INTERNATIONAL STANDARD

ISO/IEC 24778

First edition
2008-02-15

Information technology — Automatic identification and data capture techniques — Aztec Code bar code symbology specification

Technologies de l'information — Techniques d'identification automatique et de capture des données — Spécification pour la symbologie de code à barres du code Aztec

Reference number
ISO/IEC 24778:2008(E)

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

Contents

Page

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work. In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of the joint technical committee is to prepare International Standards. Draft International Standards adopted by the joint technical committee are circulated to national bodies for voting. Publication as an International Standard requires approval by at least 75 % of the national bodies casting a vote.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

ISO/IEC 24778 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 31, *Automatic identification and data capture techniques*.

Introduction

Aztec Code is a two-dimensional matrix symbology whose symbols are nominally square, made up of square modules on a square grid, with a square bullseye pattern at their center. Aztec Code symbols can encode from small to large amounts of data with user-selected percentages of error correction.

Manufacturers of bar code equipment and users of the technology require publicly available standard symbology specifications to which they can refer when developing equipment and application standards. The publication of standardised symbology specifications is designed to achieve this.

Information technology — Automatic identification and data capture techniques — Aztec Code bar code symbology specification

1 Scope

This International Standard defines the requirements for the symbology known as Aztec Code. It specifies the Aztec Code symbology characteristics including data character encodation, rules for error control encoding, the graphical symbol structure, symbol dimensions and print quality requirements, a reference decoding algorithm, and user-selectable application parameters.

2 Normative references

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO/IEC 646:1991, *Information technology — ISO 7-bit coded character set for information interchange*

ISO/IEC 15415:2004, *Information technology — Automatic identification and data capture techniques — Bar code print quality test specification — Two-dimensional symbols*

ISO/IEC 15424, *Information technology — Automatic identification and data capture techniques — Data Carrier Identifiers (including Symbology Identifiers)*

ISO/IEC 19762 (all parts), *Information technology — Automatic identification and data capture (AIDC) techniques — Harmonized vocabulary*

AIM Inc. International Technical Specification: Extended Channel Interpretations

– Part 1, Identification Schemes and Protocols

– Part 2, Registration Procedure for Coded Character Sets and Other Data Formats

– Character Set Register

3 Terms, definitions, symbols and functions

3.1 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO/IEC 19762 and the following apply.

3.1.1
bullseye
set of concentric square rings used as the finder pattern in Aztec Code

3.1.2
checkword
codeword which is included in a symbol for error correction and/or error detection

3.1.3
dataword
codeword which is part of the data message encoded in a symbol

3.1.4
domino
2-module sub-structure of the symbol character in Aztec Code which is the elemental entity used in graphical encoding of the symbol

3.1.5
Mode Message
short fixed-length, error-corrected subsidiary message within an Aztec Code symbol which directly encodes the symbol's size and data message length

3.2 Symbols and functions

3.2.1 Mathematical symbols

For the purposes of this document, the following mathematical symbols apply.

B the number of bits in each codeword

C_b the symbol capacity in number of bits

C_w the symbol capacity in number of codewords

D the number of data (message) codewords in the symbol

K the number of error correction codewords in the symbol, equal to C_w - D

L the number of data layers (1 to 32) in the symbol, defining its size

m the symbology identifier modifier value

X the X-dimension or nominal square grid spacing

x a general variable used to express error correction polynomials

(x,y) Cartesian coordinates within the module grid

3.2.2 Mathematical functions and operations

For the purposes of this document, the following mathematical functions and operations apply.

abs() is the absolute value function

div is the integer division operator

max(a,b) is the greater of a and b

mod is the remainder after integer division

4 Symbology characteristics

4.1 Basic characteristics

Aztec Code is a two dimensional matrix symbology with the following basic characteristics:

a. Encodable character set:

   1. All 8-bit values can be encoded. The default interpretation shall be:

      a. for values 0 to 127, in accordance with the U.S. national version of ISO/IEC 646;

      (NOTE: This version consists of the GO set of ISO/IEC 646 and the CO set of ISO/IEC 6429 with values 28 to 31 modified to FS, GS, RS and US respectively.)

      b. for values 128 - 255, in accordance with ISO/IEC 8859-1.

