

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.

G06T 9/00 (2006.01)

G06T 15/00 (2006.01)

H04N 13/00 (2006.01)



[12] 发明专利申请公布说明书

[21] 申请号 200480042753.3

[43] 公开日 2007 年 3 月 28 日

[11] 公开号 CN 1938727 A

[22] 申请日 2004.2.27

[21] 申请号 200480042753.3

[86] 国际申请 PCT/MX2004/000012 2004.2.27

[87] 国际公布 WO2005/083637 西 2005.9.9

[85] 进入国家阶段日期 2006.10.16

[71] 申请人 TD 视觉有限公司

地址 墨西哥墨西哥城

[72] 发明人 M·R·古铁雷斯诺韦洛

[74] 专利代理机构 北京纪凯知识产权代理有限公司

代理人 沙捷

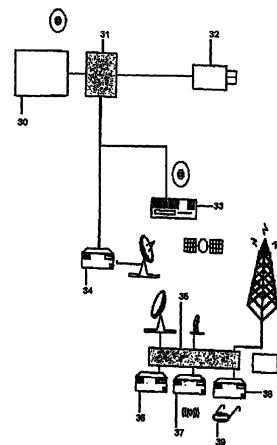
权利要求书 4 页 说明书 21 页 附图 5 页

[54] 发明名称

用于数字解码 3D 立体视频图像的方法和系统

[57] 摘要

MPEG2 可兼容立体 3D 视频图像数字解码方法和系统，使用其自己的编码算法。为了从数字视频流中获取 3D 图像，借助在解码处理的不同部分中的软件和硬件改变，对现有 MPEG2 编码器做出修改。即，通过软件对视频数据流的 video_sequence 结构进行修改，以包括 TDVision® 技术的图像类型的在比特级上的必要的标记；在解码处理中和在通过软件和硬件解码信息中进行修改，其中，激活双倍输出缓冲器，激活并行的和差别解码选择器，执行解压缩处理，显示相应的输出缓冲器；解码器必须通过软件编程以同时接收和解码每一个具有 TD-Vision® 立体识别符的两个独立的程序流。



1. 一种立体 3D 视频图像数字解码系统和方法，其中，视频数据流的 video_sequence 的结构通过软件修改，以包括图像类型的在比特级上的必要的标记，其特征在于，仅修改所述软件并且使用 user_data() 部分以存储允许重新生成立体视频信号的错误校正，由此实际上识别视频格式；在不为 TDVision® 视频的情况下，应用逻辑“与”用于 MPEG2 向后可兼容性；典型地，通过扫描所述 video_sequence 解码；当所述图像为 TDVision® 图像时：

- a) 将最近的完整图像缓冲器存储到左或右通道缓冲器中，
- b) 对于 B 类型帧信息应用差别或并行解码，
- c) 通过应用运动和颜色校正矢量，对最近获得的图像应用错误校正，
- d) 在它们各自的通道缓冲器中存储结果，
- e) 继续 video_sequence 读取。

2. 立体 3D 视频图像数字解码方法和系统，其中，如权利要求 1 所述，所述视频数据流的所述 video_sequence 结构通过软件修改，以包括所述图像类型的在所述比特级上的所述必要的标记，其特征进一步在于，解码器编译格式如下所示：

- a) 读取 video_sequence，
- b) 区分 sequence_header，如果 TDVision® 图像被识别，则激活双倍缓冲器，
- c) 在 user_data 中读取图像，如同其包含在所述结构中，
- d) 在 sequence_scalable_extension 中加入 video_sequence MPEG 信息，所述信息可被包含在所述结构中，
- e) 在 picture_header 中找到 extra_bit_picture 中的 TDVision® 图像识别符，
- f) 读取 picture_coding_extension 中的“B”类型图像，并且，如果其为 TDVision® 类型图像，则随后解码第二缓冲器，
- g) 如果所述图像是时间可扩展的，将“B”应用到所述解码器。

3. 立体 3D 视频图像数字解码方法和系统，其中，如权利要求 1 所述，所述视频数据流的所述结构和所述 video_sequence 被修改以包括所述图像类型的在所述比特级上的所述必要的标记，其特征进一步在于，当所述解码器检测到 user_data() 码时，其搜索 32 比特的 3DVision®_start_identifier = 0x000ABCD 识别符，检测到该信息时，调用比较输出缓冲器的特定解码函数，并从所述 video_sequence 的当前读取偏移处应用所述特定解码函数。

4. 立体 3D 视频图像数字解码方法和系统，其中，如权利要求 1 所述，所述视频数据流的所述 video_sequence 结构通过软件被修改，以包括所述图像类型的在所述比特级上的所述必要的标记，其特征进一步在于，所述解码器必须通过软件编程以同时接收和解码两个程序流。

5. 立体 3D 视频图像数字解码方法和系统，其中，如权利要求 1 所述，所述视频数据流的所述 video_sequence 结构通过软件被修改，以包括所述图像类型的在所述比特级上的所述必要的标记，其特征进一步在于，两个相关的视频信号可在相同的 video_sequence 中被发送；所述信号相互相关，并且来自 3DVision® 照相机；根据它们的代数相加 ($R-L=\text{delta}$)，每一个信号被存储为 B 类型帧，其解码是通过来自它们中的一个的差别进行的。

6. 立体 3D 视频图像数字解码方法和系统，其中，如权利要求 1 所述，所述视频数据流的所述 video_sequence 结构通过软件被修改，以包括所述图像类型的在所述比特级上的所述必要的标记，其特征进一步在于，两个独立的视频流 L 和 R 以同时形式存储，但以相同的 time_code 同步，并且被并行解码和显示。

7. 立体 3D 视频图像数字解码方法和系统，其中，所述视频数据流的所述 video_sequence 结构通过硬件被修改，其特征在于，通过硬件，实现 MPEG2 向后兼容性的 TDVision® 技术的属于所述

video_sequence 的所述结构、子结构和序列的特定使用，实际上，区分它为 2S 还是 3D 信号；激活双倍输出缓冲器（附加存储器）；激活并行解码选择器，激活差别解码选择器；执行图像解压缩处理，在相应的输出缓冲器中显示图像；启动对于可兼容的 MPEG2 的读取器是透明的 PICTURE_DATA3D()函数。

8. 立体 3D 视频图像数字解码方法和系统，其中，如权利要求 7 所述，所述视频数据流的所述 video_sequence 结构通过硬件被修改，其特征在于，属于所述 video_sequence 的所述结构、子结构和序列的特定使用以通过硬件实现所述 MPEG2 向后兼容性的 TDVision®技术：

a) Sequence_header

aspect_ratio_information

1001 在 TDVision®中 n/a

1010 在 TDVision®中 4:3

1011 在 TDVision®中 16:9

1100 在 TDVision®中 2.21:1

执行与 0111 的逻辑“与”，以获得与 2D 系统的向后兼容性，其中，指令被发送到规定立体重叠照相缓冲器（左或右）必须等于源的 DSP：

b) frame_rate_code

1001 TDVision®格式的 24,000/1001 (23.976)

1010 TDVision®格式的 24

1011 TDVision®格式的 25

1100 TDVision®格式的 30,000/1001 (29.97)

1101 TDVision®格式的 30

1110 TDVision®格式的 50

1111 TDVision®格式的 60,000/1001 (59.94)

执行与 0111 的逻辑“与”以获得与 2D 系统的向后兼容性，其中，指令被发送到规定立体重叠照相缓冲器（左或右）必须等于源的 DSP：

c) user_data()

sequence_scalable_extension

d) picture_header

extra_bit_picture

0=TDVision®

1=标准

e) picture_coding_extension

picture_structure

00= TDVision® 格式的图像

f) picture_temporal_scalable_extension()

