

数字视频基本知识探讨

武汉万德数码技术有限公司 程建刚 2002年9月2日

由于工作的需要，我在数字视频处理方面做了几年的开发，我所说的开发是指应用层的开发，而不是底层的研究。先后开发出了几个产品，比如数字接口摄像头、图像采集卡、TV-VGA 视频转换器、造纸行业用的在线斑点检测仪、数字伪彩色处理仪等（这些产品的资料已发布在网站 <http://www.wondur.com> 上），在开发中也了解了一些关于 CMOS 面阵传感器、CCD 面阵传感器、CCD 线阵传感器、高速摄像、高分辨率摄像、视频编解码、图像压缩等方面的知识，目前我还在继续做类似的开发。前面说的产品都有一定的销量，所以经常会收到用户打来的电话，大多数是咨询一些很基本的数字视频知识。数字视频无疑是复杂的，牵扯到很多方面，所以就有了写这篇文章的想法，我希望用尽量通俗的方式来说明一些基本的问题，以期能对读者的开发工作有一些帮助（说明一下：为了尽量地把问题说得简单一点，叙述中我作了一些简化，比如把 U 和 Cb 都说成 U、V 和 Cr 都说成 V 等。另外还需要说明的是，下面我所说的模拟视频的各个参数都是针对我国的电视标准——PAL 制式而言的，而其它的电视制式则略有差别。读者在做进一步研究时请注意）。因为我的水平十分有限，错漏难以避免，有兴趣的读者可以写信和我一起来探讨（我的 e-mail: db200@netease.com 或 cjpg@wondur.com）。

视频简单地讲就是活动图像。电影也是活动图像，电影把大量静止图像记录在胶片上，一张一张地连续显示出来，就成了我们看到的电影。对于电视视频，每秒钟包含几十帧静止图像，每一帧静止图像由几百个行组成，每一行又由几百个像素点组成。1 秒钟包含的图像帧数为帧频，1 秒钟包含的总行数为行频，1 秒钟包含的总像素数实际上就是相当于视频带宽。

首先来说模拟视频。

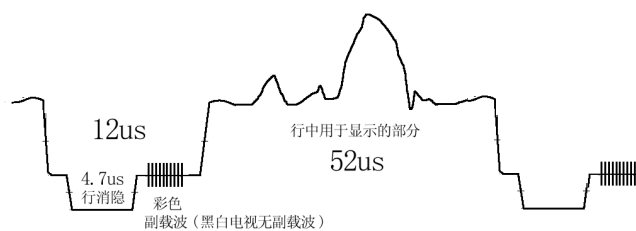
遥想起来，在当年有限的条件下，用电子管做出电视机的前辈们实在是令人佩服的，但是，用现在的眼光来看，当年的电视技术又实在是很简陋的。我国的黑白视频信号帧频为 25 Hz，就是说每秒显示 25 幅图像，之所以规定为 25 Hz，而不是其它的某个值，主要是因为两个原因。一个与人的视觉生理特点有关，当图像的刷新速度达到 5 帧/秒的时候，人开始感觉图像是活动的，而达到 24 帧/秒的时候，人感觉图像是完全连续和流畅的（电影所使用的帧频就是 24 Hz），所以视频信号帧频应大于等于 24 Hz。理论上来说帧频越高越好，但是帧频越高，对电路的要求也越高，技术越复杂，成本也越高（现在有的电脑彩显帧频已达到 200Hz），在当时的条件下，只能选择一个大于 24 的尽量小的值。另一个原因是因为我国的电网频率是 50Hz，当采用 25 Hz 帧频时，隔行扫描时的场频为 50 Hz，正好与电网同频，这样，电源对图像的干扰是固定的，人眼不容易感觉出来，或者换个角度说：电源就可以不必做得那么精密（那个时候开关电源还很少使用，整流电源缺乏稳压措施，纹波很大）。所以选择了 25 Hz 帧频。说到这里，又要说说“场频”。电视在显示图像的时候，把一帧分成了两场来显示，一个场由帧中的奇数行组成，叫做奇场，另一个场由帧中的偶数行组成，叫做偶场。之所以要这样做，主要是因为 CRT 显象管上一秒钟显示 25 帧图像时，人眼感觉到连续性还是不太好，而且还有明显的闪烁，一帧分成两场后，场频为 50 Hz，图像更加连续一些。当然还有一些别的原因，与电路设计方面有关。

