



## (12) 发明专利

(10) 授权公告号 CN 1678073 B

(45) 授权公告日 2011.01.19

(21) 申请号 200510058897.4

(22) 申请日 2005.03.30

(30) 优先权数据

04251920.7 2004.03.31 EP

(73) 专利权人 三菱电机株式会社

地址 日本东京

(72) 发明人 安德鲁·塞克

(74) 专利代理机构 北京三友知识产权代理有限公司

11127

代理人 黄纶伟

(51) Int. Cl.

H04N 7/26 (2006.01)

(56) 对比文件

CN 1132983 A, 1996.10.09, 全文.

GB 2368220 A, 2002.04.24, 全文.

US 6498810 B1, 2002.12.24, 摘要, 说明书第3栏第28行至第6栏第6行, 图4, 5.

Taubman, D. Highly scalable video compression with scalable motion coding.

Image Processing, 2003. ICIP 2003.

Proceedings. 2003 International Conference on 2. 2003, 2III-273-III-276 摘要, 3. 运动值量化, 4. 可伸缩运动编码, 5. 实验结果.

审查员 戚颖

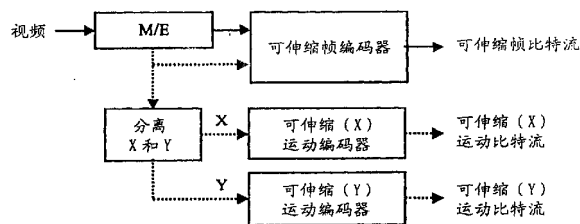
权利要求书 1 页 说明书 7 页 附图 1 页

### (54) 发明名称

用于可伸缩视频编码的方向自适应可伸缩运动参数编码

### (57) 摘要

一种运动图像数据编码方法, 特别地, 利用运动补偿 3-D 子带编码的运动图像数据编码方法, 其中相对于由运动补偿得到的运动矢量的第二分量单独地或者独立地对该运动矢量的第一分量进行可伸缩编码, 包括对于第一和第二运动矢量分量进行单独的比特率分配。



1. 一种利用运动补偿对运动图像数据进行编码的方法,其中相对于运动矢量的水平分量单独地或独立地对运动矢量的垂直分量进行可伸缩编码,并且包括对运动矢量的垂直和水平分量进行单独的比特率分配,并且

其中,所述比特率分配涉及 $D_{x,M} \approx \Psi_{R,S}^1 D_M^1 + \Psi_{R,S}^2 D_M^2$ 形式的运动诱导视频失真模型,其中 $\Psi_{R,S}^1$ 和 $D_M^1$ 是指垂直运动矢量分量, $\Psi_{R,S}^2$ 和 $D_M^2$ 是指水平运动矢量分量,

或者,所述比特率分配涉及组合了帧采样失真和运动分量失真的 $D_x \approx D_s + \Psi_{R,S}^1 D_M^1 + \Psi_{R,S}^2 D_M^2$ 形式的总重构视频失真模型。

2. 根据权利要求1所述的方法,其中利用运动补偿 3-D 子带编码对运动图像数据进行编码。

3. 根据权利要求1所述的方法,其中运动矢量的编码是可伸缩编码。

4. 根据权利要求1所述的方法,其中比特率分配考虑了图像序列中的垂直和水平内容的差异。

5. 根据权利要求1所述的方法,其中比特率分配考虑了图像序列对垂直和水平运动的敏感度的差异。

6. 根据权利要求1所述的方法,其中比特率分配涉及反应分别由水平和垂直运动误差引起的重构误差的水平和垂直运动敏感度或伸缩系数。

7. 根据权利要求1所述的方法,包括对于多个重构比特率和 / 或空间分辨率确定比特率分配。

8. 根据权利要求7所述的方法,包括导出并存储码率表,该码率表包括对于多个重构比特率和 / 或空间分辨率的运动矢量分量的比特率分配信息。

9. 根据权利要求1所述的方法,该方法包括对所述比特率分配使用优化技术。

10. 根据权利要求9所述的方法,包括识别分别分配给水平和垂直运动矢量分量的比特,并利用解码的水平和垂直运动矢量分量重构图像序列。

11. 一种利用运动补偿对运动图像数据进行编码的装置,该装置包括:

用于相对于运动矢量的水平分量单独地或独立地对运动矢量的垂直分量进行编码的装置;以及

用于对运动矢量的垂直和水平分量进行单独的比特率分配的装置,所述比特率分配涉及 $D_{x,M} \approx \Psi_{R,S}^1 D_M^1 + \Psi_{R,S}^2 D_M^2$ 形式的运动诱导视频失真模型,其中 $\Psi_{R,S}^1$ 和 $D_M^1$ 是指垂直运动矢量分量, $\Psi_{R,S}^2$ 和 $D_M^2$ 是指水平运动矢量分量,

或者,所述比特率分配涉及组合了帧采样失真和运动分量失真的 $D_x \approx D_s + \Psi_{R,S}^1 D_M^1 + \Psi_{R,S}^2 D_M^2$ 形式的总重构视频失真模型。

12. 根据权利要求11所述的装置,其中,所述用于进行编码的装置包括用于对垂直运动矢量分量进行编码的可伸缩运动编码器,以及用于对水平运动矢量分量进行编码的可伸缩运动编码器。

13. 根据权利要求11所述的装置,其中,所述用于进行单独的比特率分配的装置被设置为在编码过程中向垂直和水平运动矢量分量分配比特。

14. 根据权利要求13所述的装置,该装置包括用于在编码过程中向垂直和水平运动矢量分量分配比特的码率表。

## 用于可伸缩视频编码的方向自适应可伸缩运动参数编码

### 技术领域

[0001] 本发明涉及用于对图像序列形式的运动图像数据进行编码的方法和装置。本发明特别涉及包括空间和时间滤波以及运动补偿的 3D 子带编码,以及运动矢量的编码。

### 背景技术

[0002] 在互联网等的异构通信网络中,高效的视频通信必须提供宽范围的传输约束和视频显示参数。在同一网络上的不同用户之间信道带宽经常会变化几个数量级。另外,随着网络互联的快速发展趋势,移动电话、手持个人数字助理和台式工作站等的各具有不同显示分辨率和处理能力的设备可能会访问相同的数字媒体内容。

[0003] 可伸缩 (scalable) 视频编码旨在通过压缩原始视频内容,使得同时支持多种不同比特率和显示分辨率的有效重构,从而应对不同的视频通信网络和终端用户的兴趣。比特率可伸缩性是指在不损失压缩效率的情况下在精细划分的比特率范围上重构压缩视频的能力。这使得多个用户可以访问一个压缩比特流,每个用户利用他/她所有的可用带宽。如果没有比特率可伸缩性,则需要在网络上提供同一视频数据的多个版本,这显著地增加了存储和传输的负担。可伸缩性的其它重要形式包括空间分辨率和帧频 (时间分辨率) 可伸缩性。这些使得可以在不同的显示分辨率下有效地重构压缩视频,从而适应所有类型的终端用户设备的性能。在 D. Taubman, "Successive refinement of video: fundamental issues, past efforts and new directions", Int. Sym. Visual Comm. ImageProc. July 2003 中,介绍了可伸缩视频压缩领域中当前的推动因素,过去的实践以及未来的趋势。

[0004] 近年来,随着几个重要的发现,可伸缩视频编码的研究得到了越来越多的关注。具体而言,一种新的构建高效前馈压缩系统的架构与先前的方案相比具有显著的优势。实际上,可伸缩视频编码器终于开始实现可以与现有的不可伸缩编码方法相比的压缩性能,然而具有所有的上述这些理想的可伸缩特征。这些新的方案称为“运动补偿提升 (motion-compensated lifting)”方案,最初是由 Secker 和 Taubman (A. Secker and D. Taubman, "Lifting-based invertible motionadaptive transform (LIMAT) framework for highly scalable video compression", IEEE Trans. Image Proc. Dec 2003) 和 Pesquet-Popescu 等 (B. Pesquet-Popescu and V. Bottreau, "Three-dimensional lifting schemes for motion compensated video compression", IEEE Int. Conf. Acoustics, Speech Signal Proc., ppl 793-1796, Dec 2001) 同时提出的。

[0005] 运动补偿提升方案使得可以对视频数据应用高效的基于小波的时间变换,而不会牺牲压缩系统逆转能力。小波时间变换把原始视频帧变换为时间“子带 (subband)”帧的集合。可逆变换特别重要,因为它们使得在有足够带宽的情况下可以完美地重构视频。使用与可伸缩图像压缩中使用的基本上相同的技术对时间子带帧进行处理。这些目前已经基本上成熟的技术 (在最近的 JPEG2000 图像压缩标准中达到顶点) 包括 J. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients", IEEE Trans.