9. 立体 3D 视频图像数字解码方法和系统，其中，如权利要求 7 所述，所述视频数据流的所述 video_sequence 结构通过硬件被修改，其特征在于，当识别出 PICTURE_DATA3D()结构时，其进行由解码器直接读取信息，但其将所述信息写入到也连接到位于电子显示装置中的视频输出以外的附加的视频输出的第二输出缓冲器中。

10. 立体 3D 视频图像数字解码方法和系统，其中，如权利要求 7 所述，所述视频数据流的所述 video_sequence 结构通过硬件被修改，其特征在于，如果信号为 TDVision®类型，则识别其是否为传输流、程序流或每秒 60 帧的左或右多路信号；当其为传输流时，其在现有 2D 编码器中具有向后兼容性；其中，指令被发送到规定立体重叠照相缓冲器（左或右）必须等于源的 DSP，所述 DSP 具有显示不具有 TDVision®的 3D 特性的视频的能力。

用于数字解码 3D 立体视频图像的方法和系统

技术领域

本发明涉及 3DVisor™ 装置中的立体视频图像显示，特别涉及借助于数字数据压缩系统的视频图像解码方法，该数字数据压缩系统允许使用标准化压缩技术存储三维信息。

背景技术

目前，使用数据压缩技术以减小图像或图像序列的显示中的比特消耗。标准化工作由国际标准化组织的专家组执行。目前，这些方法通常称为 JPEG（联合图像专家组），和 MPEG（运动图像专家组）。

这些技术的公共特性在于，图像块是借助于应用适用于块的变换来处理的，该变换通常称为离散余弦变换（DCT）。对形成的块进行量化处理，随后用变长码进行编码。

变长编码是可逆过程，其允许对已经用变长码进行编码的块进行精确重建。

数字视频信号显示包括以 30 至 75 Hz 频率连续显示或表示的一定数量的图像帧（30 至 96 fps）。每一个图像帧仍然是由像素阵列根据特定系统的显示分辨率形成的图像。例如，VHS 系统的显示分辨率为 320 列 480 行，NTSC 系统的显示分辨率为 720 列 486 行，并且，高清晰电视系统（HDTV）的显示分辨率为 1360 列 1020 行。关于低分辨率的数字化形式，320 列乘 480 行的 VHS 格式，两小时长度的电影可等于 100 GB 的数字视频信息。为了比较，传统压缩光盘的容量约为 0.6 GB，磁盘的容量为 1-2 GB，并且现有压缩光盘的容量为 8 GB 或者更大。

我们在电影院和电视屏幕上看到的所有图像都是根据以很高速率显示完整图像（静态图像，如照片）的原理。当它们以每秒 30 帧的速率（30 fps）快速并且连续的方式被显示出来时，由于人眼的视觉暂留（retention），我们将它们看成动画图像。

为了将以连续方式显示的图像编码并形成视频信号，每一个图像

都需要被分成行，其中，每行依次被分为图像元素或像素，每一个像素具有两个相关联的值，即，亮度和色度。亮度表示每一点上的光强，而色度作为定义的可用三个字节表示的颜色空间（RGB）的函数表示颜色。

图像被显示在屏幕上，在水平-垂直网格（raster）中，从上到下，从左到右循环地显示。显示的线数和频率可以由于格式不同而改变，例如，NTSC、PAL、或者 SECAM。

视频信号在被传输、接收和解码以显示在例如常规电视机或 3DVisor® 的显示装置中之后，可被数字化，以数字格式存储，该处理被称为模拟-数字视频信号编码-解码。

由定义，MPEG 有两种在系统流中交织视频和音频的不同方法。

传输流被用在错误概率较高的系统中，例如卫星系统，其易于受到干扰。每一个包的长度是 188 字节，以标识头开始，这使得可以识别间隙并修复可能的错误。不同的音频和视频程序能够在单一传输流上通过传输流同时传输；由于信息头，它们可为独立的并且被单独解码并集成到许多程序中。

程序流被用在错误概率较低的系统中，如被用在 DVD 播放中。在这种情况下，包具有可变长度和实际上比传输流中所使用的包大的尺寸。作为主要特性，程序流仅允许单一的程序内容。

即使当传输流和程序流处理不同的包时，也以相同的格式解码视频和音频格式。

按顺序，有三种应用到上面的包的压缩类型，例如，时间预测、压缩和空间压缩。

解码与冗长的数学处理相关，其目的是为了减小信息量。整帧的完整的图像被分为称为宏块的单元，每一个宏块由 16 像素×16 像素的矩阵组成，并且从上到下从左到右排列和命名。尽管在屏幕上有矩阵阵列，在信息流上发送的信息遵循特定的顺序，即，宏块以升序排列，也就是，macroblock0, macroblock 1 等。

一组连续宏块表示一片；如果宏块属于单一的行，一片中可以有任意数目的宏块。以与各宏块相同的方式，这些片从左到右从上到下命名。这些片必须覆盖整个图像，因为这是 MPEG2 压缩视频的格式，

编码图像不需要对每一个像素采样。一些 MPEG 标准需要处理严格的片结构，它们应该覆盖整个图像。

1999 年 10 月 5 日授予 Katata 等人的第 5,963,257 号 USP 保护一种平面视频图像解码装置，具有根据位置区域和图像格式、底层码、预测编码顶层码来分离解码数据的装置，由此获得编码数据的分层结构；解码器具有分离分层结构中的编码的数据以获得高质量图像的装置。

2001 年 9 月 18 日授予 Shen 等人的第 6,292,588 号 USP 保护一种用于编码从小区域重建和解码的预测平面图像的装置和方法，方法是，从小区域图像数据和所述图像的优化预测数据的总和生成重建的平面图像的数据。所述对图像数据流的预测解码装置包括一维 DCT 系数的变长码。2002 年 4 月 9 日授予 Boon 的第 6,370,276 号 USP 使用与上述相似的解码方法。

2002 年 9 月 24 日授予 Lazzaro 等人的第 6,456,432 号 USP 保护一种立体 3D 图像显示系统，其从两个角度拍摄图像，将它们显示在 CRT 上，并以场顺序方式多路输出图像，而不使观察者的双眼感到闪烁。

2003 年 12 月 2 日授予 Duruoz 等人的第 6,658,056 号 USP 保护一种数字视频解码器，其包括逻辑显示部件，逻辑显示部件响应于“相邻场”命令以获得输出存储器中的指定位置的数字视频场。数字视频显示系统配备有 MPEG2 视频解码器。图像被解码为存储器缓冲器，存储器缓冲器被优化，以保持补偿可变表并访问显示为数据场的固定存储器指针表。