我国的黑白视频信号规定每帧图像共 625 行，每场为 312.5 行，行频 15625 Hz，视频带宽 6MHz。在每场的 312.5 行中，有一些行要用作场消隐，是不包含视频信号的，按照 CCIR656 标准规定的行编号方法，奇场的行号为第 1 至 312.5 行，偶场的行号为第 312.5 至 625 行，其中，奇场的第 23.5 至 310 行包含有效的视频信号，共 287.5 行，偶场的第 336 至 622.5 行包含有效的视频信号，共 287.5 行。所以一帧中有效的总行数为 576 行，由最上面的半行加上中间的 574 行加上最下面的半行组成。（说到这里又要顺便说个问题，有经验的读者在 PC 机上用视频捕捉卡捕捉图像的时候，如果分辨率设定为 720 点*576 行，会发现最上面的行左边半边是黑色的，或者最下面的行右边半边是黑色的，就是因为一场中包含有半行的无效行的原因。）

上面说的是黑白视频信号，那么彩色视频信号又是怎么回事呢？当科学家们开始研究彩色电视的时候，黑白电视已经大量使用了，所以最好彩色电视信号能够与黑白电视信号保持兼容，以便彩色电视信

号在黑白电视机上能播放出黑白的图像。这个问题是很困难的，因为色度信号也要占用较大的带宽，而在电视射频频带上，一个频道挨着一个频道，亮度信号（实际上还有调频的伴音信号）已经把频带给塞满了，幸运的是，后来终于运用压缩亮度信号

带宽、大面积着色等技术解决了这个问题。从频域的角度来看，色度信号（UV 色差信号）是插在亮度信号的频谱的间隙之中的，具体的位置是插在 4.43MHz 的地方，带宽 1.3MHz。在接收机中，简单地说，把收到的信号中的 4.43MHz 处的带宽 1.3MHz 的信号取出来就成了色度信号，而把收到的信号中的 4.43MHz 处的这个信号滤掉剩下的就是亮度信号。很多硬件工程师可能不容易理解这是什么意思，不过没关系。下面我们从时域的角度来看看它们的一行的波形，如图所示，黑白视频的亮度信号采用了调幅制，一行的周期是 64 μs，其中显示在屏幕上的信号占 52 μs，其余部分为行消隐、行同步头。对于彩色信号，还在行同步头上叠加了一小段 4.43MHz 的副载波信号，用作接收机中的 4.43MHz 信号的频率和相位基准。



黑白及彩色电视信号一行的波形

号，而把收到的信号中的 4.43MHz 处的这个信号滤掉剩下的就是亮度信号。很多硬件工程师可能不容易理解这是什么意思，不过没关系。下面我们从时域的角度来看看它们的一行的波形，如图所示，黑白视频的亮度信号采用了调幅制，一行的周期是 64 μs，其中显示在屏幕上的信号占 52 μs，其余部分为行消隐、行同步头。对于彩色信号，还在行同步头上叠加了一小段 4.43MHz 的副载波信号，用作接收机中的 4.43MHz 信号的频率和相位基准。

