



(12) 发明专利

(10) 授权公告号 CN 1909670 B

(45) 授权公告日 2011.01.19

(21) 申请号 200510105360.9

(22) 申请日 2005.09.23

(30) 优先权数据

05254924.3 2005.08.05 EP

0421217.1 2004.09.23 GB

(73) 专利权人 三菱电机株式会社

地址 日本东京

(72) 发明人 斯塔夫罗斯·帕斯卡勒基斯

米罗斯瓦夫·博贝尔

(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 李辉

(51) Int. Cl.

H04N 9/64 (2006.01)

H04N 7/26 (2006.01)

(56) 对比文件

CN 1402853 A, 2003.03.12, 权利要求4、说明书第11页第28行至第12页第13行, 第17页第12-18行、附图6, 11.

CN 1358296 A, 2002.07.10, 说明书正文第3页第2段至第5页第4段、附图1, 2.

CN 1167959 A, 1997.12.17, 说明书正文第9页第3段、附图5-7.

CN 1495670 A, 2004.05.12, 全文.

CN 1365067 A, 2002.08.21, 全文.

JP 特开 2002-318812 A, 2002.10.31, 全文.

CN 1426002 A, 2003.06.25, 全文.

US 6477269 B1, 2002.11.05, 说明书正文第5栏第18行至第6栏第59行、第9栏第18-65行、附图3-5, 8-10.

CHAO-HUANG WANG ET AL. Image retrieval using block-based color distribution. SCI 2003. 7TH WORLD MULTICONFERENCE ON SYSTEMICS, CYBERNETICS AND INFORMATICS PROCEEDINGS, ORLANDO, FL, USA. 2003, 445-446.

审查员 黄海云

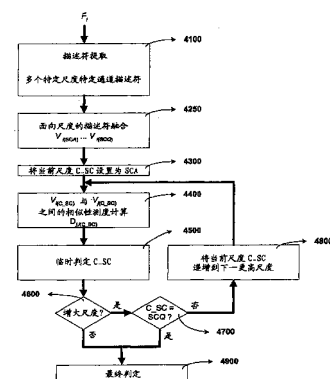
权利要求书 4 页 说明书 18 页 附图 11 页

(54) 发明名称

图像表示和分析方法

(57) 摘要

一种表示和分析图像的方法,其包括以下步骤:在一个或多个尺度下并针对一个或多个色彩通道生成多个描述符,该描述符捕获多个区域内的色彩内容和互相关信息;根据这些描述符的特征,如尺度、色彩通道、特征语义以及区域,按多种方式将它们关联起来;以及对这种图像表示进行比较以评估图像的相似性。



1. 一种表示图像的方法,其包括以下步骤:生成针对图像的对应的多个区域捕获色彩信息和色彩互相关信息的多个子描述符,至少一个区域按不同于其他区域的尺度;以及
通过将具有预定特征子描述符元素关联起来而从子描述符生成至少两个描述符,其中所述特征是色彩、色彩互相关、色彩通道、图像尺度和图像区域中的一个或多个,其中每个所述至少两个描述符具有所述子描述符的多个元素。
2. 如权利要求 1 所述的方法,其中,针对图像的多个区域生成子描述符。
3. 如权利要求 2 所述的方法,其中,每个所述区域都对应于 $n \times n$ 像素大小的块。
4. 如权利要求 1 所述的方法,其包括以下步骤:针对图像的多个非重叠区域中的每一个导出子描述符。
5. 如权利要求 1 所述的方法,其包括以下步骤:对图像进行处理以按多个尺度生成图像的多个表示,和导出针对所述多个尺度的子描述符。
6. 如权利要求 5 所述的方法,其中,所述处理包括向下采样(downsampling)。
7. 如权利要求 6 所述的方法,其中,所述向下采样包括对多个像素块进行平均。
8. 如权利要求 1 所述的方法,其包括以下步骤:针对多个色彩通道中的每个通道处理图像,和导出针对所述多个通道中的每个通道的子描述符。
9. 如权利要求 8 所述的方法,其中,对于不同色彩通道使用相同尺度的图像的区域。
10. 如权利要求 8 所述的方法,其中,对于不同通道使用不同尺度的图像的区域。
11. 如权利要求 1 所述的方法,其中,对于不同区域和/或不同尺度和/或不同色彩通道使用不同的用于导出子描述符的处理。
12. 如权利要求 1 所述的方法,其中,通过小波变换生成所述多个子描述符。
13. 如权利要求 12 所述的方法,其中,通过 Haar 或 Daubechies 小波变换生成所述多个子描述符。
14. 如权利要求 1 所述的方法,其中,使用非小波变换导出针对区域的至少一个子描述符元素。
15. 如权利要求 1 所述的方法,其中,仅使用区域中的像素子集导出针对该区域的至少一个子描述符元素。
16. 如权利要求 15 所述的方法,其中,仅使用区域中的像素子集导出针对区域的多个子描述符元素中的每一个或全部。
17. 如权利要求 1 所述的方法,其中,用于导出所述子描述符元素的计算使得所得子描述符元素数量比用于导出这些子描述符元素的像素数量要少。
18. 如权利要求 1 所述的方法,其中,利用公式 1-4:

$$LL = (F_{i(64 \times 64)}^Y(0,0) + F_{i(64 \times 64)}^Y(1,0) + F_{i(64 \times 64)}^Y(0,1) + F_{i(64 \times 64)}^Y(1,1))/2 \quad (1)$$

$$HL = (F_{i(64 \times 64)}^Y(0,0) - F_{i(64 \times 64)}^Y(1,0) + F_{i(64 \times 64)}^Y(0,1) - F_{i(64 \times 64)}^Y(1,1))/2 \quad (2)$$

$$LH = (F_{i(64 \times 64)}^Y(0,0) + F_{i(64 \times 64)}^Y(1,0) - F_{i(64 \times 64)}^Y(0,1) - F_{i(64 \times 64)}^Y(1,1))/2 \quad (3)$$

$$HH = (F_{i(64 \times 64)}^Y(0,0) - F_{i(64 \times 64)}^Y(1,0) - F_{i(64 \times 64)}^Y(0,1) + F_{i(64 \times 64)}^Y(1,1))/2 \quad (4)$$