2003 年 12 月 16 日授予 Boon 的第 6,665,445 号 USP 保护一种用于图像传输的数据结构、平面图像编码方法和平面图像解码方法。解码方法由两部分组成，第一部分编码图像格式信息数据流，第二部分为对图像数据流的像素值的解码处理，两部分能够根据平面图像信号编码进行交换。

2004 年 1 月 13 日授予 Moutin 等人的第 6,678,331 号 USP 保护一种 MPEG 解码器，其使用共用存储器。实际上，该电路包括微处理器、解码平面图像序列的 MPEG 解码器，和微处理器以及解码器的共用存储器。它还包括一种用于评估解码器延迟的电路，和用于确定微处理器或解码器的存储器优先级的控制电路。

2004年1月13日授予 Ferguson 的第 6,678,424 号 USP 保护一种实时人体视觉系统的行为模型；实际上，它处理两个不同通道中的二维图像信号，一个信号从另一个中产生。

发明内容

本发明的一个目的是提供一种立体 3D 视频图像数字解码系统和方法，包括软件的改变和硬件的改变。

本发明的另一个目的是提供一种解码方法，其中，标准 video_sequence 处理被应用于编码图像数据，即，variable_length_decoding(VLD)、inverse_scan；inverse_quantization、inverse_discrete_cosine_transform(IDCT)，和 motion_compensation。

本发明的又一个目的是改变用于解码视频格式的标识、2D 图像 MPEG2 向后兼容性、区分 TDVision® 类型图像、存储最近的图像缓冲器、应用信息解码、应用错误校正并将结果存储在各自的通道缓冲器中的软件信息。

本发明的又一个目的是提供一种用 video_sequence 处理标准形式的解码方法，用这种方式，当发现 TDVision® 类型图像时，最近完整图像的缓冲器被存储在左或右通道缓冲器中。

本发明的又一个目的是提供一种解码处理，其中，两个相关的（差别）视频信号可在相同的 video_sequence 内被发送，其中，应用信息解码并被存储为 B 类型帧。

本发明的又一个目的是提供一种解码处理，其中，在应用运动和颜色校正矢量时，对最近获得的图像应用错误校正。

本发明的又一个目的是通过软件对解码器进行编程，以同时接收和编码两个独立的程序流。

本发明的又一个目的是提供一种解码系统，其通过其中激活双输出缓冲器的硬件，解码 3D 图像信息。

本发明的又一个目标是提供一种 3D 图像信息的解码系统，其并行地并由差别激活图像解码选择器。

本发明的又一个目的是提供一种 3D 图像信息解码系统，其执行解压缩处理，并显示相应的输出缓冲器。

具体实施方式

硬件和软件算法的组合使得立体 3D 图像信息压缩成为可能，其通过发送具有立体重叠照相识别符的两个同步程序，作为对应于来自 3Dvision®照相机的左和右信号的两个独立的但具有相同 time_code 的视频信号被接收，由此促进编码-解码处理。同样，两个相关的视频信号可通过获取它们的差别被处理，差别被存储为具有图像类型识别符的“B”类型帧。由于使编码处理保持公开以促进技术发展，仅需要遵循该解码处理，即：对其中获得实质减少的编码数据应用变长解码，但是查询表必须被用于执行解码；应用反向扫描处理；应用反向量化处理，其中每一个数据被乘以一个标量；应用反余弦变换函数；应用错误校正或运动补偿步骤，并最终得到解码的图像。

从附图说明与所附的说明书，将会更好地理解本发明的新颖性特征及其结构和操作方法，其中，用相似的附图标记表示相似的部分和步骤。

图 1 表示本发明的主题涉及的技术示意图。该图显示出立体 3D 图像编码和解码系统和相应的方法。图像来自立体照相机 (32)，信息在 (31) 中编译并显示在任何适当的系统 (30) 或 (33) 中。信息在 (34) 中被编码，随后其可被传输到具有例如 (35) 的适当的在先解码过程的系统，其可为电缆系统 (36)、卫星系统 (37)、高清晰电视系统 (38) 或例如 TDVision®的 3DVisors® (39) 的立体视觉系统。

图 2 显示出其中绘出处理步骤的流程图。其目的是通过对现有 MPEG2 解码器做出修改和对解码处理 (2) 中的软件 (3) 和硬件 (4) 做出改变：解码器 (1) 必须与 MPEG2-4 兼容，从数字视频流中获取三维图像。

图 3 绘出必须被修改的结构和数据流的 video_sequence，以在比特级识别 TDVision®技术图像类型。

下面详细说明解码处理 (20) 的每一过程：

编码数据 (10) 为具有块信息的字节、宏块、场、帧、和 MPEG2 格式视频图像。

Variable_length_decoding (11) (VLC, 变长解码器) 是用较短码代替最经常使用的符号并用较长码代替较少发生的那些符号的压缩算

法。该信息的压缩版本占据较少空间，并可通过网络快速传输。然而，它并不是容易编辑的格式，并且需要使用查询表来解压缩。

例如，单词 BEETLE

字母	ASCII 码	VLC
B	01000010	0000 0010 10
E	0110 0101	11
L	0110 1100	0001 01
T	0111 0100	0100

因此，该单词的 ASCII 码是：0100 0010 0110 0101 0110 0101 0111 0100 0110 1100 0110 0101，在 VLC 中是：0000 0010 10 11 11 0100 00010 01 11。

注意到有很大的减少，然而，为了从 VLC 返回到单词“Beetle”，需要在查询表中搜索以解码该比特流，这是通过精确比较所读取的比特来进行的。

反向扫描（12）：信息必须由块分组，并且用 VLC 对信息进行编码，可获得线性流。块为 8×8 数据矩阵，所以有必要将线性信息转换为 8×8 方形矩阵。这是根据它是逐行图像还是隔行图像，通过在两种序列类型中都从上到下和从左到右的向下之字形方式进行的。

反向量化（13）：其为简单地将每一个数据值乘以一个因数。当被编码时，块中的多数数据被量化以去除人眼所不能感知到的信息，量化允许获得较大的 MPEG2 流转换，并且还需要在解码处理中执行逆处理（反向量化）。

反余弦变换（14）（IDCT, inverse_discrete_cosine_transform）：在每一个块中处理的数据属于频域，该反余弦变换允许变换回空间域的采样。一旦在 IDCT 中的数据已经被变换，能够获得像素、颜色和颜色校正。

运动补偿（15）允许校正在 MPEG 格式的解码过程之前生成的一些错误，运动补偿将之前的帧作为参考，并相对于像素计算运动矢量（它能计算高达四个矢量），并用它们创建新图像。该运动补偿被应用于 P 和 B 类型图像，其中，图像位置位于距离参考图象“t”时间处。除了运动补偿外，还应用错误校正，因为其不足以预测特定像素的位

