

DataMatrix 二维条码白皮书

Data Matrix 二维条码原名 Datacode，由美国国际资料公司(International Data Matrix, 简称 ID Matrix)於 1989 年发明。Data-Matrix 二维条码是一种矩阵式二维条码，其发展的构想是希望在较小的条码标签上存入更多的资料量。Data Matrix 二维条码的最小尺寸是目前所有条码中最小的，尤其特别适用於小零件的标识，以及直接印刷在实体上。

Data Matrix 二维条码的发展

Data Matrix 二维条码原名 Datacode，由美国国际资料公司(International Data Matrix, 简称 ID Matrix)於 1989 年发明。Data-Matrix 二维条码是一种矩阵式二维条码，其发展的构想是希望在较小的条码标签上存入更多的资料量。Data Matrix 二维条码的最小尺寸是目前所有条码中最小的，尤其特别适用於小零件的标识，以及直接印刷在实体上。

Data Matrix 二维条码又可分为 ECC000-140 与 ECC200 两种类型，ECC000-140 具有多种不同等级的错误纠正功能，而 ECC200 则透过 Reed-Solomon 演算法产生多项式计算出错误纠正码，其尺寸可以依需求印成不同大小，但采用的错误纠正码应与尺寸配合，由於其演算法较为容易，且尺寸较有弹性，故一般以 ECC200 较为普遍，本节也仅介绍 ECC200 规格的 Data Matrix 二维条码，以下所说的 Data-Matrix 二维条码事实上都是指 ECC200 而言。



图 Data Matrix 二维条码的外观

如图所示，Data Matrix 二维条码的外观是一个由许多小方格所组成的正方形或长方形符号，其资讯的储存是以浅色与深色方格的排列组合，以二位元码(Binary-code)方式来编码，故电脑可直接读取其资料内容，而不需要如传统一维条码的符号对映表(Character Look-up Table)。深色代表“1”，浅色代表“0”，再利用成串(String)的浅色与深色方格来描述特殊的字元资讯，这些字串再列成一个完成的矩阵式码，形成 DataMatrix 二维条码，再以不同的印表机印在不同材质表面上。由於 Data Matrix 二维条码只需要读取资料

的 20%即可精确辨读，因此很适合应用在条码容易受损的场所，例如印在暴露於高热、化学清洁剂、机械剥蚀等特殊环境的零件上。

Data Matrix 二维条码的尺寸可任意调整，最大可到 14 平方英寸，最小可到 0.0002 平方英寸，这个尺寸也是目前一维与二维条码中最小的，因此特别适合印在电路板的零组件上。另一方面，大多数的条码的大小与编入的资料量有绝对的关系，但是 **Data Matrix** 二维条码的尺寸与其编入的资料量却是相互独立的，因此它的尺寸比较有弹性。此外，**Data Matrix** 二维条码最大储存量为 2,000bytes，自动纠正错误的的能力较低，只适用特别的 CCD 扫描器来解读。

Data Matrix 二维条码的结构

(一) 特性

Data Matrix 二维条码具有以下特性：

可编码字元集包括全部的 ASCII 字元及扩充 ASCII 字元，共 256 个字元。

条码大小(不包括空白区)：10×10 ~ 144×144

资料容量：235 个文数字资料，1556 个 8 位元资料，3116 个数字资料。

错误纠正：透过 Reed-Solomon 演算法产生多项式计算获得错误纠正码。不同尺寸宜采用不同数量的错误纠正码。

(二) 基本结构

每个 **Data Matrix** 二维条码符号由规则排列的方形模组构成的资料区组成，资料区的四周由定位图形(Finder Pattern)所包围，定位图形的四周则由空白区包围，资料区再以排位图形(Alignment Patterns)加以分隔。

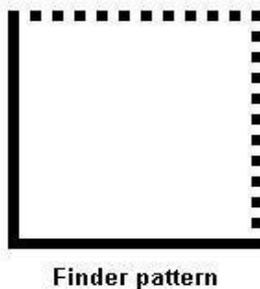


图 1 定位图形

定位图形

定位图形是资料区域的一个周界，为一个模组宽度。其中两条邻边为暗实线，主要用於限定物理尺寸；定位和符号失真。另两条邻边由交替的深色和浅色模组组成，主要用於限定符号的单元结构，但也能帮助确定物理尺寸及失真。