上面说的是彩色视频信号怎么加到原来的黑白视频信号中。那么彩色图像是怎么还原和显示出来的呢？我们知道，黑白图像的像素只能用亮度（灰度）来描述。而彩色图像的像素的描述要复杂一些，有很多不同的方法，比如在印刷行业中用的是 CMYK（青、品红、黄、黑）四色合成的方法，而在计算机或电视机的 CRT 显象管中都是用的 RGB（红、绿、蓝）三基色合成的方法，我也不明白为什么要选这三种颜色，而没选别的颜色，我猜想可能是因为人眼包含 RGB 三种色感细胞，所以用这三种基色能合成比较多的人眼能识别的彩色出来，或者因为这三种基色的荧光粉比较容易制造。通过 RGB 三基色来合成某种彩色，或者某种彩色怎样分解为三基色，这个是大家都熟悉的。用 RGB 三基色来表示彩色的确很直观，但是如果把这种方法用作图像传输则绝不是一个好的方法。第一个缺点是与黑白图像不兼容，把 RGB 三基色转换为灰度的方法是： $\text{灰度} = R \times 0.3 + G \times 0.59 + B \times 0.11$ ，这个转换过程显然是比较复杂的。对于电视机而言，就意味着必须解码出 RGB 信号才有可能得到黑白图像，而黑白电视机没有解码功能，所以不能实现兼容。第二个缺点是占用太多带宽，用 RGB 三基色表示图像，每个分量的带宽是均等的，都约等于亮度信号的带宽，所以对于每个分量，都要用较大的带宽来描述。第三个缺点是抗干扰能力差。由于 G 分量占有亮度值的 59%，所以当 G 受到干扰的时候，像素的亮度值会受到很大的影响，而人眼对亮度值的变化是十分敏感的，所以图像主观质量会明显下降。基于这些原因，在视频信号传输中采用的是 YUV 合成的方法。Y 代表亮度信息，U 代表蓝色色差（就是蓝色信号与亮度信号之间的差值），V 代表红色色差。我们来看看使用这种表示方法的优点。第一个优点是黑白图像兼容。假定一个像素是用 YUV 表示的，我们只要忽略 UV 分量，取出 Y 分量，就可以得到像素的亮度值，从而把彩色图像转换为黑白图像。这样很容易实现彩色电视信号与黑白电视信号的兼容。第二个优点是节省带宽。说这个问题的时候要先说说大面积着色原理。实验发现，人眼对亮度信息是敏感的，主要通过亮度差别来分辨物体形状的细节，而对彩色信息是不敏感的，人眼区分不出物体颜色上的细小的变化，或者说人眼不容易觉察出来图像的色彩的细节部分的变化。因此，可以对亮度信号用较高的采样频率采样，而对色度信号用较低的采样频率采样（或者用较低的量化深度），比如几个相邻的像素的亮度值不同，但是却可以使用一个相同的色度值。这就是大面积着色原理。基于这个原理，在电视信号传输中，U 或 V 信号的带宽远小于 Y 信号的带宽，这样就节约了带宽。换个方式来说，比如在计算机中，用 RGB 方式描述一个像素需要 R、G、B 共 3 个字节。而用 YUV 方式描述，则对于每 2 个像素，Y 用 2 个字节，U 取相同的值，用一个字节，V 取相同的值，用一个字节，平均每个像素 2 个字节。或者每个像素 Y 用一个字节，U 用半个字节，V 用半个字节，共 2 个字节。第三个优点是抗干扰能力强。由于亮度信号是单独表示的，所以如果色差信号受到干扰，不会影响到亮度，主观感觉噪声不会明显增加。

在电视机中，彩色视频信号首先分解为亮度信号 Y 和色度信号，色度信号再分解为 U 色差信号和 V 色差信号，最后由 YUV 三个分量经过矩阵运算变换为 RGB 信号，以便在显像管上显示。那么 YUV 具体是怎样变换为 RGB 呢？这个问题又叫做“彩色空间变换”，我将在本文的后面详细地讨论这个问题。

通过前面的讨论，我们已经知道：黑白视频信号带宽 6 MHz，由帧、场、行、像素等组成，像素用

亮度值 Y 来描述。彩色视频信号在黑白视频信号中插入了一个带宽 1.3 MHz 的色度信号，由这个信号来得到 UV 色差信号，最后把 YUV 变换为 RGB 来描述像素。

现在我们来分析一下视频信号的缺点。1、帧频低。视频信号的帧频只有 25 Hz，必然导致图像闪烁。2、分辨率低。在一帧中有效的行只有 576 行。由于采用了隔行扫描，一帧图像要由连续的两场来拼合而成，而实际上很难保证两场中的行准确地错开（对准间隙），这进一步导致垂直方向的分辨能力损失。3、亮色串扰。亮度信号和色度信号混合在一起，解码的时候不能很好地分开，导致亮度信号和色度信号互相干扰。4、缺乏改进余地。除非重新制定新的标准，否则前面提到的三个问题在现有基础上都很难改进。视频信号格式有这么多缺点，是因为受当年制定这个标准时的技术条件制约造成的，而近年来通过在电视中加入一些数字处理的手段，比如倍频扫描（100 Hz 场频）、使用数字梳状滤波器等方法，电视的图像质量也得到一些提高。而现在正在研究的数字电视，是重新制定的全新的标准，用以获得胶片质量的图像，可能会彻底淘汰现在的视频标准和电视设备，当然这是将来的事情，也不是我讨论的主题。