或公式 6-9:

$$\begin{aligned} a_1 = & h_0 \cdot (h_0 \cdot F_{i(64 \times 64)}^Y(0,0) + h_1 \cdot F_{i(64 \times 64)}^Y(1,0) + h_2 \cdot F_{i(64 \times 64)}^Y(2,0) + h_3 \cdot F_{i(64 \times 64)}^Y(3,0)) \\ & + h_1 \cdot (h_0 \cdot F_{i(64 \times 64)}^Y(0,1) + h_1 \cdot F_{i(64 \times 64)}^Y(1,1) + h_2 \cdot F_{i(64 \times 64)}^Y(2,1) + h_3 \cdot F_{i(64 \times 64)}^Y(3,1)) \\ & + h_2 \cdot (h_0 \cdot F_{i(64 \times 64)}^Y(0,2) + h_1 \cdot F_{i(64 \times 64)}^Y(1,2) + h_2 \cdot F_{i(64 \times 64)}^Y(2,2) + h_3 \cdot F_{i(64 \times 64)}^Y(3,2)) \end{aligned}$$

$$+h_3 \cdot (h_0 \cdot F_{i(64 \times 64)}^Y(0, 3) + h_1 \cdot F_{i(64 \times 64)}^Y(1, 3) + h_2 \cdot F_{i(64 \times 64)}^Y(2, 3) + h_3 \cdot F_{i(64 \times 64)}^Y(3, 3)) \quad (6)$$

$$\begin{aligned} a_2 = & h_0 \cdot (h_3 \cdot F_{i(64 \times 64)}^Y(0, 0) - h_2 \cdot F_{i(64 \times 64)}^Y(1, 0) + h_1 \cdot F_{i(64 \times 64)}^Y(2, 0) - h_0 \cdot F_{i(64 \times 64)}^Y(3, 0)) \\ & + h_1 \cdot (h_3 \cdot F_{i(64 \times 64)}^Y(0, 1) - h_2 \cdot F_{i(64 \times 64)}^Y(1, 1) + h_1 \cdot F_{i(64 \times 64)}^Y(2, 1) - h_0 \cdot F_{i(64 \times 64)}^Y(3, 1)) \\ & + h_2 \cdot (h_3 \cdot F_{i(64 \times 64)}^Y(0, 2) - h_2 \cdot F_{i(64 \times 64)}^Y(1, 2) + h_1 \cdot F_{i(64 \times 64)}^Y(2, 2) - h_0 \cdot F_{i(64 \times 64)}^Y(3, 2)) \\ & + h_3 \cdot (h_3 \cdot F_{i(64 \times 64)}^Y(0, 3) - h_2 \cdot F_{i(64 \times 64)}^Y(1, 3) + h_1 \cdot F_{i(64 \times 64)}^Y(2, 3) - h_0 \cdot F_{i(64 \times 64)}^Y(3, 3)) \end{aligned} \quad (7)$$

$$\begin{aligned} a_3 = & h_3 \cdot (h_0 \cdot F_{i(64 \times 64)}^Y(0, 0) + h_1 \cdot F_{i(64 \times 64)}^Y(1, 0) + h_2 \cdot F_{i(64 \times 64)}^Y(2, 0) + h_3 \cdot F_{i(64 \times 64)}^Y(3, 0)) \\ & - h_2 \cdot (h_0 \cdot F_{i(64 \times 64)}^Y(0, 1) + h_1 \cdot F_{i(64 \times 64)}^Y(1, 1) + h_2 \cdot F_{i(64 \times 64)}^Y(2, 1) + h_3 \cdot F_{i(64 \times 64)}^Y(3, 1)) \\ & + h_1 \cdot (h_0 \cdot F_{i(64 \times 64)}^Y(0, 2) + h_1 \cdot F_{i(64 \times 64)}^Y(1, 2) + h_2 \cdot F_{i(64 \times 64)}^Y(2, 2) + h_3 \cdot F_{i(64 \times 64)}^Y(3, 2)) \\ & - h_0 \cdot (h_0 \cdot F_{i(64 \times 64)}^Y(0, 3) + h_1 \cdot F_{i(64 \times 64)}^Y(1, 3) + h_2 \cdot F_{i(64 \times 64)}^Y(2, 3) + h_3 \cdot F_{i(64 \times 64)}^Y(3, 3)) \end{aligned} \quad (8)$$

$$\begin{aligned} a_4 = & h_3 \cdot (h_3 \cdot F_{i(64 \times 64)}^Y(0, 0) - h_2 \cdot F_{i(64 \times 64)}^Y(1, 0) + h_1 \cdot F_{i(64 \times 64)}^Y(2, 0) - h_0 \cdot F_{i(64 \times 64)}^Y(3, 0)) \\ & - h_2 \cdot (h_3 \cdot F_{i(64 \times 64)}^Y(0, 1) - h_2 \cdot F_{i(64 \times 64)}^Y(1, 1) + h_1 \cdot F_{i(64 \times 64)}^Y(2, 1) - h_0 \cdot F_{i(64 \times 64)}^Y(3, 1)) \\ & + h_1 \cdot (h_3 \cdot F_{i(64 \times 64)}^Y(0, 2) - h_2 \cdot F_{i(64 \times 64)}^Y(1, 2) + h_1 \cdot F_{i(64 \times 64)}^Y(2, 2) - h_0 \cdot F_{i(64 \times 64)}^Y(3, 2)) \\ & - h_0 \cdot (h_3 \cdot F_{i(64 \times 64)}^Y(0, 3) - h_2 \cdot F_{i(64 \times 64)}^Y(1, 3) + h_1 \cdot F_{i(64 \times 64)}^Y(2, 3) - h_0 \cdot F_{i(64 \times 64)}^Y(3, 3)) \end{aligned} \quad (9)$$

其中,

$$h_0 = \frac{1+\sqrt{3}}{4\sqrt{2}}, h_1 = \frac{3+\sqrt{3}}{4\sqrt{2}}, h_2 = \frac{3-\sqrt{3}}{4\sqrt{2}}, h_3 = \frac{1-\sqrt{3}}{4\sqrt{2}},$$