置，但可能存在其颜色的改变。由此，获得解码图像（16）。

为了解码 P 或 B 类型图像，拍摄参考图像，运动矢量被代数相加，以计算下一图像，并最终应用错误校正数据，由此成功地生成解码图像。实际上，在 video_sequence 中，存在两个相关的视频信号，“ $R-L=\delta$ ”， δ 差别是被存储为具有 TDVision® 识别符的 B 类型立体重叠照相帧的差别，并且在解码时通过图像的差别创建。也就是， $R-\delta=L$ 和 $L-\delta=R$ ，左图像从与右图像的差别创建，相反，右图像从与左图像的差别创建。

以如下方法说明之前的处理：获取左或右信号，二者均存储在邻时缓冲器中，随后，计算左和右信号之间的差别，随后将其编码为 B 类型图像，存储在 video_sequence 中，稍后由所述图像的差别解码。

在解码处理中，可得出，由 VLC 过程输入的数据远小于由相同过程输出的数据。

MPEG 视频序列结构：这是 MPEG2 格式中使用的最大结构，并具有以下格式：

视频序列 (Video_Sequence)

序列头 (Sequence_header)

序列扩展 (Sequence_Extension)

用户数据 (0) 和扩展 (Extension_and_User_Data(0))

图像组头 (Group_of_Picture_Header)

用户数据 (1) 和扩展 (Extension_and_User_Data(1))

图像头 (Picture_Header)

编码图像扩展 (Picture_Coding_Extension)

用户数据 (2) 和扩展 (Extension_and_User_Data(2))

图像数据 (Picture_Data)

片 (Slice)

宏块 (Macroblock)

运动矢量 (Motion_Vectors)

编码块图案 (Coded_Block_Pattern)

块 (Block)

最终序列编码 (Sequence_end_Code)

这些结构组成视频序列。视频序列被应用于 MPEG 格式，以区分每个版本，其中，必须验证紧接在序列头后面是否存在序列扩展；如果序列扩展不接在序列头后面，则该流为 MPEG1 格式。

在视频序列的开始，sequence_header 和 sequence_extension 出现在 video_sequence 中。sequence_extension 的重复必须与第一次相同，并且 sequence_header 的“s”的重复与第一次出现相比变化很少，仅定义量化矩阵的部分应该变化。如果序列重复允许对视频流的随机访问，即，如果解码器想在视频流的中间开始播放，这是可以完成的，因为仅需要找到该时刻以前的 sequence_header 和 sequence_extension，以解码后续图像。这也发生在不能从开头开始的视频流的情况，例如在传输时间后打开的卫星解码器。

整个视频信号编码-解码处理由以下步骤组成：

数字化视频信号，其可以 NTSC、PAL 或 SECAM 格式完成。

将视频信号存储为数字格式

传输信号

在物理媒介（DVD、VCD、小型 DV）中记录数字视频流

接收信号

播放视频流

解码信号

播放信号

将由适当的 DSP 处理的存储器加倍并有可能设置高达 8 个输出缓冲器是重要的，这允许立体图像在例如 TDVision® 的 3DVisor® 的装置上的先前的和同时的显示。

实际上，两个通道必须在调用 DSP 的程序 API 时被初始化，例如，示例性情况为 Texas Instruments TMS320C62X DSP。

MPEG2VDEC_create (const IMPEG2VDEC_fxns*fxns, const MPEG2VDEC_Params*params)。

其中，IMPEG2VDEC_fxns y MPEG2VDEC_Params 为定义每一个视频通道的操作参数的指针结构，例如：

3DLhandle=MPEG2VDEC_create(fxns3DLEFT, Params3DLEFT)。

3DRhandle=MPEG2VDEC_create(fxns3DRIGHT, Params3DRIGHT)。

由此使两个视频通道被解码，并获得两个视频处理器，一个用于左-右立体通道。

需要双倍显示输出缓冲器，借助于软件，将定义两个缓冲器中的哪一个必须通过调用 AP 函数显示输出：

即，MPEG2VDEC_APPLY(3DRhandle, inputR1, inputR2, inputR3, 3doutright_pb, 3doutright_fb)。

MPEG2VDEC_APPLY(3DLhandle, inputL1, inputL2, inputL3, 3doutleft_pb, 3doutleft_fb)。

该相同的步骤可被实现以用于任何 DSP、微处理器或有类似功能的电子装置。

其中，3DLhandle 为指向由 DSP 的创建 (create) 函数返回的句柄的指针，input1 参数为 FUNC_DECODE_FRAME 或 FUNC_START_PARA 地址，input2 为指向外部输入缓冲器地址的指针，并且，input3 为外部输入缓冲器的大小。

3doutleft_pb 为参数缓冲器的地址，并且，3doutleft_fb 为将要存储解码图像的输出缓冲器的开始。

时间码和时间标签将被用于以顺序的、同步的方式输出到最后的装置。

将由 DSP 处理的存储器加倍并有可能设置高达 8 个输出缓冲器是重要的，这允许立体图像在例如 TDVision® 公司的 3DVisor® 的装置上的先前的和同时的显示。

软件和硬件处理的集成是通过称为 DSP 的装置执行的，DSP 执行多数硬件处理。这些 DSP 通过由制造商提供的 C 和汇编语言混合编程。每一个 DSP 有其自己的 API，由位于 DSP 中并被软件调用的功能列表或过程调用组成。

由该参考信息，做出用于 MPEG2 格式可兼容的 3D 图像解码的本申请。

实际上，在视频序列的开始，总出现序列头 (sequence_header) 和序列扩展。序列扩展的重复必须与第一个相同。相反，与第一次出现相比，序列头重复变化很少，仅定义量化矩阵的部分应该变化。

图 4 显示出用于 TDVision® 解码方法的编译软件格式 (40)，其中，

在 `sequence_header` (42) 中识别数字立体图像视频流的 `video_sequence` (41), 其可为相关的或独立的 (并行图像)。如果图像为 TDVision®, 则激活双倍缓冲器并识别 `aspect_ratio_information` 中的变化。对应于此处找到的图像的信息在 `user_data` (43) 中被读取。`sequence_scalable_extension` (44) 识别包含在其中的和基层以及增强层中的信息, 此处可定位 `video_sequence`, 定义 `scalable_mode` 和层识别符。`extra_bit_picture` (45) 识别 `picture_structure`, `picture_header`, 并且 `picture_coding_extension` (46) 读取“B”类型图像, 并且如果它为 TDVision® 类型图像, 则它解码第二缓冲器。`picture_temporal_scalable_extension()` (47), 在具有时间可扩展性的情况下, 被用于解码 B 类型图像。

即, 序列头 (`sequence_header`) 提供视频流上的较高的信息级, 为了明确说明, 还指出对应于每一个信息级的比特数, 最高有效比特位于序列扩展 (`Sequence_Extension`) 结构中, 它是由以下结构形成的:

Sequence_Header

字段	#bits	描述
<code>Sequence_Header_Code</code>	32	<code>Sequence_Header</code> 开始 0x00001B3
<code>Horizontal_Size_Value</code>	12	较低有效比特用于宽度*
<code>Vertical_Size_Value</code>	12	12 个较低有效比特用于高度
<u>Aspect Ratio Information</u> (纵横比信息)	4	图像方面
		0000 禁用
		0001 n/a TDVision®
		0010 4:3 TDVision®
		0011 16:9 TDVision®
		0100 2.21:1 TDVision®
		0111 将执行逻辑“与”以获得与 2D 系统的向后兼容性。
		0101...1111 保留