我想讨论的是“数字视频”，而前面说的都是模拟视频，这是因为我所说的数字视频是现有的模拟视频的数字表示，先搞清楚了模拟视频，下面的讨论就很简单了。

视频信号起初是以模拟信号的形式保存在录像磁带上的，而现在随着数字技术的发展，可以变成数字信号保存在光盘或计算机硬盘中，当然这些离不开功能强大的计算机，实际上，在嵌入式领域，数字视频也是可以应用的，比如用单片机或 DSP 来处理数字视频数据。下面，我们就来讨论数字视频数据的格式、转换、保存、显示等方面的问题。

从哪个问题开始呢？我们先来讨论一下模拟视频信号解码并量化成连续的数字视频流后所包含的一些信号元素。前面我们已经讨论了，模拟视频信号每秒 25 帧，每帧周期 40ms，而每帧又分为 2 场，每场 20ms，先输出的是奇场，然后是偶场，奇场的行号为第 1 至 312.5 行，偶场的行号为第 312.5 至 625 行，其中，奇场的第 23.5 至 310 行包含有效的视频信号，偶场的第 336 至 622.5 行包含有效的视频信号。把模拟视频信号解码为 YUV 分量后，分别进行 A/D 量化采样，转换为数字视频流，时间上也按上面的顺序依次输出。有几个相关的国际标准对此作了一些约定。比如 1994 年国际无线电咨询委员会出台了 CCIR601 标准，这个标准主要是针对演播室的要求制定的，其中规定亮度信号的采样频率是 13.5 MHz，色度信号的采样频率是 6.75 MHz，采用 8 位 PCM 编码。这样，每行的 Y 采样数为 864 个，其中有效的 Y 采样数为 720 个。而每行的 U 或 V 采样数为 432 个，有效的 U 或 V 采样数为 360 个。平均描述每个像素的 YUV 的位数分别为 8bit、4bit、4bit，又叫做 YUV422 编码方案（当然还有很多别的方案，如 YUV411 等）。国际无线电咨询委员会还出台了 CCIR656 建议，其中规定视频数据的量化值中的 0 和 255 保留不用，而量化数据串行输出的顺序是：U0、Y0、V0、Y1、U2、Y2、V2、Y3、U4、Y5、V4、Y6，如此循环。通过上面的讨论，我们可以看出，数字视频流应包含的元素有：奇偶场指示信号 FI（有的称为 ODD）、场同步信号、行同步信号、像素时钟、YUV 数据输出。在此我们顺便计算一下数字视频流的数据量，每秒的数据量 = (720 像素 * 576 行 * 25 帧) * 2 字节 = 20736000 字节，数据率约 165Mbps，由此可见数字视频的数据量之大，数据率之高！

可以想象，保存和描述数字视频流的最简单办法当然就是记录和描述连续的一帧帧的静止图像。保存静止图像的最简单格式是 BMP 格式，就是位图。我们现在就来分析一下 BMP 文件格式。BMP 文件记录图像的方式实际上有很多种，甚至可以直接在其中记录 YUV 分量，但是我不准备讨论那么多。BMP 文件由文件头、图像特征描述、颜色表、图像数据四部分构成。为了简单起见，下面用的是 VB 的语法，下面用到的数字没特别说明的都是十进制表示。

BMP 文件头数据结构如下:

Type BitMapFileHeader	‘共 14 字节
bfType As Integer	‘2 个字节, 填入字符“BM”, 即 4D42 (16 进制)
bfSize As Long	‘4 个字节, 填入整个 BMP 文件的字节数大小
bfReserverd1 As Integer	‘2 个字节, 保留, 填入 0
bfReserverd2 As Integer	‘2 个字节, 保留, 填入 0
bfOffBits As Long	‘4 个字节, 表示图像数据在整个 BMP 文件中的起始位置
End Type	

图像特征描述块的数据结构如下:

Type BitMapInfoHeader	‘共 40 字节
biSize As Long	‘4 个字节, 表示本结构的字节数大小, 填入固定的值 40
biWidth As Long	‘4 个字节, 填入图像水平方向像素数, 其值必须是 4 的整数倍
biHeight As Long	‘4 个字节, 填入图像垂直方向像素数
biPlanes As Integer	‘2 个字节, 填入固定的值 1
biBitCount As Integer	‘2 个字节, 表示每个像素的位数, 灰度图填 8, 24 位真彩图填 24
biCompression As Long	‘4 个字节, 填入 0 表示不压缩
biXSizeImage As Long	‘4 个字节, 表示图像的像素总数
biXPelsPerMeter As Long	‘4 个字节, 填入固定值 3780
biYPelsPerMeter As Long	‘4 个字节, 填入固定值 3780
biClrUsed As Long	‘4 个字节, 填入固定值 0
biClrImportant As Long	‘4 个字节, 填入固定值 0
End Type	

在 BMP 中记录的图像数据不一定是 RGB 或 Y 的值, 可以只是一个“编号”, 而这个“编号”对应的实际 RGB 颜色要从颜色表中查出来。24 位真彩图因为直接记录了 RGB 值, 所以是不需要颜色表的, 因此在 24 位真彩 BMP 文件中没有颜色表这部分, 而其它的情况下就必须有一个颜色表。颜色表一共 256 项, 每一项 4 个字节, 前三个字节分别代表 B、G、R 值, 最后一个字节为 0。我们来以 8 位灰度图举例说明, 8 位灰度图是黑白图像, 记录的实际上就是亮度分量 Y, 根据 RGB 三基色原理, R=Y、B=Y、G=Y 的时候合成的就是一个亮度为 Y 的灰色像素, 极限的情况, 当 R=255、G=255、B=255 时表示最白的像素, 当 R=0、G=0、B=0 时表示最黑的像素。因此颜色表的值应为: 0, 0, 0, 0, 1, 1, 1, 0, 2, 2, 2, 0, 3, 3, 3, 0..... 255, 255, 255, 0。

BMP 文件的最后部分记录图像数据, 也是数据量最大的部分。一帧图像的像素数据填入 BMP 文件的顺序是: 先填最下面的一行, 从左到右依次一个个像素依次填入, 这样一行行一直到填完最上面的一行。对于 8 位灰度图, 对每个像素数据只需填入一个字节的 Y 值。对于 24 位真彩图, 要填 3 个字节, 注意: 先填 B 值, 再填 G 值, 最后是 R 值。

总结一下: 灰度图的 BMP 文件由“文件头+图像特征描述块+颜色表+图像数据 Y”组成。24 位真彩图由“文件头+图像特征描述块+图像数据 BGR”组成。如果读者对于这些不熟悉, 可以在 WINDOWS 的画图工具中画一副图, 保存为上面的格式, 然后用二进制编辑工具比如 UltraEdit 来观察和分析文件的内容。关于 BMP 文件的更多知识, 读者可以进一步查阅资料。

现在再回头讨论一下视频保存为 BMP 图像的格式。对于数字视频流，由于一帧中有效的行数是固定的 576 行，所以如果把它转换成 576 行的 BMP 图像，当然最简单，而且可以取得较好的效果。而如果取别的值，比如 600 行或者 400 行，那么就必须进行插值运算，依据现有的行计算出假想中的行，这样的转换其运算量很大，对图像的质量也有一定损失，一般要用专门的硬件来实现（一般高档的视频捕捉卡带有这样的功能）。那么是不是说只能转换成 576 行呢？当然也不是，比如可以转换成 288 行，隔一行取一行，实际上就是只采集了一场。另外还可以丢弃一些行，比如只取中间的 480 行，而上下各丢掉 48 行，这样得到的图像虽然被裁剪了，但清晰度不会下降。再来看看一行中应该采多少个像素。由于模拟视频信号幅值在一行中是连续变化的，因此采样数没有行数那样的限制。比如我们可以采 400 个像素，当然也可以采 401 个像素，这取决于我们对水平分辨率的要求，一行中所采得的像素越多，则分辨率越高，但需要注意亮度信号的带宽是有限的，采样率高到一定程度后，再提高就没有意义了。前面说了，CCIR601 标准的规定是一行采 720 个有效像素。除了这个标准外，比较常用的另一种标准是一行采 768 个有效像素。说到这儿，顺便说一下，在计算机上用视频捕捉卡捕捉图像时，驱动程序提供的图像格式是固定的几种，比如 768*576、720*576、384*288、320*288 等，为什么只提供这些格式，以前不明白的读者现在知道了吧！