或公式 14-17 :

$$x_1 = (F_{i(64 \times 64)}^Y(0, 0) + F_{i(64 \times 64)}^Y(1, 0) + F_{i(64 \times 64)}^Y(0, 1) + F_{i(64 \times 64)}^Y(1, 1)) / 4 \quad (14)$$

$$x_2 = (F_{i(64 \times 64)}^Y(0, 0) - F_{i(64 \times 64)}^Y(1, 0)) / 2 \quad (15)$$

$$x_3 = (F_{i(64 \times 64)}^Y(0, 0) - F_{i(64 \times 64)}^Y(0, 1)) / 2 \quad (16)$$

$$x_4 = (F_{i(64 \times 64)}^Y(0, 0) - F_{i(64 \times 64)}^Y(1, 1)) / 2 \quad (17)$$

或公式 I-IV :

$$\Phi_1 = (F_{i(64 \times 64)}^Y(0, 0) + F_{i(64 \times 64)}^Y(1, 0) + F_{i(64 \times 64)}^Y(0, 1) + F_{i(64 \times 64)}^Y(1, 1)) / 4 \quad (I)$$

$$\Phi_2 = (F_{i(64 \times 64)}^Y(0, 0) - F_{i(64 \times 64)}^Y(1, 0)) / 2 \quad (II)$$

$$\Phi_3 = (F_{i(64 \times 64)}^Y(1, 0) - F_{i(64 \times 64)}^Y(1, 1)) / 2 \quad (III)$$

$$\Phi_4 = (F_{i(64 \times 64)}^Y(1, 1) - F_{i(64 \times 64)}^Y(0, 1)) / 2 \quad (IV)$$

或公式 18 和 19 :

$$\Psi_1 = F_{i(64 \times 64)}^Y(0, 0) + F_{i(64 \times 64)}^Y(1, 0) + F_{i(64 \times 64)}^Y(0, 1) + F_{i(64 \times 64)}^Y(1, 1)) / 4 \quad (18)$$

$$\begin{aligned} \Psi_2 = & \max (F_{i(64 \times 64)}^Y(0,0), F_{i(64 \times 64)}^Y(1,0), F_{i(64 \times 64)}^Y(0,1), F_{i(64 \times 64)}^Y(1,1)) - \min (F_{i(64 \times 64)}^Y \\ & (0,0), \\ & F_{i(64 \times 64)}^Y(1,0), F_{i(64 \times 64)}^Y(0,1), F_{i(64 \times 64)}^Y(1,1)) \end{aligned} \quad (19)$$

导出子描述符元素,

其中, $F_{i(64 \times 64)}^Y(0,0)$ 、 $F_{i(64 \times 64)}^Y(1,0)$ 、 $F_{i(64 \times 64)}^Y(0,1)$ 以及 $F_{i(64 \times 64)}^Y(1,1)$ 为局部邻域的像素。

19. 如权利要求 1 所述的方法,其包括以下步骤:通过去除特定描述符元素减少子描述符的维数。

20. 如权利要求 1 所述的方法,其中,针对单个特征值或针对特征值的范围将多个子描述符或子描述符元素关联起来。

21. 如权利要求 1 所述的方法,其中,不进行解码而直接从 MPEG 编码 I 帧导出所述子描述符。

22. 一种表示图像组的方法,其包括以下步骤:使用前述任一权利要求所述的方法导出各图像的表示;以及使用所述表示来表示所述图像组。

23. 如权利要求 22 所述的方法,其包括以下步骤:导出所述表示的函数以表示所述图像组,其中,所述函数包括平均、中值、平均值、协方差的统计测度。

24. 一种对图像和/或图像组进行比较的方法,其包括对使用前述任一权利要求导出的所述的各表示进行比较的步骤,其中,该比较步骤包括对根据特征而相对应的多个描述符进行比较的步骤。

25. 如权利要求 24 所述的方法,其用于确定相似性。

26. 如权利要求 24 或 25 所述的方法,其包括并行地进行多个描述符比较的步骤。

27. 如权利要求 24 所述的方法,其包括串行地进行多个描述符比较的步骤。

28. 如权利要求 27 所述的方法,其包括以下步骤:利用两个或更多个描述符进行临时评估;以及利用该临时评估的结果确定是否执行涉及附加描述符的进一步评估。

29. 如权利要求 28 所述的方法,其中,临时评估和进一步评估涉及针对相同特征的不同值的描述符或针对不同特征的描述符。

30. 如权利要求 29 所述的方法,其包括按由粗到细的增加尺度进行连续评估的步骤。

31. 如权利要求 29 或 30 所述的方法,其使用基于以下信息中的一个或更多个的描述符:(i) 根据亮度提取的色彩内容,(ii) 根据亮度提取的色彩互相关,(iii) 根据色度提取的色彩内容,以及(iv) 根据色度提取的色彩互相关。

32. 如权利要求 31 所述的方法,其中,将基于(i)和(iii)的描述符用于至少一个评估,而将基于(ii)和(iv)的描述符用于至少一个连续评估。

33. 如权利要求 24 所述的方法,其包括对作为图像序列的一部分的多幅图像进行比较的步骤,并且其中,将比较结果用于确定是否已发生帧不连续性的事件。

34. 如权利要求 24 所述的方法,其用于图像检索。

35. 如权利要求 33 或 34 所述的方法,其包括以下步骤:对色度和亮度进行评估,以区分拍摄转场与照明效果,或者用于对在不同照明条件下捕获的相似图像进行检索。

36. 利用如权利要求 1 到 23 中的任何一项所述的方法对图像或图像组进行编码的方法。

37. 对利用如权利要求 36 所述的方法编码的图像或图像序列进行解码的方法。
38. 用于根据如权利要求 1 所述的方法对图像或图像组进行表示的设备。
39. 用于根据如权利要求 25 所述的方法对图像和 / 或图像组进行比较的设备。
40. 如权利要求 38 或 39 所述的设备,其包括用于存储图像数据和 / 或图像描述符的存储装置,和用于执行所述方法的处理装置。

图像表示和分析方法

技术领域

[0001] 本发明涉及用于表示图像的方法和装置,以及用于对图像间的相似性进行评估的方法和装置。

背景技术

[0002] 在 Lienhart, R., “Comparison of Automatic Shot Boundary Detection Algorithms”, In Proceedings of Image and Video Processing VII 1999, Proc. SPIE 3656-29, pp. 290-301, Jan. 1999 中,提出了一种用于检测视频序列中的拍摄转场 (shot transition) 的方法。针对序列中的每个帧,创建 RGB 色空间中的 3 维直方图。然后将序列中的多个连续帧的差异计算为它们相应的直方图之间的差异、计算为绝对二元 (bin-wise) 差异的总和。然后通过搜索预定固定阈值以上的距离来识别拍摄转场。因此,该方法仅根据空间非敏感的色彩内容信息来检测拍摄转场。因此,该方法未利用存在于空间排列和色彩互相关 (interrelation) 中的信息资源。