   This interpretation corresponds to ECI 000003.

   2. Two non-data characters can be encoded, FNC1 for compatibility with some existing applications and ECI escape sequences for the standardized encoding of message interpretation information.

b. Representation of data: A dark module is a binary one and a light module is a binary zero.

c. Symbol size:

   1. The smallest Aztec Code symbol is 15 x 15 modules square, and the largest is 151 x 151.

   2. No quiet zone is required outside the bounds of the symbol.

d. Data capacity (at recommended error correction level):

   1. The smallest Aztec Code symbol encodes up to 13 numeric or 12 alphabetic characters or 6 bytes of data.

   2. The largest symbol encodes up to 3832 numeric or 3067 alphabetic characters or 1914 bytes of data.

e. Selectable error correction:

   1. User-selectable, from 5 % to 95 % of the data region, with a minimum of 3 codewords.

   2. Recommended level is 23 % of symbol capacity plus 3 codewords.

f. Code type: Matrix

g. Orientation independent: Yes

4.2 Summary of additional features

The following summarizes additional features that are inherent or optional in Aztec Code:

a. Reflectance Reversal (Inherent): Though Aztec Code symbols are shown and described in this specification always with the finder's center dark and with dark modules encoding binary 1s throughout, symbols exhibiting the opposite reflectance characteristics are easily autodiscriminated and decoded with the standard reader.

b. Mirror Image (Inherent): Images which contain an Aztec Code symbol in mirror reversal, either because they are obtained using a reflected optical path, a reversed scan direction, or from behind through a clear substrate, are easily autodiscriminated and decoded with the standard reader.

c. Extended Channel Interpretation (Optional): The ECI mechanism enables characters from various character sets (e.g. Arabic, Cyrillic, Greek, Hebrew) and other data interpretations or industry-specific requirements to be represented.

d. Structured Append (Optional): Structured Append allows files of data to be represented logically and continually in up to 26 Aztec Code symbols. The symbols may be scanned in any sequence to enable the original data to be correctly reconstructed.

e. Reader Initialization Symbols (Optional): A distinct format of Aztec Code symbol is available for use in barcode menus for reader initialization. The encoded message in these special symbols is never passed on to an application.

f. Aztec "Runes" (Optional): a series of 256 small, machine-readable marks compatible with Aztec Code are available for special applications. See Annex A.

5 Symbol description

Aztec Code symbols are nominally square, made up of square modules on a square grid, with a square bullseye pattern at their center. Figure 1 shows two representative Aztec Code symbols, a small 1-layer symbol on the left which encodes 12 digits with 47 % error correction and a larger 6-layer symbol on the right which encodes 168 text characters with 30 % error correction.

Figure 1 — Representative Aztec Code symbols

These symbols illustrate the two basic formats of Aztec Code symbols: on the left is a "compact" Aztec Code symbol, visually characterized by a 2-ring bullseye, useful for encoding shorter messages efficiently, while on the right is a "full-range" Aztec Code symbol, visually characterized by a 3-ring bullseye, which supports much larger symbols for longer data messages. Since encoders can autoselect and decoders autodiscriminate between the two formats, a seamless transition is achieved to cover the full spectrum of applications.

5.1 Symbol structure

The underlying structure of a compact Aztec Code symbol is shown in Figure 2, and that of a full-range Aztec Code symbol is shown in Figure 3. In both cases, the Aztec Code symbol has at its center a Core Symbol which is then surrounded by data fields on all four sides.