Signal Proc., vol 41, pp 3445-3462, Dec. 1993, D. Taubman and A. Zakhor, "Multi-rate 3-d subbandcoding of video", IEEE Trans. Image Proc., vol. 3, pp. 572-588, Sept 1994, A. Said and W. Pearlman, "A new, fast and efficient image codec based on set partitioning in hierarchical trees", IEEE Trans. Circ. Sys. Video Tech., pp. 243-250, June 1996, and D. Taubman, E. Ordentlich, M. Weinberger and G. Seroussi, "Embedded Block Coding in JPEG2000", Signal Processing-Image Communication, vol. 17, No. 1, pp. 49-72, Jan 2002 中所介绍的。还可参照本申请人的未授权申请 EP03255624.3 (P047), 在此以引用的方式并入这些内容。

[0006] 运动补偿提升变换的高压缩性能的关键在于其非常有效地开发运动的能力, 以及其适应于任何运动模型的能力。在文献中提出了大量的运动模型, 任何一种模型都可以引入到提升变换架构中。提出了多种方法以对使用参数化运动模型而得到的边信息 (side information) 进行表达和编码。但是, 通常边信息的量很大, 并且通常进行无损的编码, 这会显著地降低完全压缩系统的比特率可伸缩性。

[0007] 为了在很宽范围的比特率 (从几千比特 / 秒 (kbps) 到几兆比特 / 秒 (Mbps)) 上实现比特率可伸缩性, 运动信息的表达精度也必须是可伸缩的。如果没有运动可伸缩性, 则在低比特率下运动参数的编码成本会消耗过多的可用带宽。相反, 在高比特率下无法以足够的精度表达运动从而实现最大的编码增益。还需注意, 运动信息处理精度的伸缩能力是时间可伸缩性的一个自然延伸。这是因为细化重构视频序列的时间信息不仅要涉及到细化时间采样率, 还要涉及到利用运动自适应时间合成滤波器阵对这些时间采样进行插值的精度。

[0008] Secker 和 Taubman 最近在将于 IEEE Trans. Image Proc 中发表的 A. Secker and D. Taubman, "Highly scalable video compression with scalable motion coding" 中讨论了可伸缩运动编码, 也在作者的网站 [www.ee.unsw.edu.au/~taubman/](http://www.ee.unsw.edu.au/~taubman/) 中公布。在这个论文中他们提供了一种用于压缩和联合伸缩运动参数和视频采样的新颖的架构。他们的方法涉及利用与时间子带帧编码中使用的类似的可伸缩图像编码技术对与运动补偿提升变换相关的运动参数进行压缩。

[0009] Secker 和 Taubman 的工作涉及两个主要的贡献。首先, 他们描述了运动信息的可伸缩压缩方法, 其次, 他们提供了一种在视频帧编码花费的比特和运动参数编码花费的比特之间达到最佳平衡的架构。部分地, 可伸缩运动编码方法涉及按照与传统的可伸缩图像编码系统中处理标量图像采样的相同方式处理各个运动矢量分量。运动信息通常包括二维矢量的二维数组 (对应于视频帧之间的垂直和水平位移)。通过提取垂直和水平运动分量, 并将它们排列成二维标量场, 可以把它们压缩为标量图像。虽然对标量运动分量场进行空间小波变换, 但是把所得到的变换运动分量重新组合成矢量, 并联合地进行嵌入式量化和编码。这使得嵌入式编码级可以利用变换后的运动矢量分量之间的冗余。不光 Secker 和 Taubman 的可伸缩运动编码方案受人关注, 他们的在运动和视频采样比特率之间达到最佳平衡的方法也受人关注。和现有的可伸缩视频编码方案 (会涉及到产生可伸缩视频采样比特流) 不一样, 除了不可伸缩的运动参数比特流之外, Secker 和 Taubman 的方法产生两个可伸缩比特流; 一个对应于视频采样, 另一个对应于运动参数, 如图 1 所示。

[0010] 使用原始运动参数来生成可伸缩视频采样比特流。压缩之后伸缩运动信息表示使

用与压缩时不同的运动参数来进行重构。这个不一致导致了附加的重构视频失真。但是，这个附加失真可以量化，并且可以与由伸缩视频采样比特流导致的失真达到平衡，从而可以求得运动和采样比特率的最优组合。