[0003] 在 Zabih, R., Miller, J., Mai, K., “A Feature-Based Algorithm for Detecting and Classifying Scene Breaks”, In Proceedings of 1995 3rd ACM International Conference on Multimedia, San Francisco, CA USA, pp. 189-200, 1995 中,提出了一种用于检测视频序列中的拍摄转场的不同方法。针对序列中的每个帧,计算出边沿图。然后根据存在于第一帧中而不存在于第二帧中的边沿的数量以及存在于第二帧中而不存在于第一帧中的边沿的数量,来计算序列中多个连续帧之间的差异。然后,该差异测度的时间序列中的峰点 (sharp peak) 表示拍摄转场的存在性。因此,该方法仅根据边沿信息 (其作为一种类型的空间互相关信息) 检测拍摄转场。尽管其原理是正确的,但是该方法未利用存在于帧的色彩内容中的信息资源。此外,边沿图创建过程很耗计算机时并且意味着仅揭示帧内最强的色彩不连续性。此外,该方法对运动相当敏感。因此,作者建议使用图象配准技术来抵消该缺点,但是这种处理很耗计算机时。

[0004] 在 Dailianas, A., Allen, R. B., England, P., “Comparison of Automatic Video Segmentation Algorithms”, SPIE Integration Issues in Large Commercial Media Delivery Systems, vol. 2615, pp. 2-16, Oct. 1995 中,提出了用于检测视频序列中的拍摄转场的另一方法。将序列中的多个连续帧之间的差异计算为绝对像素方面 (pixel-wise) 差异的总和。然后通过搜索预定固定阈值以上的距离来识别拍摄转场。因此,该方法仅根据空间敏感的色彩内容信息来检测拍摄转场。尽管其原理是正确的,但是该方法未利用存在于色彩的空间互相关中的信息资源。此外,对视频结果的这种简单处理导致对噪声和运动高度敏感性。一种运动补偿算法可以解决该运动敏感性问题,但是这种处理很耗计算机时。

[0005] 在 Xiong, W., “Shot Boundary Detection”, US 2003/0091235 A1, published 15 May, 2003 中,提出了一种基于不同类型信息的组合检测拍摄转场的方法。该方法包括计算两个帧之间的块基 (block-based) 差异,如果该块基差异超过固定阈值,则表明是候选拍摄转场。在此情况下,通过规定这两个帧之间的色彩和 / 或边沿差异也超过预定阈值,验证

拍摄转场。为了对块基差异进行计算,将帧划分成多个块并计算块平均。然后,对相应块之间的差异进行阈值比较以确定两个块是相似的还是不同的,并对两个帧之间的不同块的数量进行阈值比较以确定两个帧是相似的还是不同的。色差是绝对双向差异的总和,而边沿差异利用边沿直方图,以捕获边沿量级和方向信息。

[0006] 在 Nakajima, Y., Sugano, M., Yanagihara, H., for KDDICORPORATION(JP), “Picture Searching Apparatus”, US 2004/0091044A1, published 13 May, 2004 中,提出了一种用于检测拍摄转场的方法,其基于以下信息:(a) 图像之间的相关性,(b) 子采样图像之间的相关性,(c) 图像之间的运动,以及(d) 子采样图像之间的运动。其中,将图像之间和子采样图像之间的相关性测量为像素向差异或直方图差异,并且根据各种运动矢量差测量图像之间和子采样图像之间的运动。

[0007] 在 Jafarkhani, H., Shahraray, B., for AT&T CORP. (US), “Method for Analyzing Video”, US 6,542,619 B1, granted 1 April, 2003 中,提出了一种拍摄转场检测方法,其包括以下步骤:创建视频帧的两个一维投影,即行和列投影;对每个投影执行小波变换并只保留高频分量(即,小波系数);以及对每次变换的高频分量进行自相关。对于一系列视频帧,当所得自相关系数时间曲线显示出预定最高值时,表示拍摄转场。因此,该方法使用了由小波变换提供的空间敏感色彩内容和互相关信息,但是该信息不是关于帧的而是关于帧投影的,导致很大的信息损失。

[0008] 在 Jacobs, C. E., Finkelstein, A., Salesin, D. H., “Fast Multiresolution Image Querying”, In Proceedings of 1995 ACM SIGGRAPH Conference, Los Angeles CA, USA, Aug. 9-11, pp. 277-286, 1995 中,提出了一种用于检索与给定图像相似的图像的方法。根据该方法,最初通过图像的 Haar 小波分解来表示图像。然后,截取该分解,即,仅保留标度函数系数(平均强度)和很少数量的最大幅度小波系数。然后,对截取的分解进行量化,即,仅保留小波系数的符号。由此,形成了表征图像的单个图像描述符以用于图像检索。

[0009] 在 Zhuang, Z.-Y., Hsu, C.-T., Chen, H.-Y., Ouhyoung, M., Wu, J.-L., “Efficient Multiresolution Scene Change Detection by Wavelet Transformation”, In Proceedings of 1997 IEEE International Conference on Consumer Electronics ICCE '97, Taipei, Taiwan, Jun. 11-13, pp. 250-251, 1997 中,提出了一种用于检测拍摄转场的方法,其按与在“Fast Multiresolution Image Querying(快速多分辨率图像查询)”中所描述的不同方式对视频帧进行表征。“Fast Multiresolution Image Querying”方法与“Efficient Multiresolution Scene Change detection by Wavelet Transformation(利用小波变换的高效多分辨率场景变化检测)”方法之间的差别在于,根据后一方法,舍弃帧的周边并将帧缩减到只剩它们的中央部分。这种方法会引起很大的信息损失,并且当在视频中存在显著运动时会导致假视频分割和/或很大的过分割。