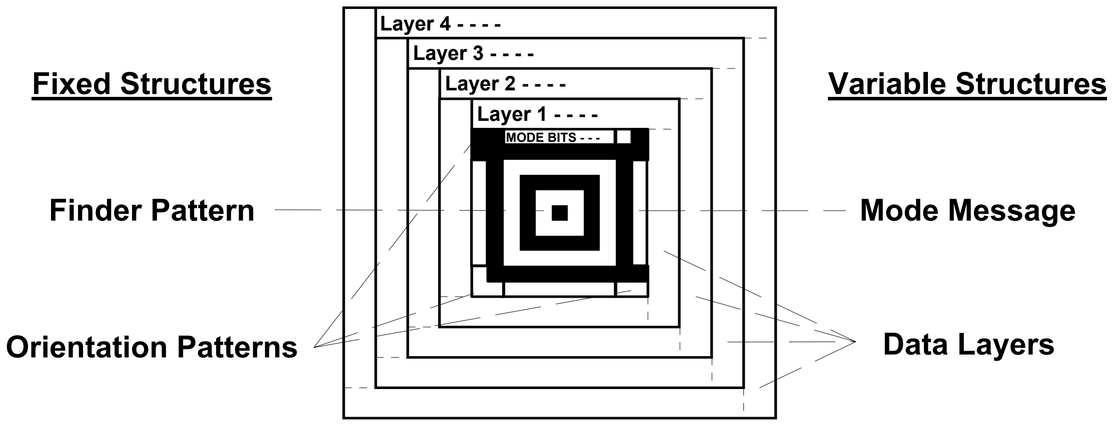

Figure 2 — Structure of a “compact” Aztec Code symbol

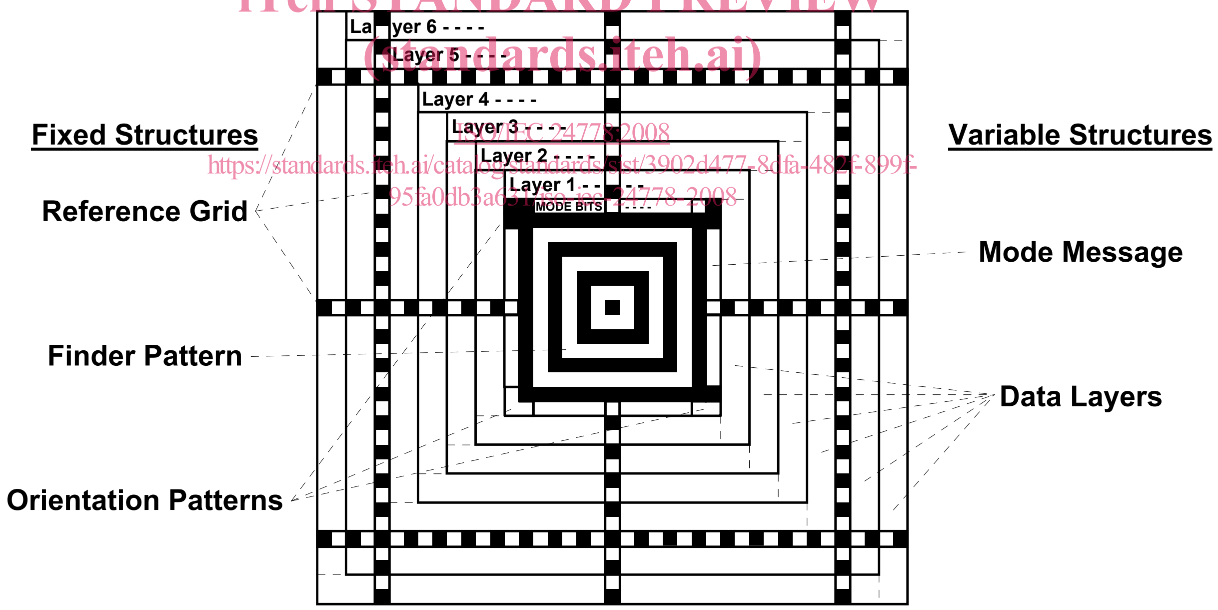

Figure 3 — Structure of a “full-range” Aztec Code symbol

5.1.1 Core Symbol

The Core Symbol, always square and at the exact center of an Aztec Code symbol, consists of a finder pattern, orientation patterns, and a Mode Message. This core covers an 11x11 module area in compact symbols and a 15x15 module area in full-range symbols. It is called the Core “Symbol” because it must be successfully found and decoded before decoding can proceed into the surrounding data fields.

