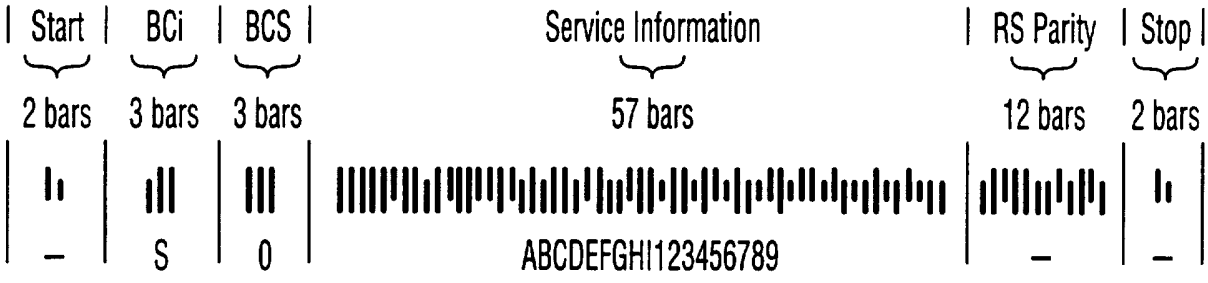


FIG. 6


FIG. 7

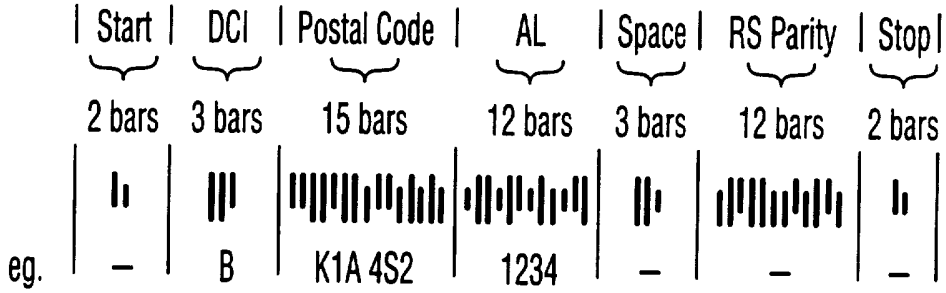


FIG. 8


FIG. 9

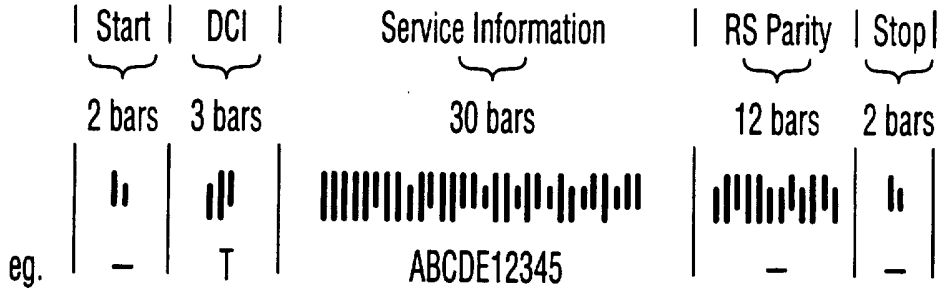


FIG. 10

# INTERNATIONAL SEARCH REPORT

Inte. nal Application No
PCT/CA 95/00606

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 G06K19/06 B07C3/18

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 G06K B07C

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category * | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| A | US,A,5 288 976 (CITRON HOWARD ET AL) 22 February 1994<br>see column 3, line 24 - line 26<br>see column 6, line 29 - line 38<br>see column 7, line 54 - line 66<br>see figures 3,4 | 1,14 |
| A | US,A,5 298 731 (ETT ALLEN H) 29 March 1994<br>see column 1, line 20 - line 26<br>see column 7, line 45 - line 55<br>see figure 10E | 1,14 |
| A | US,A,5 153 929 (ITAGAKI HIROSHI) 6 October 1992<br>see column 2, line 10 - line 22<br>see column 25, line 23 - line 26 | 1,14 |

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

1

Date of the actual completion of the international search
18 January 1996

Date of mailing of the international search report
30 JANUARY 1996

Name and mailing address of the ISA
European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax (+31-70) 340-3016

Authorized officer
Goossens, A

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

| Category * | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| A | GB,A,2 140 948 (GCA CORP) 5 December 1984<br>see page 3, line 39 - line 58<br>see figure 5<br>--- | 1,14 |
| A | US,A,5 324 927 (WILLIAMS ROBERT L) 28 June 1994<br>see column 1, line 6 - line 12<br>see column 2, line 58 - column 4, line 8<br>see figure 2<br>----- | 1,14 |

1

# INTERNATIONAL SEARCH REPORT

...information on patent family members

Ir... ...nal Application No
PCT/CA 95/00606

| Patent document cited in search report | Publication date | Patent family member(s) | | Publication date |
|---|---|---|---|---|
| US-A-5288976 | 22-02-94 | NONE | | |
| US-A-5298731 | 29-03-94 | NONE | | |
| US-A-5153929 | 06-10-92 | JP-A- | 3285459 | 16-12-91 |
| GB-A-2140948 | 05-12-84 | US-A- | 4567361 | 28-01-86 |
| | | DE-A- | 3418753 | 29-11-84 |
| | | FR-A- | 2546645 | 30-11-84 |
| | | JP-A- | 60041173 | 04-03-85 |
| US-A-5324927 | 28-06-94 | AU-B- | 5992294 | 15-08-94 |
| | | CA-A- | 2152082 | 21-07-94 |
| | | EP-A- | 0679111 | 02-11-95 |
| | | WO-A- | 9415725 | 21-07-94 |