

集，使用 3 个字节的转义序列可以切换到 JIS X 0201 (JIS-Roman) 字符集或大得多的 JIS X 0208-1978 和 JIS X 0208-1983 字符集。

表 16-3 中列出了这些转义序列。实际上，日文文本以 ESC \$ @ 或 ESC \$ B 开始，以 ESC ( B 或 ESC ( J 结束。

表16-3 iso-2022-jp的字符集切换转义序列

| 转义序列     | 转义后的字符集                     | 每个代码的字节数 |
|----------|-----------------------------|----------|
| ESC ( B  | US-ASCII                    | 1        |
| ESC ( J  | JIS X 0201-1976 (JIS Roman) | 1        |
| ESC \$ @ | JIS X 0208-1978             | 2        |
| ESC \$ B | JIS X 0208-1983             | 2        |

在 US-ASCII 或 JIS-Roman 模态下，每个字符使用单个字节。当使用更大的 JIS X 0208 系列的字符集时，每个字符代码使用 2 个字节。该编码把发送的字节的值域范围限制在 33~126 之间<sup>13</sup>。

#### 4. euc-jp

euc-jp 是另一种流行的日文编码。EUC 代表 “Extended Unix Code” (扩展 Unix 代码)，最早是为了在 Unix 操作系统上支持亚洲字符而开发的。

和 iso-2022-jp 类似，euc-jp 编码也是变长的，允许使用几种标准的日文字符集。但和 iso-2022-jp 不同的是，euc-jp 编码不是模态的。没有转义序列可以在不同模态之间切换。

euc-jp 支持 4 种编码后的字符集：JIS X 0201 (JIS-Roman，对 ASCII 进行一些日文替换)、JIS X 0208、半宽片假名（最早在日文电报系统中使用的 63 个字符）以及 JIS X 0212。

编码 JIS Roman (它和 ASCII 兼容) 的时候使用 1 个字节，对 JIS X 0208 和半宽片假名则使用 2 个字节，而对 JIS X 0212 使用 3 个字节。这种编码有点浪费空间但处理起来很简单。

表 16-4 概括了此编码的格局。

注 13：尽管每个字节只能有 94 个不同的值(33~126)，这也足够覆盖 JIS X 0208 系列字符集里面的所有字符了，因为这些字符集是按照 94×94 的网格来组织代码值的，所以 2 个字节足以覆盖 JIS X 0208 字符集中的全部字符代码。