[0011] 在上述 A. Secker 和 D. Taubman 的“Highly scalable videocompression with scaiable motion coding”中，作者描述了尽管运动误差和所导致的视频失真之间有着复杂的相互关系，但是也可以利用线性方法对该行为进行近似的建模。这个重要的发现证明了可以独立地构建可伸缩运动和视频比特流，因为可以在压缩视频帧之后确定运动和采样比特流的最优组合。根据 Secker 和 Taubman，由重构视频序列中的运动误差导致的总方差  $D^{(M)}$  可以由以下线性模型表示：

$$[0012] \quad D^{(M)} \approx \Psi_{R,S} D_M \quad (1)$$

[0013] 其中  $DM$  表示运动矢量中由于压缩后伸缩导致的均方差。伸缩系数  $\Psi_{R,S}$  取决于视频信号重构的空间分辨率  $S$ ，并取决于精度，或视频采样重构的比特率  $R$ 。

[0014] 运动信息和采样数据之间的最优码率分配需要知道与子带帧的可伸缩编码期间生成的嵌入式表示的前  $L^{(S)}$  个比特相关的重构视频采样失真  $D^{(S)}$ 。另外，码率分配也需要知道由于把运动参数比特流截为长度  $L^{(M)}$  而导致的重构视频失真  $D^{(M)}$ 。根据 Lagrange 乘子法，

对于总长度  $L^{\max}$ ，当对于失真 - 长度斜率  $\lambda > 0$  有  $\frac{-\Delta D^{(S)}}{\Delta L^{(S)}} \geq \lambda$  且  $\frac{-\Delta D^{(M)}}{\Delta L^{(M)}} \geq \lambda$ ，且  $L^{(S)} + L^{(M)}$  尽

可能大但是不超过  $L^{\max}$  时，实现了运动和采样比特之间的最优分配。这里， $\Delta D^{(S)} / \Delta L^{(S)}$  和  $\Delta D^{(M)} / \Delta L^{(M)}$  是采样和运动比特流截取点处的失真 - 长度斜率的离散近似。实际上，通常知道  $D^{(S)}$  就够了， $L^{(S)}$ 、 $D^{(M)}$  和  $L^{(M)}$  只用于一些有限的可能的比特流截取点，以对于任意  $L^{\max}$  得到近乎最优的码率分配。

[0015] 根据公式 (1)，只要  $\Psi_{R,S}$  在  $L^{(M)}$  的小幅变化之下相对恒定，可以根据  $\frac{-\Delta D^{(S)}}{\Delta L^{(S)}} \geq \lambda$

和  $\Psi_{R,S} \frac{-\Delta D_M}{\Delta L^{(M)}} \geq \lambda$  来等价地执行码率分配。根据上述的 Secker 和 Taubman，通常是这种情况，

从而编码运动数据的码率失真最优性基本上与采样数据无关，并且可以与可伸缩采样比特流无关地构建可伸缩运动比特流。可以在压缩之后根据运动敏感系数  $\Psi_{R,S}$  求得运动和采样数据之间的最优码率分配。

[0016] 尽管这个码率失真优化模型可以应用于任何可伸缩视频编码方法，但 JPEG2000 采用的 EBCOT 算法提供了用于运动和采样比特流的编码和联合伸缩的优异的构架。在 D. Taubman, E. Ordentlich, M. Weinberger and G. Seroussi, “Embedded Block Coding in JPEG2000”, Signal Processing-Image Communication, vol 17, no 1 pp. 49-72, Jan 2002 中可以找到 EBCOT 算法的完整论述。EBCOT 算法产生组织为嵌入式“质量层”的比特流。任何层边界处的比特流截取都能产生满足上述码率失真优化目标的重构信号。涉及部分质量层的进一步重构减少重构失真，但不一定是码率失真最优形式。只要使用足够多的质量层，这个次优性通常并不要紧。

[0017] 目前的对运动参数和视频数据进行联合伸缩的方法仅考虑运动矢量失真的幅值，而没有考虑其方向。但是，视频序列通常会表现出各向异性功率谱，从而垂直和水平运动误差的效果会显著不同。当发生这种情况时，在现有的方案中垂直和水平运动矢量分量之间