<u>Frame rate code</u> (帧速率码)	4	0000 禁用 0001 TDVision® 格式中的 24,000/1001(23.976) 0010 TDVision® 格式中的 24 0011 TDVision® 格式中的 25 0100 30,000/1001 (29,97) “ 0101 TDVision® 格式中的 30 0110 TDVision® 格式中的 50 0111 60,000/1001 (59.94) “ (将执行逻辑 “与” 以获得与 2D 系统的向后兼容性。) 1000 60 1111 保留
Bit_rate_value	18	Video_stream 比特率的 18 个较 低有效比特 (bit_rate= 400×bit_rate_value + bit_rate_extension << 18 = 最 高有效比特位于 sequence_extension 结构中。
Marker_bit	1	总为 1(防止 start_code_失效)。
Vbv_buffer_size_value	10	vbv_buffer_size 的 10 个较低有 效比特, 其确定视频缓冲检验 器 (VBV) 的大小, VBV 为一 种结构, 其用于确保数据流可 被用于解码有限大小的缓冲器 而不超过或在缓冲器中保留过 多自由空间。
Constrained_parameters_flag	1	总为 0, 不用在 MPEG2 中。
Load_intra_quantizer_matrix	1	指示内部编码的 (intra-coded) 的量化矩阵是否有效。
If(load_intra_quantizer_matrix) Intra_quantizer_matrix(64)	8×64	如果指示量化矩阵, 则这里必

		须指出，它为 8×64 的矩阵。
Load_non_intra_quantizer_matrix	1	如果
		load_non_intra_quantized_matrix
		x
If	8×64	如果先前的标记被激活，形成
load_non_intra_quantizer_matrix		量化矩阵的 8×64 数据被存储
Non_intra_quantizer_matrix(64)		在这里。

*最高有效比特位于 sequence_extension 结构内。

Picture_coding_extension

字段	bits#	描述
Extension_start_code	32	总为 0x000001B5
Extension_start_code_identifier	4	总为 1000
F_code(0)(0)	4	用于解码运动矢量；当它为 I 类型图像时，该数据由 1111 填充。
F_code(0)(1)	4	
F_code(1)(0)	4	解码运动矢量(B)中的信息，当它为(P)类型图像时，它必须被设置为 1111，因为没有向后运动。
F_code(1)(1)	4	解码运动矢量中的信息，当它为 P 类型图像时，它必须被设置为 1111，因为没有向后运动。
Intra_dc_precision	2	反向量化 DC 离散余弦变换系数的精度。
		00 8 比特精度
		01 9 比特精度
		10 10 比特精度
		11 11 比特精度
Picture_structure	2	指示图像被分成多场还是整帧。
		00 保留(TDVision®格式的图像)

		01 顶场
		10 底场
		11 逐帧图像
Top_field_first	1	0=先解码底场 1=先解码顶场
Frame_pred_frame_dct	1	
Concealment_motion_vectors	1	
Q_scale_type	1	
Intra_vic_format	1	
Alternate_scan	1	
Repeat_first_field	1	0=显示逐行帧 1=显示两个相同的逐行帧
Chrome_420_type	1	如果色度格式为 4:2:0, 则它必须 等于逐行帧 (progressive frame), 否则它必须等于零。
Progressive_frame	1	0=隔行的 1=逐行的
Composite_display_flag	1	警告初始编码的信息
V_axis	1	
Field_sequence	3	
Sub_carrier	1	
Burst_amplitude	7	
Sub_carrier_phase	8	
Next_start_code()		

Picture_temporal_scalable_extension()

在具有时间可扩展性的情况下, 存在两个空间分辨率流, 底层提供视频帧的较少索引版本, 而顶层可以被用于获取相同视频的多帧版本。低质量、低成本或免费解码器可使用时间可扩展性, 而相同付费下每秒可传输更多帧。

Picture_temporal_scalable_extension()

字段	bits#	描述
Extension_start_code_identifier	4	总为 1010
Reference_select_code	2	用于指示参考图像将被用于解码内部编码的 (intra_coded) 图像 对于 O 类型图像 00 增强最近的图像 01 按显示顺序的较低和最近的帧层 10 按禁止显示顺序的下一较低帧层 11 禁用 对于 B 类型图像 00 禁用 01 增强模式下的最近解码的图像 10 增强模式下的最近解码的图像 11 按显示顺序的底层中的最近图像
Forward_temporal_reference	10	时间参考
Marker_bit	1	
Backward_temporal_reference	10	时间参考
Next_star_code()		

Picture_spatial_scalable_extension()

在图像空间可扩展的情况下，增强层包含数据，其允许基层的更好的分辨率，使其可被重建。当增强层作为动作补偿的参考被用作基层的函数时，底层必须被提高并补偿，以获得增强层的更大分辨率。

Picture_spatial_scalable_extension()

字段	bits#	描述
Extension_start_code_identifier	4	总为 1001
Lower_layer_temporal_reference	10	参考较低层的时间图 像
Marker_bit	1	1
Lower_layer_horizontal_offset	15	水平补偿 (Offset)
Marker_bit	1	1
Lower_layer_vertical_offset	15	垂直补偿 (Offset)
Spatial_temporal_weight_code_table_index	2	预测细节
Lower_layer_progressive_frame	1	1=逐行的 0=隔行的
Lower_layer_desinterlaced_field_select	1	0=使用顶场 1=使用底场

Next_start_code()

Copyright_extension()

Extension_start_code_identifier	4	总为 010
Copyright_flag	1	如果它等于 1 则它使用版权 如果它为零(0)，不需要附加的版 权信息
Copyright_identifier	8	1=初始 0=复制
Original_or_copy	1	
Reserved	7	
Marker_bit	1	
Copyright_number_1	20	授予的版权号
Marker_bit	1	
Copyright_number_2	22	授予的版权号
Marker_bit	1	

Copyright_number_3 22 授予的版权号

Next_start_code()

Picture_data()

这是简单结构，它本身中不具有字段。

Slice()

包含相同的垂直位置中的一个或多个宏块的信息。

Slice_start_code	32
Slice_vertical_position_extension	3
Priority_breakpoint	7
Quantizer_scale_code	5
Intra_slice_flag	1
Intra_slice	1
Reserved_bits	7
Extra_bit_slice	1
Extra_information_slice	8
Extra_bit_slice	1

Macroblock()

Macroblock_modes()

Motion_vectors()

Motion_vector()

Coded_block_pattern()

Block()

EXTENSION_AND_USER_DATA(2)