[0010] 上述两种方法共有的缺陷是这样的假设,即,通过只保留很少数量的最大幅度系数可以有效地截取小波分解。为了阐述这一点,从 128×128 像素开始到 2×2 像素的图像平面的多尺度小波分解将产生 16383 个小波系数。如本领域技术人员所了解的,根据量级将该序列截取为很小数量的系数(例如,如作者所建议的 40 或 60 个最大量级系数),会得到这样的描述符,即,其极易受噪声影响、易受图像检索和视频分割的部分遮挡的影响、以及易受视频分割的高视频运动和内摄照明效应的影响,存在一些问题。通过只保留符号对

所截取序列进行量化会放大该问题。

[0011] 根据这些方法的另一显著问题在于,未利用 Haar 小波分解的系数所附的语义信息 (semantic information)。这种语义信息包括:系数所表示的特定色彩信息,例如 RGB 的 R 或 YC_bC_r 的 Y;其中存在系数的特定图像尺度,例如,它是在捕获精细细节的高图像尺度下的系数或是在捕获粗图像信息的低图像尺度下的系数等。

[0012] 其中,对图像之间的相似度进行评估的方法例如以从与给定图像相似的一组图像检索图像或者以检测数字视频中的帧不连续性(如拍摄转场或照明和其他效应)开始。这些方法依赖于以下操作:提取这样的图像描述符,即,其捕获了在一个或多个图像尺度下的并且在一个或多个图像通道上的空间敏感色彩内容和互相关信息;接着,不按单个描述符而按多个描述符(由语义内容区分这些描述符)组合描述符;以及在有效利用所述语义内容的多重判定框架中使用这些描述符。因此,与先前的方法不同,可以建立图像之间的复杂关联,例如建立这样的关系:两幅图像表示同一场景,但是一幅图像存在很明显的遮挡(如一个人走到相机跟前);或者两幅图像表示同一场景,却是在不同的照明条件下拍摄的;或者两个帧属于同一次拍摄,但是由于整体照明效果而看起来很不相同。

发明内容

[0013] 在所附权利要求中陈述了本发明的多个方面。

[0014] 根据本发明一个方面,根据对关于图像的一个或多个区域的色彩信息和/或色彩互相关信息进行捕获的子描述符的特征(例如,区域、尺度、色彩、色彩互相关、色彩通道等),把这些子描述符关联起来以形成两个或多个描述符。

[0015] 子描述符可以具有也捕获色彩和/或色彩互相关信息的元素。在此情况下,这些子描述符元素也可以构成子描述符。例如,图像区域的 Haar 小波变换的四个测度 LL、HL、LH 以及 HH 一起形成该图像区域的子描述符。然而,也可以把 LL 本身或 HH 本身看成它们本身的子描述符。

[0016] 所述“将子描述符关联起来”的步骤可以包括以下步骤:将它们拆分成它们的构成元素并根据对应特征对所述元素进行分组或关联。

附图说明

[0017] 以下参照附图对本发明实施例进行描述,附图中:

[0018] 图 1 示出了根据本发明一实施例的针对 Y 通道的块平均再采样过程;

[0019] 图 2A-2D 示出了在本发明一个实施例中对尺度描述符进行的计算;

[0020] 图 3A-3D 示出了根据本发明另一实施例对尺度描述符进行的计算;

[0021] 图 4 是根据本发明一实施例的用于对多幅图像进行比较的流程图;

[0022] 图 5 是根据本发明另一实施例的用于对多幅图像进行比较的流程图;

[0023] 图 6 是根据本发明另一实施例的用于对多幅图像进行比较的流程图;

[0024] 图 7 是根据本发明另一实施例的用于对多幅图像进行比较的流程图;

[0025] 图 8 是根据本发明另一实施例的用于对多幅图像进行比较的流程图;

[0026] 图 9 是根据本发明另一实施例的用于对多幅图像进行比较的流程图;以及

[0027] 图 10 是适于实现本发明实施例的处理设备的示意图。

具体实施方式

[0028] 考虑数字图像 $F_i(x, y)$, 其中 (x, y) 表示空间坐标, $x = 0 \cdots M-1$ 并且 $y = 0 \cdots N-1$ 。在本发明一个实施例中, 在色彩空间 YC_bC_r 中表示 F_i , 但是这并非限制性的, 本发明实施例可以适用任何数量个通道的任何色彩空间。在本发明一个实施例中, F_i 的空间分辨率是 720×576 个像素, 即, $M = 720$ 并且 $N = 576$, 但是这并非限制性的, 本发明实施例可以适用任何空间分辨率。在本发明一个实施例中, 把要提取描述符所用最高分辨率选择为 64×64 个像素, 但是这并非限制性的, 本发明实施例可以适用其他分辨率。因此, 对 F_i 进行再采样以得到 64×64 像素图像 $F_i(64 \times 64)$ 。在本发明一个实施例中, 该再采样过程是如图 1 所示的简单块平均过程, 但是对再采样机制的选择也并非限制性的。更具体来说, 图 1 示出了针对 Y 通道的块平均再采样过程, 但是这并非限制性的, 而是可以将该过程用于任何色彩空间的任何通道。在本发明另一实施例中, 要提取描述符所用的最高尺度与 F_i 的分辨率相一致并且最初不必进行再采样。

[0029] 图 2 示出了在本发明一个实施例中对该尺度的描述符进行的计算。更具体来说, 图 2 示出了对 Y 通道的描述符的计算, 但是这并非限制性的, 而是可以将类似方法应用于所有色彩通道。在图 2A 中, 将 2×2 窗口应用于图像的左上角。该窗口指定了局部邻域。对于该邻域中的像素 $F_{i(64 \times 64)}^Y(0, 0)$ 、 $F_{i(64 \times 64)}^Y(1, 0)$ 、 $F_{i(64 \times 64)}^Y(0, 1)$ 以及 $F_{i(64 \times 64)}^Y(1, 1)$, 可以将 2 维 Haar 变换系数 LL、HL、LH 以及 HH 计算成:

$$[0030] \quad LL = (F_{i(64 \times 64)}^Y(0, 0) + F_{i(64 \times 64)}^Y(1, 0) + F_{i(64 \times 64)}^Y(0, 1) + F_{i(64 \times 64)}^Y(1, 1)) / 2 \quad (1)$$

$$[0031] \quad HL = (F_{i(64 \times 64)}^Y(0, 0) - F_{i(64 \times 64)}^Y(1, 0) + F_{i(64 \times 64)}^Y(0, 1) - F_{i(64 \times 64)}^Y(1, 1)) / 2 \quad (2)$$

$$[0032] \quad LH = (F_{i(64 \times 64)}^Y(0, 0) + F_{i(64 \times 64)}^Y(1, 0) - F_{i(64 \times 64)}^Y(0, 1) - F_{i(64 \times 64)}^Y(1, 1)) / 2 \quad (3)$$

$$[0033] \quad HH = (F_{i(64 \times 64)}^Y(0, 0) - F_{i(64 \times 64)}^Y(1, 0) - F_{i(64 \times 64)}^Y(0, 1) + F_{i(64 \times 64)}^Y(1, 1)) / 2 \quad (4)$$