的比特分配是次优的。解决这个问题可以带来更大的压缩效率,从而减小与可伸缩运动信息相关的性能损失。

## 发明内容

[0018] 本发明的原理是通过对各个运动矢量分量单独地执行码率分配,从而改善完全视频编码器的码率失真优化。本质上,这涉及对重构视频数据最为敏感的运动分量花费更多的比特。例如,对于在垂直方向上包含绝对高的频率能量的视频数据,对于垂直运动分量的编码花费更多的比特,而对于水平运动矢量分量花费更少的比特。相反,当视频序列包含绝对多数的水平纹理信息时,把大部分的运动比特花费在水平运动矢量分量的编码上,从而对于水平运动误差更为敏感。

[0019] 本发明在于对现有技术的运动诱导视频失真模型的改进。该改进模型引入了各个运动矢量分量 MSE 的各个项,而不仅仅是对应于各个运动矢量幅值 MSE 的一个项。该改进模型表达为:

$$[0020] \quad D_{x,M} \approx \Psi_{R,S}^1 D_M^1 + \Psi_{R,S}^2 D_M^2$$

[0021] 其中  $\Psi_{R,S}^1$  和  $D_M^1$  指垂直运动矢量分量,  $\Psi_{R,S}^2$  和  $D_M^2$  指水平运动矢量分量。对于不相干的运动和采样误差,可以使用下面的添加失真模型来量化总重构视频失真,作为单独的运动分量失真和帧采样失真之和。

$$[0022] \quad D_x \approx D_S + \Psi_{R,S}^1 D_M^1 + \Psi_{R,S}^2 D_M^2$$

[0023] 可以自然地对现有的运动信息编码和码率分配方法进行扩展以便于应用该改进模型。下面对这些扩展进行说明。

[0024] 总的来说,本发明的一个方面涉及一种利用运动补偿来编码运动图像数据的方法,该方法包括单独地考虑水平和垂直运动矢量分量(例如在重构/重构误差中)的影响。这可以通过单独地对水平和垂直运动矢量分量进行编码来实现,例如,优先编码对重构图像/帧的质量有更重要贡献的分量。该优先编码可以涉及平移或伸缩,例如比特平面中的比特平面平移或者部分比特平面编码。该优先编码可以基于比特率分配,即对于更重要的运动矢量分量分配更多的比特,例如使用优化技术,例如对于不同的比特率和/或空间分辨率使重构误差最小。在运动矢量的可伸缩编码,特别是与 3-D 子带编码相关的环境中,本发明特别有用。

[0025] 根据本发明的另一个方面,提供了一种运动图像数据编码的方法,特别是运动补偿 3-D 子带编码的方法,其中单独地,或者与运动矢量的第二分量无关地对由运动补偿得到的运动矢量的第一分量进行可伸缩编码,该方法包括对于运动矢量的第一和第二分量的单独的比特率分配。运动矢量是由运动预测技术得到的。

[0026] 在所附的权利要求中阐述了本发明的这些和其它方面。

## 附图说明

[0027] 下面参照附图对本发明的实施例进行说明,其中:

[0028] 图 1 是现有技术的编码系统的框图;

[0029] 图 2 是根据本发明一个实施例的编码系统的框图。

## 具体实施方式

[0030] 本发明和现有技术之间的主要差别在于分别地控制垂直和水平运动矢量分量中的失真。是这样实现的：首先将运动矢量场分解为与各个图像维相对应的标量场，分别地进行编码，从而生成两个可伸缩运动分量比特流，如图 2 所示。

[0031] 可以采用文献中介绍的任何一种可伸缩图像压缩技术来对各个图像分量比特流进行可伸缩编码。具体而言，最好使用在上述描述的 Secker 和 Taubman 的由最近的 JPEG2000 图像压缩标准导出的方法来有效地对运动数据进行操作。在最简单的形式中，本发明不涉及在嵌入式量化和编码之前对运动矢量分量进行重组。注意，这与现有技术不同，在现有技术中使用 JPEG2000 的部分比特平面编码技术的变种联合地对各个运动矢量进行嵌入式量化和编码。

[0032] 视频的高效重构需要编码采样信息和两个编码运动分量表示之间精确的码率分配。这是由辅助码率分配信息实现的，辅助码率分配信息指定了取决于空间分辨率和比特率等的预期重构参数的运动和采样数据的最优组合。重构所需的辅助率信息（由视频服务器提供或来自压缩文件）包括一组与上述现有技术类似的表。但是，在本发明中，码率表对于各个所需的重构比特率和空间分辨率确定了两个（不是一个）运动比特率，以及视频采样比特率。