5.1.1.1 Finder pattern

The finder pattern in Aztec Code is a set of concentric square rings. Centered on a single dark module, there is a ring of light modules surrounded by a larger ring of dark modules, and so forth outward to a second 9x9 module dark ring in compact symbols and to a third 13x13 module dark ring in full-range symbols.

5.1.1.2 Orientation patterns

Four 3-module chevron-shaped orientation patterns are located at the corners of the finder pattern. The upper lefthand pattern is all dark and the lower lefthand pattern is all light, while the upper righthand pattern has 2 modules dark and the lower right pattern has just 1 module dark, as shown in Figures 2 and 3.

5.1.1.3 Mode Message

The single layer of bits adjoining the finder pattern, excluding the orientation patterns and in full-range symbols also excluding the center bit along each side (which is part of the reference grid), comprises an error-corrected Mode Message wrapped in a clockwise direction starting from the upper left corner. This message explicitly encodes both the number of data layers in the overall symbol (and thus its size) and the number of datawords in those layers, the rest being error correction checkwords for that message. Encoding details for the Mode Message are given in 7.2.

5.1.2 Data fields

The data fields symmetrically surround the Core Symbol with one or more data layers. In full-range symbols, a reference grid also threads throughout the data fields.

5.1.2.1 Reference grid

The reference grid, clearly evident in Figure 3, is a ladder-like extension of the dark/light periodicity in the finder along every 16th row and column of the symbol, extending to the limit of the data fields. Its regular structure provides outlying reference points often needed to accurately map the data field in the larger full-range symbols. Compact Aztec Code symbols, which are of limited size, have no reference grid structure.

5.1.2.2 Data layers

The message data themselves plus their error correction words are laid into 2-module thick layers, spiralling clockwise from the upper left corner of the Core Symbol outward, and in full-range symbols necessarily skipping over module positions occupied by the reference grid. Compact Aztec Code symbols can have from one to four data layers, while full-range Aztec Code symbols can have from one up to 32 data layers. Details of the message encoding, codeword formation, error correction encoding, and the final laying of codewords into the data layers are given in 7.3.

5.2 Symbol character structure and sequence

In order to enhance Reed-Solomon error correction performance, the codewords, and thus the symbol characters, vary in size from 6-bits up to 12-bits depending on the overall symbol size, as illustrated in Figure 4.

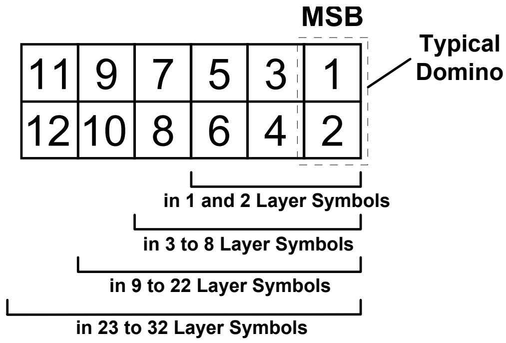

Figure 4 — Symbol character structure

While spiralling around the core, turning corners and occasionally skipping across the reference grid, the symbol characters' actual shapes vary widely. However, if they are regarded as sequences of "dominos", each 2 modules tall by 1 module wide, then the sequence of symbol characters becomes a sequence of dominos whose placement is highly systematic throughout the data fields, and thus easy to place during encoding and easy to map during decoding. Figure 5 shows how the sequence of dominos is positioned as it turns corners and transitions between data layers.