[0034] 在上述公式中, LL 捕获色彩内容信息, 而 HL、LH 以 HH 捕获色彩互相关信息。本领域的技术人员将理解, 例如通过首先在水平方向上然后在垂直方向上执行计算, 可以对公式 (1)-(4) 进行修改以减少所需计算次数。根据公式 (1)-(4) 或它们的最优化但是以值 4 而非值 2 作为分母来计算 Haar 系数, 也是很常见的。

[0035] 以 $V_{i(64 \times 64)}^Y$ 表示 $F_{i(64 \times 64)}^Y$ 的描述符, 可以如下将这些值写入描述符中:

$$[0036] \quad V_{i(64 \times 64)}^Y(1) = LL, V_{i(64 \times 64)}^Y(2) = HL, V_{i(64 \times 64)}^Y(3) = LH, V_{i(64 \times 64)}^Y(4) = HH \quad (5)$$

[0037] 接下来, 如图 2B 所示, 所述 2×2 窗口向右滑动 2 个像素以形成包括像素 $F_{i(64 \times 64)}^Y(2, 0)$ 、 $F_{i(64 \times 64)}^Y(3, 0)$ 、 $F_{i(64 \times 64)}^Y(2, 1)$ 以及 $F_{i(64 \times 64)}^Y(3, 1)$ 的新邻域。针对新像素值再计算 Haar 变换系数 LL、HL、LH 以及 HH, 然后把它们存储在描述符位置 $V_{i(64 \times 64)}^Y(5)$ 到 $V_{i(64 \times 64)}^Y(8)$ 中。如图 2C 和 2D 所示, 一旦完成了对一组行的所有计算, 则滑动窗口向下移动两个像素并再从左向右滑动, 直到处理完整幅图像为止。这完成了对描述符 $V_{i(64 \times 64)}^Y$ 的计算, 该描述符 $V_{i(64 \times 64)}^Y$ 在 64×64 像素的尺度下捕获了针对 Y 平面的图像空间敏感色彩内容和互相关。

[0038] 本发明多个实施例涉及对捕获空间敏感色彩内容和互相关信息的描述符的计算。

[0039] 因此, 在以上描述中, 针对 LL、HL、LH 以及 HH 的公式 (1)-(4) 仅仅是在邻域内可以进行的度量的一个示例, 并且仅代表一个实施例。Haar 变换是小波变换的一个简单示例, 其生成了信号的多分辨率时频表示。本发明另一实施例使用另一类似的但是更复杂的小波

变换,即 2 维 Daubechies 小波变换,而非使用 Haar 变换。与前一变换不同,该变换按 4×4 窗口以 2 个像素的滑动步长进行操作,即,所得邻域是相交叠的。图 3 示出了该变换。在图 3A 中,将 4×4 窗口应用于图像的左上角。该窗口指定了局部邻域。对于该邻域中的像素,计算以下描述符元素:

$$[0040] \quad a_1 = h_0 \cdot (h_0 \cdot F_{i(64 \times 64)}^Y(0,0) + h_1 \cdot F_{i(64 \times 64)}^Y(1,0) + h_2 \cdot F_{i(64 \times 64)}^Y(2,0) + h_3 \cdot F_{i(64 \times 64)}^Y(3,0)) + h_1 \cdot (h_0 \cdot F_{i(64 \times 64)}^Y(0,1) + h_1 \cdot F_{i(64 \times 64)}^Y(1,1) + h_2 \cdot F_{i(64 \times 64)}^Y(2,1) + h_3 \cdot F_{i(64 \times 64)}^Y(3,1)) + h_2 \cdot (h_0 \cdot F_{i(64 \times 64)}^Y(0,2) + h_1 \cdot F_{i(64 \times 64)}^Y(1,2) + h_2 \cdot F_{i(64 \times 64)}^Y(2,2) + h_3 \cdot F_{i(64 \times 64)}^Y(3,2)) + h_3 \cdot (h_0 \cdot F_{i(64 \times 64)}^Y(0,3) + h_1 \cdot F_{i(64 \times 64)}^Y(1,3) + h_2 \cdot F_{i(64 \times 64)}^Y(2,3) + h_3 \cdot F_{i(64 \times 64)}^Y(3,3)) \quad (6)$$

$$[0041] \quad a_2 = h_0 \cdot (h_3 \cdot F_{i(64 \times 64)}^Y(0,0) - h_2 \cdot F_{i(64 \times 64)}^Y(1,0) + h_1 \cdot F_{i(64 \times 64)}^Y(2,0) - h_0 \cdot F_{i(64 \times 64)}^Y(3,0)) + h_1 \cdot (h_3 \cdot F_{i(64 \times 64)}^Y(0,1) - h_2 \cdot F_{i(64 \times 64)}^Y(1,1) + h_1 \cdot F_{i(64 \times 64)}^Y(2,1) - h_0 \cdot F_{i(64 \times 64)}^Y(3,1)) + h_2 \cdot (h_3 \cdot F_{i(64 \times 64)}^Y(0,2) - h_2 \cdot F_{i(64 \times 64)}^Y(1,2) + h_1 \cdot F_{i(64 \times 64)}^Y(2,2) - h_0 \cdot F_{i(64 \times 64)}^Y(3,2)) + h_3 \cdot (h_3 \cdot F_{i(64 \times 64)}^Y(0,3) - h_2 \cdot F_{i(64 \times 64)}^Y(1,3) + h_1 \cdot F_{i(64 \times 64)}^Y(2,3) - h_0 \cdot F_{i(64 \times 64)}^Y(3,3)) \quad (7)$$