[0033] 实际上，只需指定与所选择的重构比特率相应的运动分量和采样比特率就够了。或者，码率表可以指定要用于所选择的重构比特率和空间分辨率的运动分量的数目和采样质量层的数目。如果预期重构率落在码率表所指定的总比特率之间，则使用与低一档的总比特率相对应的码率分配，并把剩下的比特分配给采样数据。这个规则具有总是按照与运动分量质量层的总数相应的比特率重构运动比特流的属性，这保证了运动比特流本身是码率失真最优的。另外，这促进了运动信息的保留性分配，意味着运动和采样数据的平衡有利于发送稍多的运动信息，而不是稍少。或者，可以将 R-D 斜率近似为在质量层之间线性地变化，从而质量层之间的失真 - 长度曲线成为二次多项式。这个近似提供了利用部分运动和采样质量层分配比特率的手段。

[0034] 可以对可组合的运动分量和采样比特率的每一种组合重构一次视频，以获得各个重构比特率，从而确定码率表。对于这种方式，需要把搜索限制为仅包含与质量层的总数相应的运动分量和采样率。在特定总比特率的每一次重构之后，测量 PSNR 并选择具有最高 PSNR 的组合作为最优码率分配。

[0035] 优选地，首先进行粗略搜索（其中把两个运动分量比特率约束为相同）来加速上述方法。这会涉及到按照与上述现有技术中描述的相同方式（上述的 Secker 和 Taubman），对于各个总比特率仅测试比特率对（运动和采样比特率）。这个方法会产生很好的初始猜测结果，因为最优运动分量比特率通常是在同一量级。通过尝试与初始猜测所确定的比特率接近的几个运动分量比特率来细化该初始猜测结果。同样，把搜索限制为与总运动分量层相应的运动比特率。

[0036] 为了以更高的运算效率确定码率表，可以利用当对各个数据源进行截取，使得在各个截取点处各个比特流的失真长度斜率相同时，三个数据源的组合最优的事实。即，利用了 Lagrangian 优化目标涉及截取三个比特流，使得对于一些斜率  $\lambda > 0$ ,  $\frac{-\Delta D^{(s)}}{\Delta L^{(s)}} \geq \lambda$  且

$\Psi_{R,S}^1 \frac{-\Delta D_M^1}{\Delta L^{(M,1)}} \geq \lambda$  且  $\Psi_{R,S}^2 \frac{-\Delta D_M^2}{\Delta L^{(M,2)}} \geq \lambda$ , 其中  $L^{(S)} + L^{(M,1)} + L^{(M,2)}$  尽可能大, 但不超过  $L^{\max}$ 。这个问题类似于 A. Secker 和 D. Taubman 在“Highly scalable video compression with scalable motion coding”中描述的问题, 其解决方法称为“基于模型的码率分配”。本发明本质上涉及相同的码率分配过程, 只是使用了两个运动分量比特流, 以及两个运动敏感系数。通过计算下面的积分而求得运动敏感系数, 其中  $SR, S(\omega_1, \omega_2)$  是利用恰当的功率谱预测方法确定的。

$$[0037] \quad \Psi_{1,R,S} = \frac{1}{(2\pi)^2} \iint S_{R,S}(\omega_1, \omega_2) \omega_1^2 d\omega_1 d\omega_2$$

$$[0038] \quad \Psi_{2,R,S} = \frac{1}{(2\pi)^2} \iint S_{R,S}(\omega_1, \omega_2) \omega_2^2 d\omega_1 d\omega_2$$

[0039] 如上述 Secker 和 Taubman 所报导并由这些公式所指出的, 高效的码率分配通常需要对各个空间分辨率  $S$  和对于所选择的重构比特率  $R$  使用不同的运动敏感系数对。

[0040] 前一实施例的限制在于分别地对运动矢量分量进行编码和传输会降低整个运动矢量集的压缩效率。这有两个原因。第一个是各个运动比特流需要报头信息来表示各种重构参数, 如空间维度, 以及与比特流的最优截取相关的信息。后者以各种形式存在, 包括码块、质量层、时空子带等的识别标记。当本发明中的两个运动分量比特流代替现有技术中的一个运动矢量比特流时, 这个开销大约要翻倍。为了减小两个运动分量比特流所需的信令开销, 最好将这两个分量比特流折叠成一个比特流, 使得在各个运动分量之间共享各种标记。这通常至少会包括时空子带标记、维度信息、时空分解和嵌入式编码参数。