图像可显示在以下设备上：

DVD（数字通用盘）

DTV（数字电视）

HDTV（高清晰电视）

CABLE（DVB 数字视频广播）

SATELLITE (DSS 数字卫星系统); 并且它是软件和硬件处理的集成。

图 5 的硬件 (50) 部分中的解码编译格式, 在 DSP 输入存储器中被复制, 同时, 允许对应于由立体 TDvision® 照相机拍摄的左-右立体现有信号的两个独立的或相关的视频信号的同时输入。在处理步骤中, 检测 video_sequence (51), 以交替左和右帧或并行发送它们, sequence_header 识别 (52)、识别图像类型 (53), 其传输到标准视频流 (54), 随后其经过错误校正处理 (55), 视频图像信息被发送到交替共享左通道 (57) 和右通道 (58) 并将信息分配到左通道 (57) 和右通道 (58) 的输出缓冲器 (56), 并且视频流信息在所述通道中被显示为 3D 或 2D 图像。

由于以同时形式存储 L (左) 和 R (右) 视频流作为两个独立的视频流, 但用相同的 time_code 同步, 因此, 它们稍后可被解码并在大容量输出缓冲器中并行播放。它们也可为相关的并通过差别解码。

关于硬件, 多数处理是由被称为 DSP (数字信号处理器) 的装置执行的。作为一个实例, 即, 可使用 Motorola 和 Texas Instruments (TMS320C62X) 模型。

这些 DSP 通过由所讨论的制造商提供的 C 和汇编语言的混合语言编程。每一个 DSP 具有其自己的 API, 由位于 DSP 中被软件调用的功能列表或过程调用组成。从该参考信息, 3D 图像被编码, 其可与 MPEG2 格式和它们自身的编码算法兼容。当信息被编码时, DSP 负责运行预测、比较、量化和 DCT 函数应用程序处理, 以形成 MPEG2 压缩视频流。

为了从数字视频流中获取三维图像, 已经通过在解码处理的不同部分中的软件和硬件改变, 对现有 MPEG2 解码器进行了一些修改。视频数据流的结构和 video_sequence 必须被修改以包括必要的标记以在比特级上识别 TDVision® 技术图像类型。

在接下来的解码步骤中做出修改。

软件:

一 视频格式识别

- 在不是 TDVision® 视频的情况下，应用逻辑“与”用于实现 MPEG2 向后可兼容性。
- 扫描 video_sequence，以标准方式（现有技术）图像解码。
- 在 TDVision® 类型图像的情况下：
 - _ 区分它们为相关的还是独立的视频信号
 - _ 将最近的完整图像缓冲器存储到左或右通道缓冲器中。
 - _ 应用 B 类型帧信息解码。
 - _ 通过应用运动和颜色校正矢量对最近获得的图像应用错误校正。
 - _ 在它们各自的通道缓冲器中存储结果。
 - _ 继续视频序列读取。

硬件：

- 当通过硬件解码信息时：
 - _ 区分图像是 2D 还是 3D
 - _ 激活双倍输出缓冲器（存储器增加）。
 - _ 激活差别解码选择器。
 - _ 激活并行解码选择器。
 - _ 执行解压缩处理。
 - _ 图像被显示在其相应的输出缓冲器中。

以下结构、子结构和序列，将以特定方式被使用；它们属于 MPEG2 向后兼容 TDVision® 技术的硬件实现的 video_sequence 结构。

实际上：

Sequence_header

Aspect_ratio_information

1001 n/a 在 TDVision® 中

1010 4:3 在 TDVision® 中

1011 16:9 在 TDVision® 中

1100 2.21:1 在 TDVision® 中

与 0111 执行逻辑“与”以获得与 2D 系统的向后兼容性，当这发

生时，指令被发送到立体重叠照相（左或右）的缓冲器必须等于源的 DSP，所以所有解码的图像将被发送到两个输出缓冲器，以允许任何装置中的图像显示。

Frame_rate_code

1001 TDVision®格式中的 24,000/101 (23.976)

1010 TDVision®格式中的 24.

1011 TDVision®格式中的 25.

1100 TDVision®格式中的 30,000/1001 (29.97).

1101 TDVision®格式中的 30.

1110 TDVision®格式中的 50.

1111 TDVision®格式中的 60,000/1001 (59.94)

与 0111 执行逻辑“与”以获得与 2D 系统的向后兼容性。

User_data()

Sequence_scalable_extension

Picture_header

Extra_bit_picture

0=TDVision®

1=标准

Picture_coding_extension

Picture-structure

00= TDVision®格式的图像

Picture_temporal_scalable_extension()

在编码信息时，使用 DSP，DSP 负责执行预测、比较和量化处理，应用 DCT 以形成 MPEG2 压缩视频流，并区分 2D 或 3D 图像。

两个视频信号以独立形式被编码但具有相同的 time_code，对应于来自 3DVision®照相机的左信号和右信号的信号，同时发送具有 TDVision®立体重叠照相识别符的两个程序。该类型的解码被称为“通过并行图像”，因为同时存储左和右（L 和 R）视频流两者作为两个独立的视频流，但是为 time_code 同步的。随后，它们将并行地被解码和

播放。只有解码软件必须被解码，传输流的编码和压缩算法将与当前算法相同。

解码器中的软件修改。

在解码器中，利用程序 API，如在使用 TMS320C62X 系列 Texas Instruments DSP 的实例中，两个程序流必须被同时编程，或两个相关的视频信号，即，从被存储为具有识别符的 B 类型帧的二者之间的差别创建。

DSP 的编程算法和方法。

— 在启动 DSP 时，创建两个处理通道（当调用 API 时，主和辅助缓冲器或者左和右缓冲器）。

— 得到每一个通道的 RAM 存储器指针（存储分配图中的 RAM 地址）

— 当获得 TDVision® 类型视频序列时，

它被认为是 B 类型

图像被实时解码

变化或差别被应用到补偿缓冲器

结果被存储在辅助缓冲器中。

在与 video_sequence 数据流中的软件相关的软件中，需要实现两个选项：

1.- 一个仅修改软件并使用 user_data() 部分，以存储允许重新生成立体信号的错误校正。

2.- 另一个通过硬件启动对于 MPEG2 可兼容读取器透明的 PICTURE_DATA3D() 函数，并且其可通过 TDVision® 可兼容 DSP 解码。

在 MPEG2 解码器检测到 user_data() 码时，其将搜索 3DVISION_START_IDENTIFIER = 0X000ABCD 32 比特识别符，该码非常高并且难于再现码，或者其不表示数据。随后，将考虑要读取的 3D 块长度，该长度为 32 比特“n”数据。当在 USER_DATA() 中检测到该信息时，将调用特定解码函数，特定解码函数随后与输出缓冲器比较，并在 video_sequence 的当前读取偏移处应用，n 字节作为对 B 类型帧的典型校正。该校正的输出被发送到直接与位于电子显示装置中的视频输出以外的附加的视频输出相关的其它输出地址。

如果识别出 PICTURE_DATA3D()结构，则进行由解码器直接读取信息；但其将信息写入到也连接到位于电子显示装置中的视频输出以外的附加的视频输出的第二输出缓冲器中。