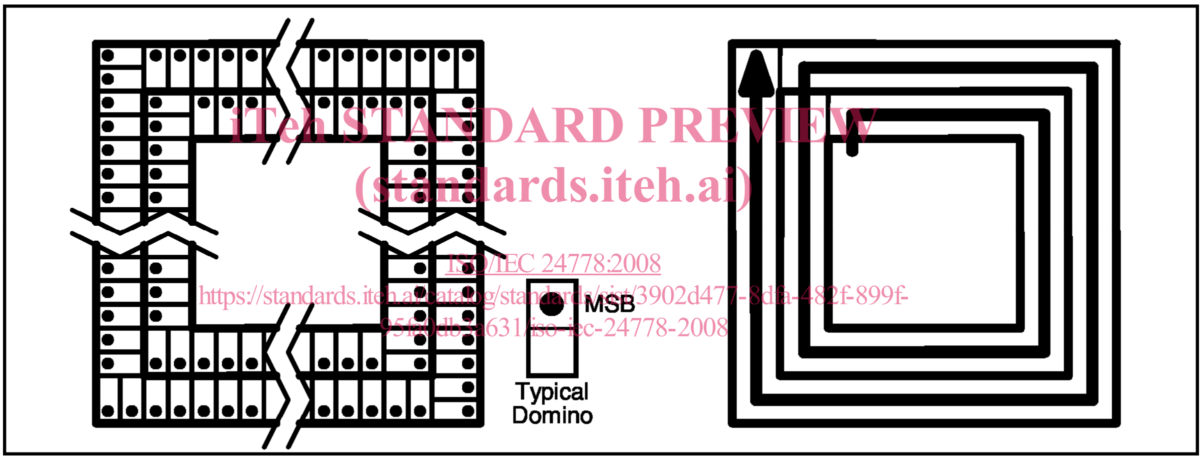

Figure 5 — "Domino" layout and sequencing

The sequence of codewords that spirals outward from the Core Symbol is in fact reversed from its natural order: the first codeword in the spiral is the last Reed-Solomon checkword, followed then by the immediately preceeding checkword, and so forth through the check words and then through the message codewords, until the final codeword in the outermost data layer is the first codeword of the encoded message. This arrangement enhances error correction by locating the data codewords (which have erasure detection because data codewords with all zeros or all ones are illegal) near the symbol's perimeter where erasures are more likely to occur.

5.3 Symbol size and capacity

Table 1 lists the overall size and capacities of the different sized Aztec Code symbols.

The data capacities shown are based on the recommended error correction levels. They represent approximate limits because the message encoding efficiency depends in a detailed way on message content and because the error correction level is user adjustable.

Table 1 — The size and capacities of Aztec Code symbols

# of Data Layers	Symbol Size (in x)	Codeword Count x Size	Symbol Bit Capacity	Symbol Data Capacities		
				Digits	Text	Bytes
1*	15 x 15	17 x 6	102	13	12	6
1	19 x 19	21 x 6	126	18	15	8
2*	19 x 19	40 x 6	240	40	33	19
2	23 x 23	48 x 6	288	49	40	24
3*	23 x 23	51 x 8	408	70	57	33
3	27 x 27	60 x 8	480	84	68	40
4*	27 x 27	76 x 8	608	110	89	53
4	31 x 31	88 x 8	704	128	104	62
5	37 x 37	120 x 8	960	178	144	87
6	41 x 41	156 x 8	1248	232	187	114
7	45 x 45	196 x 8	1568	294	236	145
8	49 x 49	240 x 8	1920	362	291	179
9	53 x 53	230 x 10	2300	433	348	214
10	57 x 57	272 x 10	2720	516	414	256
11	61 x 61	316 x 10	3160	601	482	298
12	67 x 67	364 x 10	3640	691	554	343
13	71 x 71	416 x 10	4160	793	636	394
14	75 x 75	470 x 10	4700	896	718	446
15	79 x 79	528 x 10	5280	1008	808	502
16	83 x 83	588 x 10	5880	1123	900	559
17	87 x 87	652 x 10	6520	1246	998	621
18	91 x 91	720 x 10	7200	1378	1104	687
19	95 x 95	790 x 10	7900	1511	1210	753
20	101 x 101	864 x 10	8640	1653	1324	824
21	105 x 105	940 x 10	9400	1801	1442	898
22	109 x 109	1020 x 10	10200	1956	1566	976
23	113 x 113	920 x 12	11040	2116	1694	1056
24	117 x 117	992 x 12	11904	2281	1826	1138
25	121 x 121	1066 x 12	12792	2452	1963	1224
26	125 x 125	1144 x 12	13728	2632	2107	1314
27	131 x 131	1224 x 12	14688	2818	2256	1407
28	135 x 135	1306 x 12	15672	3007	2407	1501
29	139 x 139	1392 x 12	16704	3205	2565	1600
30	143 x 143	1480 x 12	17760	3409	2728	1702
31	147 x 147	1570 x 12	18840	3616	2894	1806
32	151 x 151	1664 x 12	19968	3832	3067	1914