现在我们知道了模拟视频变换为数字视频流后的格式，也知道了怎样创建和保存 BMP 文件，但是我们还不能把从数字视频流中截取的图像数据保存为 BMP 文件，因为数字视频流是用 YUV 描述的，而 BMP 文件是用 RGB 描述的。它们之间怎么转换呢？这就是色彩空间转换的问题。

RGB 和 YUV 的对应关系用代数方程式表示如下：

$$Y = 0.299R + 0.587G + 0.114B$$

$$U = -0.147R - 0.289G + 0.436B$$

$$V = 0.615R - 0.515G - 0.100B$$

或者：

$$R = Y + 1.14V$$

$$B = Y + 2.03U$$

$$G = Y - 0.58V - 0.39U$$

CCIR601 标准中定义的值与上面略有不同。在考虑了人的视觉系统和 CRT 显示器的非线性特性之后，推荐的转换方程如下：

$$R = Y + 1.371 V$$

$$G = Y - 0.698 * V - 0.336 U$$

$$B = Y + 1.732 U$$

不过读者要注意，上面的方程中的 UV 值因为符号扩展的原因平移了 128，实际应该使用下面的这个修正的方程：

$$R = Y + 1.371 * (V - 128)$$

$$G = Y - 0.698 * (V - 128) - 0.336 * (U - 128)$$

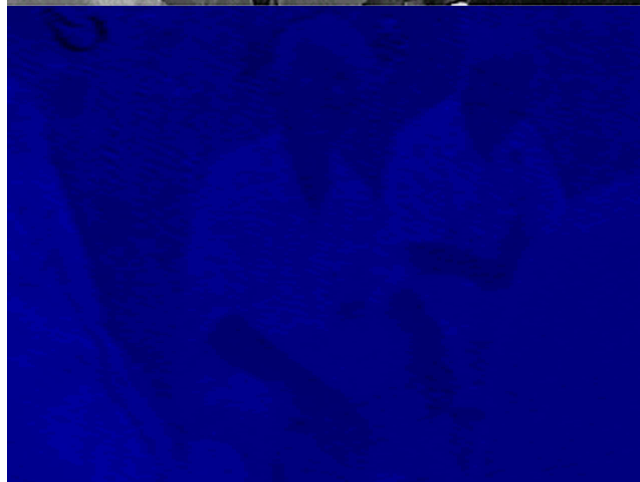
$$B = Y + 1.732 * (U - 128)$$

对于上面的方程有几点说明：1、有一些采集卡会指定 U、V 为有符号数或无符号数，使用上面的方程时不要考虑这些，而认为 Y、U、V 是从采集卡得到的一个 0 到 255 之间的无符号的值。2、用上面的公式计算出的 R、G、B 值可能会超出 0 到 255 的范围，应该在计算完后进行检查，如果值小于 0，则修正为 0，如果值大于 255，则修正为 255。上面的转换方法经过我的实际使用，证明其效果是良好的。

空谈了这么多，来点更直观的吧！我们来看一看一帧实际的图像。下边这帧图像是用武汉万德数码技术有限公司生产的 VC302 嵌入式图像采集卡采得的，信号源是电视机的视频输出（为了减小文件大小，右边的图片被压缩了 10 倍，原始图像要比现在显示的更清晰）。图像的像素分辨率是 320×240 （采集了一场中的 288 行，每行 360 个像素，但只从中截取了 240 行，每行截取了 320 个像素），第一幅是 Y 分量描述出的黑白图像，第二幅是 U 分量描述，第三幅是 V 分量描述，第四幅是 YUV 合成后的彩色图像。黑白图像与如此模糊的 UV 分量图合成后竟然可以得到色彩这么真实的彩色图像，你是不是有点难以相信？我也不敢相信，但这就是事实！从这些图的比较中我们可以得出这样的结论：1、Y 分量基本上保留了彩色图像的轮廓细节，其分辨率是较高的。2、UV 分量主观感觉十分模糊，分辨率低，只是大约地描述了整块的色彩，这正是所谓的“大面积着色”。3、我们在观察合成的彩色图像的时候并没有觉得图像的色彩模糊，这说明人眼对色彩细节的分辨能力是多么差！