符号尺寸

ECC000-140 符号有奇数行与奇数列。符号外观为一方形矩阵，尺寸从 **9×9** 至 **49×49**，不包括空白区。这些符号可透过右上角深色方格识别出来。

ECC200 符号有偶数行与偶数列。有些符号是正方形，尺寸从 **10×10** 至 **144×144**，不包括空白区。有些是长方形，尺寸从 **8×18** 至 **16×48**，不包括空白区。所有的 **ECC200** 符号都可以透过右上角浅色方格识别出来。

资料表示方法

Data Matrix 二维条码按以下步骤来表示资料：

资料编码

先分析要表示的资料，选取合适的编码方案，按所选定的方案将资料流转为字码流，并加入必要的填字，如果使用者未规定矩阵寸，则应选取能满足要存放资料的最小尺寸。

DataMatrix 二维条码共有 **6** 种编码方案，即 **6** 种字码集，见表 **10.1**。

Data Matrix 二维条码的编码方案与相对应之字元集

编码方案	字元集
ASCII	十进位数字 ASCII 值 0~127 扩展 ASCII 值 128~255
C40	基本大写文数字型
Text	基本小写文数字型
EDIFACT	32~94
Base256	0~255
X12	ANSI X12 EDI 数据集

错误检测和纠正字码(ECC)的产生

对少於 255 个字码的 Data Matrix 二维条码，错误纠正字码可由资料字码计算得出。对於多於 255 个字码的符号，应将资料字码分成多个模组，然後再产生每一个模组的错误纠正字码。错误纠正字码能够纠正两种类误字码，包括 E 错误(已知位置上的错误字码)，以及 T 错误(未知位置上的错误字码)。换句话说，E 错误是不能被扫描或不能被解码的符号字元，T 错误则是被错误解码的符号字元。

图像处理上的预处理：

对图像进行增强处理，提高图像的识别准确率。实验中通过平滑滤波方法，减少图像噪声，很好地解决了采集金属零件的 DataMatrix 码时，由于码符号边沿亮度过亮影响图像分割问题；然后对图像进行直方图均衡化，扩大对比度的动态范围，解决由于光照或摄像头的原因，造成采集的图像偏暗，对比度不够显著，引起图像中明暗模糊不清的问题。

由于采集后的图像有很多无用背景，DataMatrix 符号所在区域只占整个图像很上的比重。采用遮罩的方法，用一个固定位置的子缓冲区限制图像处理区域，忽略区域外的图像，实现 DataMatrix 的符号提取。最后用函数直接译码，并将译码结果放在指定的字符串中，用显示语句在屏幕上打印结果。

金属零件上的 DataMatrix 码是气动打印而成的成点阵式，与标准的 DataMatrix 符号

不完全一样，其点间空隙大。如对这种码毫无处理地进行识别，则识别率会很低。为了解决这个问题，采用数学形态学的膨胀算法。为了提高识别准确度，可以将 DataMatrix 符号膨胀若干次，缩小数据单位之间的空隙。这样，计算机在“寻找” DataMatrix 的“L”型寻边区时就容易准确得多。

Data Matrix (DM) 编码规则



图 2 DM 码的符号结构

如图 2 所示，DM 码的符号结构由寻边区和数据区组成。其中，寻边区包括 L 型的实心边界（图 2 (b)）和反 L 型的虚线边界（图 2 (c)）。在寻边区外层有宽度为一个数据单位的静区。寻边区是 DM 二维条码的边界，主要用于限定 DM 码的物理尺寸，定位和符号失真。反 L 型的虚线边界同样主要用于限定符号的单元结构，但也能帮助确定物理尺寸及失真。图 2 (d) 所示的数据区则包含着编码信息，是待编码的符号，包括数字、字母和汉字等按照一定的编码规则生成。DM 码由黑白两种颜色组成的点阵组合，每一个相同大小的黑色或白色方格称为一个数据单位，分别代表二进制的 0 和 1。图 2 (e) 是结构连接的 DM 码，中间一个黑白交替的十字型称为铁路线，在取样时需要利用，以提高识别率。