在程序流的情况下，两个信号（左和右）由 time_code 同步，两个信号将被具有足够的同时多视频通道解码能力的 MPEG 解码器并行地解码，或它们可以在相同 video_sequence 内发送两个相关的视频信号，例如，“R-L=delta”，其中，delta 为存储为具有立体重叠照相 TDVision® 识别符的“B”类型帧的差别，并且，其可在由图像的差别解码时被重建，即，“R-delta=L”或“L-delta=R”，如前述 Texas Instruments DSP 的情况，其可作为例证但不是限制性的实例。

包含单个视频序列的视频也被实现；但以每秒 60 帧速率交替左和右帧（每一个 30 帧），并且在解码后将视频缓冲器图像置于相应的左或右通道中。

如果信号为 TDVision® 类型，则解码器还具有通过硬件检测的能力，如果是这种情况，则，可识别信号是传输流、程序流还是每秒 60 帧的左-右多路信号。

在传输流的情况下，在现有解码器中具有向后兼容性系统，该向后兼容性系统具有显示没有 3d 特性而是仅为 2D 的相同视频的能力，在这种情况下，DSP 被禁止在任何 TDVision® 或现有技术装置中显示图像。

在程序流的情况下，使用未修改的编码器，例如那些目前用在卫星传输系统中的编码器；但是接收器和解码器具有 TDVision® 标记识别系统，因此使第二视频缓冲器形成左-右对。

最后，在多路视频的情况下，启动具有两个视频缓冲器（左-右）的 MPEG 解码器，识别适当的帧并分离每秒 30 帧的每个信号，因此，由于视频流是连续的，并且由于人眼的特性滞留波，多路作用不明显，可提供无闪烁图像。

已经示例并说明了本发明的具体实施例，在不脱离本发明的范围内做出一些修改或改变，对于本领域的那些技术人员是显而易见的。所有这些修改和改变意在使所附权利要求覆盖，使所有改变和修改落入本发明的范围内。

