BARCODES

The essential feature of a barcode is that it may be read even though the speed of the scanning device may vary wildly. This is achieved by comparing the width of successive bars and spaces with the preceding ones since the speed cannot vary much between adjacent elements. Detecting transitions from black to white and white to black is rather more difficult than it may appear to the eye which sees the whole of a bar.

Several ways have been devised to interpret the sequence of bars and spaces as characters. These are known as barcode symbologies and may be classified in two ways, two width or multi width and intelligent or non intelligent.

A wide to narrow bar ratio of 3:1 is relatively easy to determine but if there are four different widths they have to be printed accurately and measured accurately.

Intelligent symbologies may be directly decoded into characters like signals over a telephone line whereas non intelligent ones are just patterns which have to be looked up in a table. At first glance it might appear that non intelligent codes might be more efficient since any pattern could be printed. However this is not the case since to know you have a character to look up you have to have a fixed number of bars. This further means that there has to be a fixed number of wide elements otherwise the width of the characters would vary.

In devising a system or deciding which one to use it is necessary to consider:

How many characters is it able to represent?
How much space does it take up?
How reliable is it?
How easy is it to print?

Apart from Plessey and Telepen, which are British, all the symbologies devised are of American origin with no other country having come up with a viable system.

The only codes worth considering are Telepen, Code39, EAN/UPC, Code128 and Interleaved 2 of 5 (I2/5). Plessey and Codabar still exist in some legacy systems but are effectively obsolete.

Telepen, although devised early in 1972 without any knowledge of other codes, is the only symbology to be able to directly represent the full ASCII character set of 128 characters. For numeric only applications two digits, as well as 0X to 9X may be packed into one character enabling 12 digits per inch to be printed normally, although up to 20 is feasible. There is a start code, a stop code and a check character. Telepen is very reliable with a virtually zero chance of misreading. It is also the most tolerant with early labels being successfully printed on poor quality dot matrix printers. Most readers and printers handle it but it is not always referred to in American documentation.

Code39 was devised a few years later motivated by the lack of an alphanumeric symbology of American origin. This is more or less a non intelligent version of Telepen providing a range of 43 characters instead of 128 and consequently there is no double density numeric version. With the optional check character it is adequately reliable and modern printers are able to produce it without any problem.

The rather lax specification of the wide to narrow ratio, for historic reasons, has compromised the decode algorithm since the reader has to take account of this possible variation.

To provide a wider range of characters Code128 was devised although subsets are needed to obtain the full 128 characters. This is a multi width symbology which requires high quality printing. It has a much better packing density than Code39 but at the expense of print tolerance and ease of reading.

Due to the inefficiency of Code39 for numeric data an early American symbology I2/5 is still used in some applications. This is limited to only the ten digits in pairs and is very unreliable since there are no proper end codes and there is no check digit. Readers are normally configured to check the length of the symbol read.

The retail symbology, EAN, is a derivative of the American UPC. This multi width symbology is not used in any application outside retail and until the advent of high quality printers had to be printed from photo masters. Although security is not very good misreads will usually be rejected since only a small range of products will be accepted.

Many symbologies specify various ways in which data characters, or pairs of characters, may be interpreted. These just provide optional ways of pre processing the data and are nothing to do with the structure of the barcode itself.

More recently, so called, two dimensional barcodes have appeared which enable large amounts of data to be printed in machine readable form. With this technology, which is really quite different, a camera interprets an image as data.

George Sims

March 2006