[0041] 分别地对运动矢量分量进行编码会降低压缩效率的第二个原因在于这阻止了我们利用这两个运动分量之间的冗余。为了减小这个效应, 本发明的另一个实施例涉及到在嵌入式量化和编码之前对两个运动矢量分量进行重组。注意, 这意味着我们不能独立地在两个运动分量比特流之间分配比特, 从而码率分配是次优的。但是, 这可以通过利用两个比特流之间的独立性而实现的更高编码效率而得到补偿。具体而言, 我们希望利用这个事实: 当一个运动分量为 0 时, 另一个运动分量很可能也为 0。这可以利用上下文编码 (context coding) 法来实现, 类似于上述 Secker 和 Taubman 所提出的。但是, 和现有技术不一样, 本发明还涉及到利用两个运动分量对重构视频失真的相对重要性。这可以通过对 Secker 和 Taubman 所描述的部分比特平面编码操作进行进一步的改进来实现。例如, 我们可以在比特平面编码之前对一个运动矢量分量进行伸缩操作。实现这个的一个简单的方式是将所有的垂直运动矢量分量采样向左移位多个比特  $N$ , 其中

$$[0042] \quad N = \log_2 \left( \frac{\Psi_1}{\Psi_2} \right)$$

[0043] 或者, 当  $N$  为负数时, 我们可以把水平运动矢量分量向左移位  $-N$ 。这个方法有效地改进了两个运动矢量分量之间的比特平面扫描顺序, 并且在概念上与 MPEG-4 Fine Granularity Scalability 编码方案中为实现基于内容的可伸缩性而使用的比特平面移位方法类似, 如 M. van derSchaar and Y-T Lin, “Content-based selective enhancement for streaming video”, IEEE Int. Conf. Image Proc. Vol. 2, pp. 977-980, Sept 2001 中所述。注意, 把比特移位参数传输给解码器, 从而可以在解压缩时恢复正确的幅值, 但是发送



这些参数所需的比特数很小,对于压缩性能没有多大的影响。

[0044] 本发明例如可以在基于计算机的系统中实施,或者利用适当的硬件和 / 或软件实施,或者在专用设备或专用模块(如芯片)中实施。在图 2 中示出了编码器以及具有用于执行相反解码操作的相应部件的相应解码器。

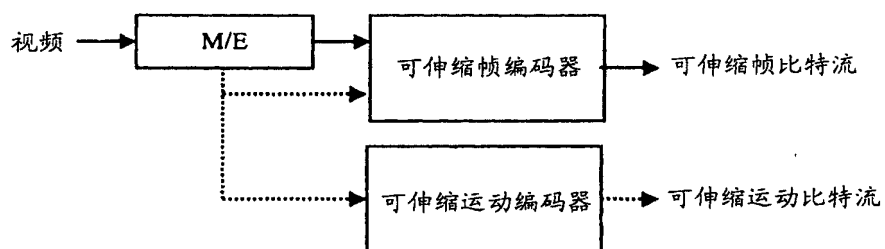


图 1

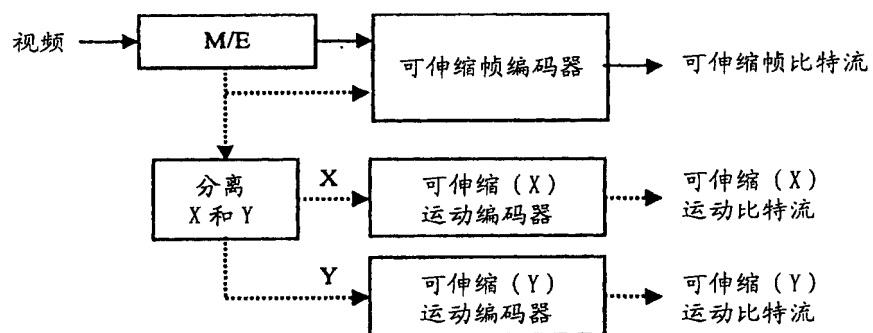


图 2