$$[0042] \quad a_3 = h_3 \cdot (h_0 \cdot F_{i(64 \times 64)}^Y(0,0) + h_1 \cdot F_{i(64 \times 64)}^Y(1,0) + h_2 \cdot F_{i(64 \times 64)}^Y(2,0) + h_3 \cdot F_{i(64 \times 64)}^Y(3,0)) - h_2 \cdot (h_0 \cdot F_{i(64 \times 64)}^Y(0,1) + h_1 \cdot F_{i(64 \times 64)}^Y(1,1) + h_2 \cdot F_{i(64 \times 64)}^Y(2,1) + h_3 \cdot F_{i(64 \times 64)}^Y(3,1)) + h_1 \cdot (h_0 \cdot F_{i(64 \times 64)}^Y(0,2) + h_1 \cdot F_{i(64 \times 64)}^Y(1,2) + h_2 \cdot F_{i(64 \times 64)}^Y(2,2) + h_3 \cdot F_{i(64 \times 64)}^Y(3,2)) - h_0 \cdot (h_0 \cdot F_{i(64 \times 64)}^Y(0,3) + h_1 \cdot F_{i(64 \times 64)}^Y(1,3) + h_2 \cdot F_{i(64 \times 64)}^Y(2,3) + h_3 \cdot F_{i(64 \times 64)}^Y(3,3)) \quad (8)$$

$$[0043] \quad a_4 = h_3 \cdot (h_3 \cdot F_{i(64 \times 64)}^Y(0,0) - h_2 \cdot F_{i(64 \times 64)}^Y(1,0) + h_1 \cdot F_{i(64 \times 64)}^Y(2,0) - h_0 \cdot F_{i(64 \times 64)}^Y(3,0)) - h_2 \cdot (h_3 \cdot F_{i(64 \times 64)}^Y(0,1) - h_2 \cdot F_{i(64 \times 64)}^Y(1,1) + h_1 \cdot F_{i(64 \times 64)}^Y(2,1) - h_0 \cdot F_{i(64 \times 64)}^Y(3,1)) + h_1 \cdot (h_3 \cdot F_{i(64 \times 64)}^Y(0,2) - h_2 \cdot F_{i(64 \times 64)}^Y(1,2) + h_1 \cdot F_{i(64 \times 64)}^Y(2,2) - h_0 \cdot F_{i(64 \times 64)}^Y(3,2)) - h_0 \cdot (h_3 \cdot F_{i(64 \times 64)}^Y(0,3) - h_2 \cdot F_{i(64 \times 64)}^Y(1,3) + h_1 \cdot F_{i(64 \times 64)}^Y(2,3) - h_0 \cdot F_{i(64 \times 64)}^Y(3,3)) \quad (9)$$

其中,

$$[0044] \quad h_0 = \frac{1 + \sqrt{3}}{4\sqrt{2}} \quad (10)$$

$$[0045] \quad h_1 = \frac{3 + \sqrt{3}}{4\sqrt{2}} \quad (11)$$

$$[0046] \quad h_2 = \frac{3 - \sqrt{3}}{4\sqrt{2}} \quad (12)$$

$$[0047] \quad h_3 = \frac{1 - \sqrt{3}}{4\sqrt{2}} \quad (13)$$

[0048] 在以上公式中, a_1 是标度函数系数,捕获色彩内容信息, a_2 、 a_3 以及 a_4 是小波变换系数,捕获色彩互相关信息。本领域的技术人员将理解,可以对公式 (6)–(13) 进行修改以减少所需计算次数。可以沿公式 (5) 的行把这些值写入描述符 $V_{i(64 \times 64)}^Y$ 中,然后,如图 3B 所

示,将该 4×4 窗口向右滑动2个像素以形成新邻域,针对新像素值再计算描述符元素,依此类推。如图3C所示,一旦滑动窗口到达图像的右边沿,那么就没有足够的像素以形成完整的 4×4 邻域。可以通过多种方法处理该问题。这些方法中的一个假设对数据进行镜像映射,如图3C所示。另一方法是假设数据是周期性的。还存在其他方法。当滑动窗口到达图像的下边沿时也会引起同样的问题,并可以如上所述地处理该问题。如图3D所示,一旦完成了对一组行的所有计算,则滑动窗口向下移动两个像素并再从左向右滑动,直到处理完整幅图像为止。

[0049] 在以上描述中,针对描述符的计算使用了小波变换。另选实施例可以依靠用于计算色彩内容和色彩互相关描述符元素的其他过程。例如,一另选实施例可以依靠产生色彩内容和互相关描述符元素的描述符提取过程,但是,与上述过程不同,每像素邻域所计算出的描述符元素数量比该邻域中的像素数量要少,从而得到更小的描述符,这更便于存储和处理但是仍然给出了对图像的空间均匀覆盖。再一另选实施例可以基于这样的描述符提取过程,即,与上述过程不同,通过该描述符提取过程,仅根据邻域像素的子集计算邻域的描述符元素,这将导致显著的加速。

[0050] 例如,对于如图2A所示的局部邻域的像素 $F_{i(64 \times 64)}^Y(0,0)$ 、 $F_{i(64 \times 64)}^Y(1,0)$ 、 $F_{i(64 \times 64)}^Y(0,1)$ 以及 $F_{i(64 \times 64)}^Y(1,1)$,本发明一另选实施例将描述符元素 x_1 到 x_4 计算成:

$$[0051] \quad x_1 = (F_{i(64 \times 64)}^Y(0,0) + F_{i(64 \times 64)}^Y(1,0) + F_{i(64 \times 64)}^Y(0,1) + F_{i(64 \times 64)}^Y(1,1)) / 4 \quad (14)$$

$$[0052] \quad x_2 = (F_{i(64 \times 64)}^Y(0,0) - F_{i(64 \times 64)}^Y(1,0)) / 2 \quad (15)$$

$$[0053] \quad x_3 = (F_{i(64 \times 64)}^Y(0,0) - F_{i(64 \times 64)}^Y(0,1)) / 2 \quad (16)$$

$$[0054] \quad x_4 = (F_{i(64 \times 64)}^Y(0,0) - F_{i(64 \times 64)}^Y(1,1)) / 2 \quad (17)$$

[0055] 因此,表示局部像素平均的元素 x_1 捕获色彩内容信息,而各表示一像素与另一像素的差别的元素 x_2 、 x_3 以及 x_4 捕获色彩互相关信息。可以沿公式(5)的行把这些值写入描述符 $V_{i(64 \times 64)}^Y$ 中,然后,如图2B所示,将该 2×2 窗口向右滑动2个像素以形成新邻域,针对新像素值再计算描述符元素,依此类推。