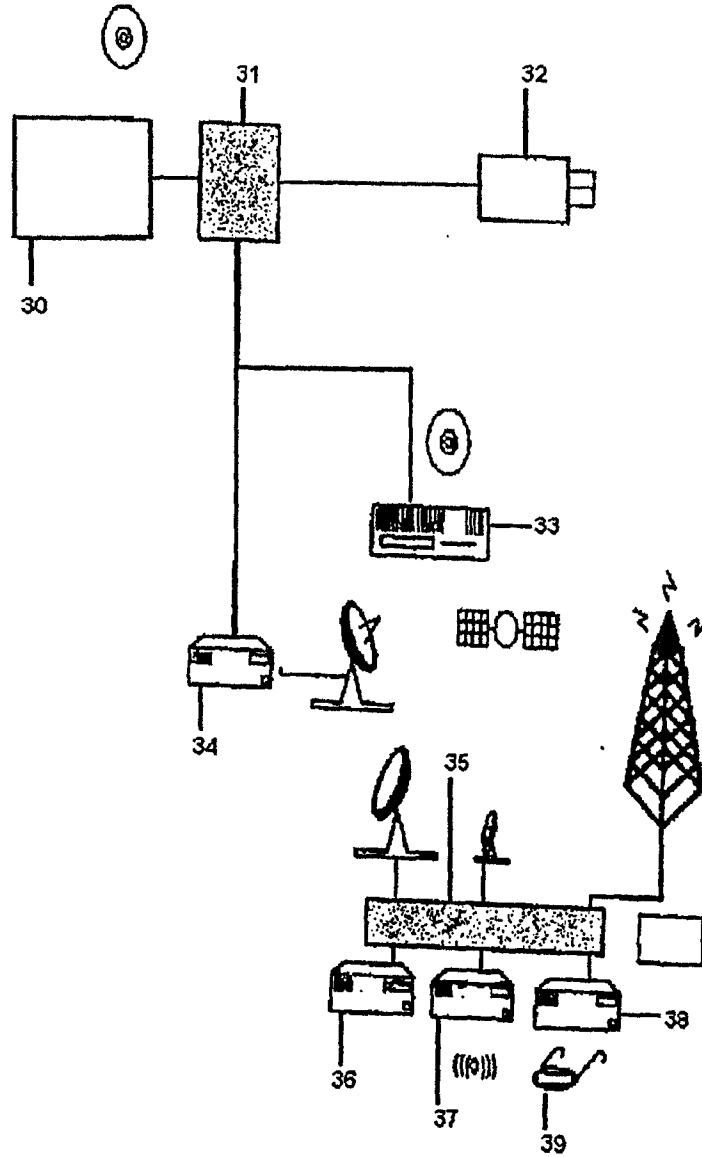


图1

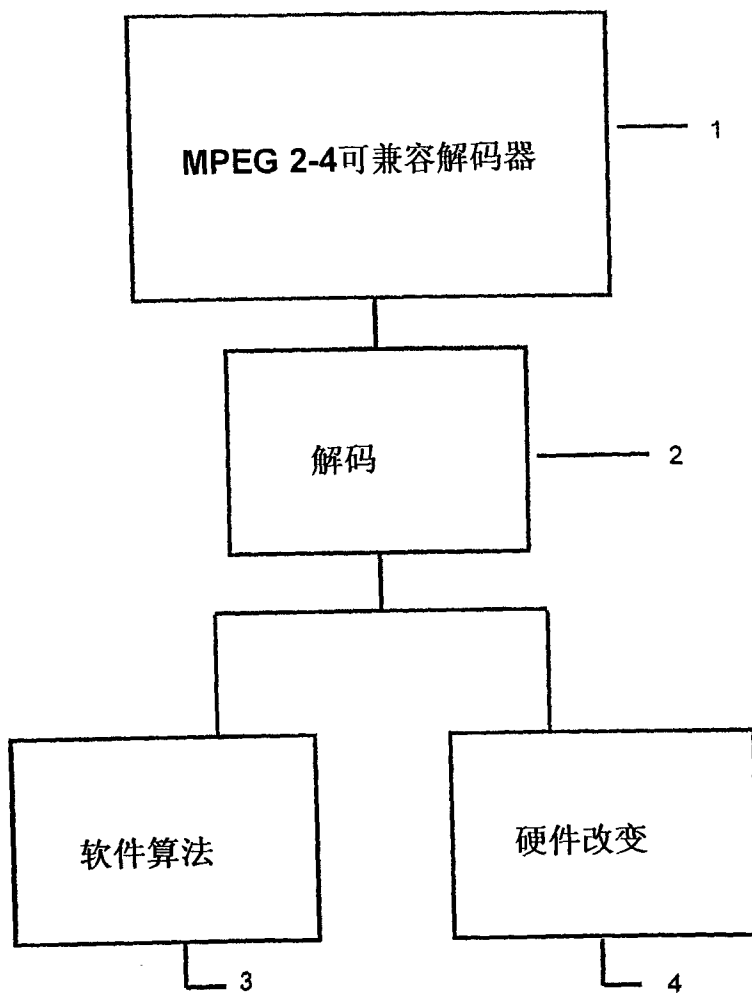


图2

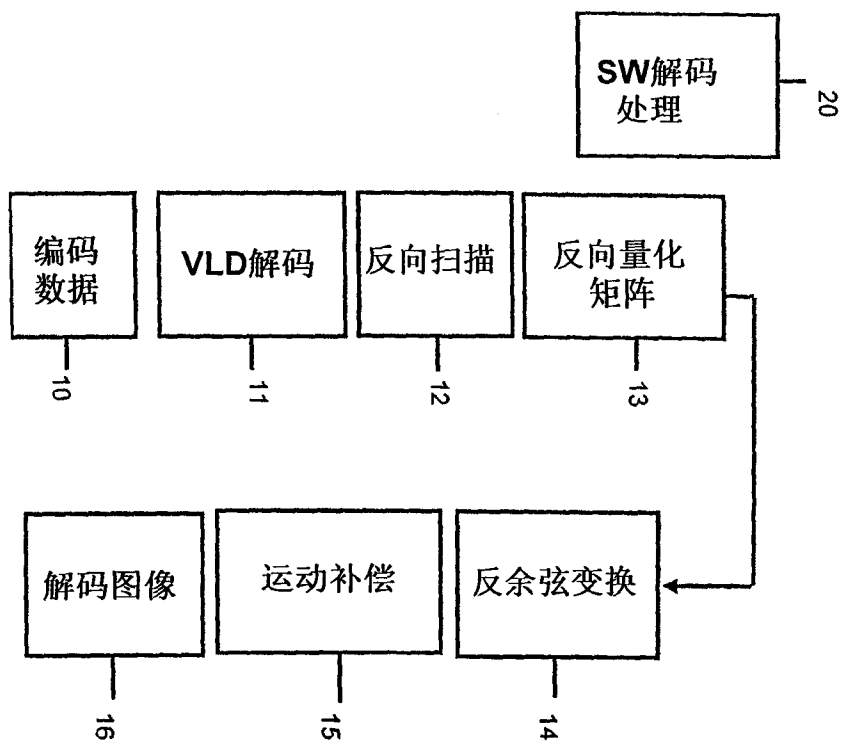


图3

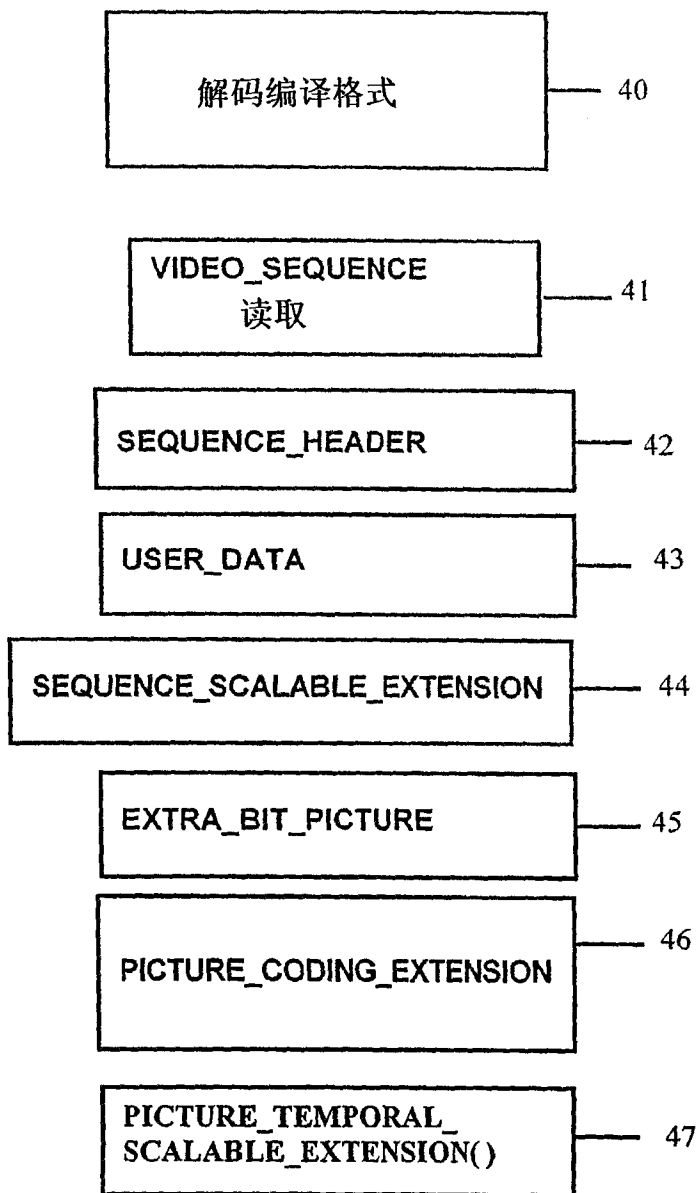


图4

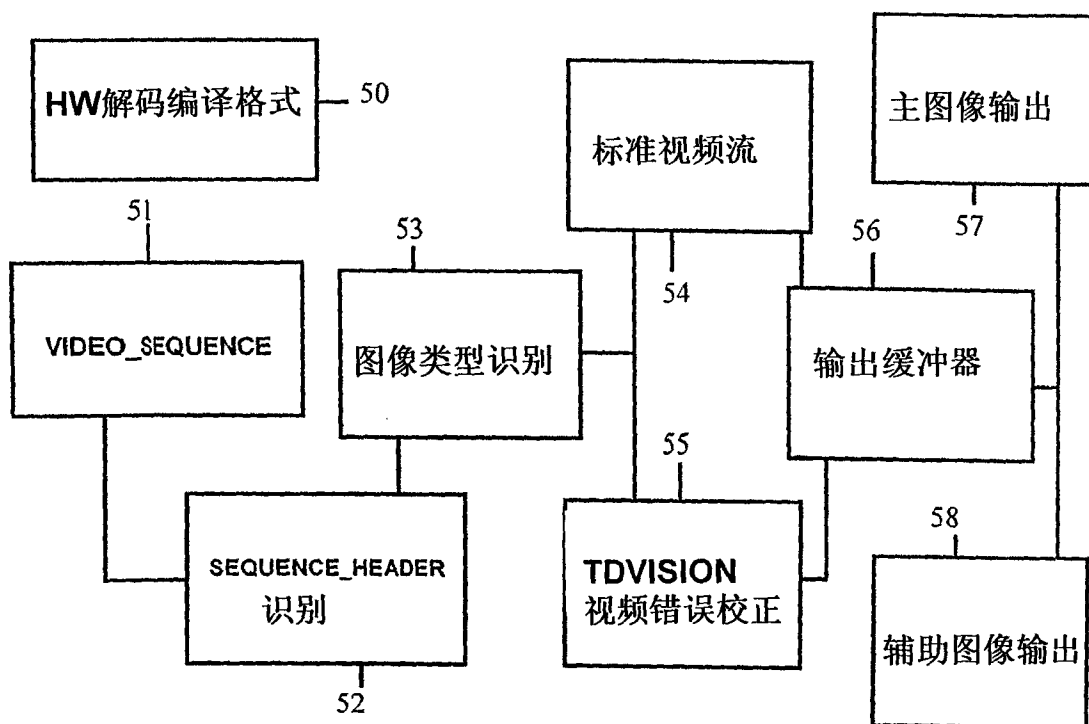


图5