* An asterisk indicates a “compact” symbol; the rest are “full-range” symbols.

NOTE: Full range symbols with 1, 2, or 3 layers are useful only for reader initialization.

6 General encodation procedures

The following steps are required to convert data into the encoded form represented in an Aztec Code symbol. The following clauses of this specification specify all the rules and procedures. An encoding example is shown in Annex G.

1. Data from a 256 character set may be encoded in Aztec Code. The input message is presented in a stream of byte values reading from left to right. Special FNC1 or ECI flag characters may be inserted at any points in the stream.

2. Each message character is translated into 4, 5, or 8 bits, preceded by additional 4 or 5-bit shift and latch codes as needed, forming a long continuous data bit stream.

3. The minimum number of bits to be encoded is computed by taking the length of the data bit stream and adding as many bits as needed to reach either the default or user-specified error correction percentage. From this calculation, the format and minimum size (number of data layers L) of the symbol is selected using Table 1. This in turn establishes both the codeword size B and overall symbol capacity in codewords C_w.

4. The data bit stream is laid into codewords, systematically avoiding the formation of any codewords containing all 0's or all 1's, thus creating D message codewords.

5. The number K of checkwords becomes C_w minus D. Systematic Reed-Solomon encoding, based on a Galois Field of size 2^B and using a generator polynomial of order K, is employed to generate K additional check codewords which are appended to the sequence of message codewords.

6. The binary values of L and D are formed into a Mode Message, and systematic Reed-Solomon encoding based on GF(16) is employed to generate additional check bits.

7. Graphically, the L-layer symbol is constructed by placing modules first for the fixed structures of the finder, orientation patterns, and (if full-range) reference grid, then for the Mode Message wrapping around the finder, and finally for the spiralling layers of dominos which constitute the sequence of datawords and checkwords taken in reverse order.

7 Symbol structure

7.1 Fixed pattern structures

An Aztec Code symbol contains three types of fixed pattern - the finder, orientation bits, and if full-range a reference grid. These are all shown in Figures 2 and 3. Their specification is facilitated by regarding the symbol grid as a Cartesian plane with (0,0) at the symbol's center and with the x-axis pointing to the right and the y-axis pointing upward.

7.1.1 The finder

The finder pattern is a square bullseye target filling the central square region whose corner modules are centered at (-F,-F), (-F,F), (F,F), and (F,-F) where F is 4 in compact symbols and 6 in full-range symbols. Where the value "0" represents a light module and "1" represents a dark module, all the modules within the finder region can be defined as encoding:

$$((\max(\text{abs}(x),\text{abs}(y))) + 1) \bmod 2$$

7.1.2 The orientation bits

The orientation bits are four groups of three modules attached to the corners of the finder. With F as defined above, six dark modules are located at (-F-1,F), (-F-1,F+1), (-F,F+1), (F+1,F+1), (F+1,F) and (F+1,-F) and six light modules are located at (F,F+1), (F+1,-F-1), (F,-F-1), (-F,-F-1), (-F-1,-F-1) and (-F-1,-F).

7.1.3 The reference grid

The reference grid extends throughout full-range Aztec Code symbols occupying every location where x is a multiple of 16 or y is a multiple of 16, i.e. ((x mod 16) = 0) or ((y mod 16) = 0). Again where the value "0" represents a light module and "1" represents a dark module, all the modules within the reference grid can be defined as encoding:

$$(x + y + 1) \bmod 2.$$

Where the reference grid and finder overlap, they have the same values.