接着，我们来讨论一下图像的显示问题。这个问题属于软件的范畴，大部分读者都比较熟悉，我不准备说太多。一种方法是把得到的 YUV 数据转换成 RGB 后来显示。另一种办法是直接利用 YUV 数据来显示，现在的显卡都具备把 YUV 数据通过硬件直接转换为 RGB 的能力，通过使用 Direct Draw 技术，把 YUV 数据直接提交给显卡，这样就省掉了软件转换所花费的 CPU 时间。在嵌入式领域，可以使用数字接口的液晶显示器显示，要根据其具体情况来处理。

最后，说一下视频开发中常见的一些问题。1、首先要注意区分视频信号的制式。我国的电视台发送的信号是 PAL 制信号，所以电视机的视频输出是 PAL 制。但是电视机也可以接受其它制式的信号，比如 VCD 机的视频输出有很多是 NTSC 制的，在电视机上也可以播放。在我国市场上销售的摄像机（我说的是工业摄像机，而不是家用的摄录像机）大部分是 PAL 制的，但也有的是 NTSC 制，而且有的可以通过机身上的拨码开关来设定制式。2、有些 DVD 机采用了一些特别的办法来提高清晰度，比如所谓的“逐行扫描”，它们输出的信号与标准信号相比有一些差别，在电视机上播放是不会有问题的，但是一些采集卡却不能识别，导致采集的图像紊乱，所以在调试过程中要慎用 DVD 作信号源。3、分辨率的问题。分辨率一般是



指垂直方向能分辨的等间隔排列的黑白条纹的最大线数。我们先来看看模拟视频，视频信号行正程的时间是 $52\mu\text{s}$ ，而视频信号的带宽最大 6MHz ，假定 1Hz 可以描述 2 个像素点，这样算来最多的线数 $=52\mu\text{s} \times 6\text{M} \times 2 = 624$ 线。而实际上在编辑、存储、传输、还原过程中会有很多损失，所以电视机上的图像分辨率是远低于这个值的，一般约在 240-340 线左右。对于摄像机的视频输出，有的黑白摄像机标称分辨率高达 600 线，理论上这个分辨率是有可能达到的。而彩色摄像机标称的分辨率有 380 线、420 线、480 线等规格。再来看看数字视频流，数字视频中一行采样的像素数是固定的，CCIR601 的规定是 720 个像素，如果用线来表示，极限值是 720 线，可见大于模拟视频的分辨率。在安防监控工程中，最常用的视频源是 420 线或 480 线的彩色摄像机，扣除传输过程中的损失，最后能达到 380 线就不错了，保存为图像时每行取 360 像素比较恰当，再增加像素数，清晰度虽然仍有提高，但已不太明显。在工业领域，有时对分辨率有特别苛刻的要求，这时采用标准视频信号输出的摄像机已经不可能满足要求，需要采用非标准的摄像机，它们大多是逐行扫描的，采用专门的高分辨率 CCD 面阵传感器，常见的分辨率如 2048 行*2048 像素/帧，帧频可以高达 100Hz，当然这样的视频信号无法在电视机上显示出来，一般是采用专用的采集卡采集，在计算机上处理。

4、场采集与帧采集。这个问题比较特别，理论上讲，摄像机首先应拍取一帧图像，然后分解出两场，还原的时候，两场再拼合为一帧。但实际上，常不是这样的，有的摄像机为了简化设计，是一场一场单独采集的，两场之间有 20ms 的时间差，所以两场拼合的时候，图像上活动的部分会出现错位。另外，一些视频编辑设备也是对场单独处理的，所以也存在这个问题。看电视的时候，因为图像连续变化，所以我们只会感到图像是活动的，但是如果把图像截取一帧下来看，错位就十分明显。为了避免这个问题，在对分辨率要求不高的时候可以单场采集，即连续的两场中只取一场，事实上很多采集卡也正是这样处理的。另外可以在只采一场的时候在水平方向采集较多的像素，比如共采 288 行，但每行采 720 个像素，这样至少在水平方向有较高的分辨率，可以改善模式识别的效果。而对分辨率要求更高的时候就只好使用非标准摄像机了。

(全文完)