[0056] 本发明再一另选实施例将公式(14)–(17)替换成以下所示的公式(I)–(IV):

$$[0057] \quad \phi_1 = (F_{i(64 \times 64)}^Y(0,0) + F_{i(64 \times 64)}^Y(1,0) + F_{i(64 \times 64)}^Y(0,1) + F_{i(64 \times 64)}^Y(1,1)) / 4 \quad (I)$$

$$[0058] \quad \phi_2 = (F_{i(64 \times 64)}^Y(0,0) - F_{i(64 \times 64)}^Y(1,0)) / 2 \quad (II)$$

$$[0059] \quad \phi_3 = (F_{i(64 \times 64)}^Y(1,0) - F_{i(64 \times 64)}^Y(1,1)) / 2 \quad (III)$$

$$[0060] \quad \phi_4 = (F_{i(64 \times 64)}^Y(1,1) - F_{i(64 \times 64)}^Y(0,1)) / 2 \quad (IV)$$

[0061] 在以上所示的公式中, ϕ_1 捕获色彩内容信息,而 ϕ_2 – ϕ_4 捕获色彩互相关信息。

[0062] 本发明再一另选实施例将公式(14)–(17)替换成以下所示的公式(18)和(19):

$$[0063] \quad \Psi_1 = (F_{i(64 \times 64)}^Y(0,0) + F_{i(64 \times 64)}^Y(1,0) + F_{i(64 \times 64)}^Y(0,1) + F_{i(64 \times 64)}^Y(1,1)) / 4 \quad (18)$$

$$[0064] \quad \Psi_2 = \max(F_{i(64 \times 64)}^Y(0,0), F_{i(64 \times 64)}^Y(1,0), F_{i(64 \times 64)}^Y(0,1), F_{i(64 \times 64)}^Y(1,1))$$

$$1)) - \min(F_{i(64 \times 64)}^Y(0,0), F_{i(64 \times 64)}^Y(1,0), F_{i(64 \times 64)}^Y(0,1), F_{i(64 \times 64)}^Y(1,1)) \quad (19)$$

[0065] 在以上所示的公式中, Ψ_1 捕获色彩内容信息, 而 Ψ_2 捕获色彩互相关信息。

[0066] 此外, 尽管以上描述和图考虑了 2×2 和 4×4 像素大小的滑动邻域窗口, 但是这并非限制性的。另选实施例可以为邻域窗口指定任何尺寸和 / 或几何形状和 / 或滑动步长, 以例如适合于给定计算需要或邻域量度过程本身。

[0067] 在以上描述中, 在 $F_{i(64 \times 64)}^Y$ 的所有位置处使用相同的邻域窗口和描述符提取过程。在另选实现中不必如此。另选实施例可以在图像的不同位置处使用不同的邻域参数和 / 或描述符提取过程。例如, 本发明一另选实施例在图像的周边使用快速描述符提取过程, 如公式 (1)–(4) 的过程或公式 (14)–(17) 的过程, 而在图像的中央处使用较慢但是更强大的描述符提取过程, 如公式 (6)–(9) 的过程。

[0068] 接下来, 将 $F_{i(64 \times 64)}^Y$ 再采样到新的较小尺度。在一个实施例中, 该新尺度是 32×32 像素。然后, 按与用于 64×64 图像的方式完全相同的方式计算新描述符 $V_{i(32 \times 32)}^Y$ 。注意, 可能该再采样过程实际上不需要进行计算。如可从公式 (1) 看到的, 先将新像素值计算为 LL 值并将其存储在 $V_{i(64 \times 64)}^Y(1)$ 、 $V_{i(64 \times 64)}^Y(5)$ 、 $V_{i(64 \times 64)}^Y(9) \dots V_{i(64 \times 64)}^Y(4093)$ 。因此, 在一个实施例中, 计算出描述符 $V_{i(64 \times 64)}^Y$ 、 $V_{i(32 \times 32)}^Y$ 、 $V_{i(16 \times 16)}^Y$ 、 $V_{i(8 \times 8)}^Y$ 、 $V_{i(4 \times 4)}^Y$ 以及 $V_{i(2 \times 2)}^Y$ 。

[0069] 这里所概述的本发明实施例包括按图像 F_i 的一个或更多个尺度提取描述符。

[0070] 因此, 在以上描述中, 最细或最大或最高的尺度、最粗或最小或最低的尺度、尺度数量和尺度间的关联仅代表了本发明一个可能的实施例, 另选实施例可以使用不同的这种参数, 以例如与系统的计算能力相匹配。

[0071] 在本发明一个实施例中, 在随后处理之前从这些特定尺度的描述符中去除某些元素。在本发明一个实施例中, 对于公式 (1)–(4) (即, Haar 变换) 所概述的标识符提取过程, 在 $2^b \times 2^b$ 、 $2^{b-1} \times 2^{b-1}$ 、 $2^{b-2} \times 2^{b-2}$ 等尺度 (例如, 64×64 、 32×32 、 16×16 等) 上, 除了在最低尺度下的 LL 元素以外, 可以去除 V_i^Y 中的所有 LL 元素。这是因为根据在 $2^{b-k-1} \times 2^{b-k-1}$ 尺度上的 LL 值和在 $2^{b-k} \times 2^{b-k}$ 尺度下的 HL、LH 以及 HH 元素可以直接计算出在 $2^{b-k} \times 2^{b-k}$ 尺度下的 LL 元素。注意, 并非总是存在这个性质。例如, 如果在 $2^b \times 2^b$ 、 $2^{b-2} \times 2^{b-2}$ 、 $2^{b-4} \times 2^{b-4}$ 等尺度 (例如, 64×64 、 16×16 、 4×4 等) 下计算描述符, 那么该性质无效。此外, 如果只计算 LL、HL、LH 以及 HH 元素的一子集作为标识符提取过程的一部分, 则该性质也无效。

[0072] 等价性质也适用于 Daubechies 小波变换。即, 在本发明一个实施例中, 对于在 $2^b \times 2^b$ 、 $2^{b-1} \times 2^{b-1}$ 、 $2^{b-2} \times 2^{b-2}$ 等尺度 (例如, 64×64 、 32×32 、 16×16 等) 下的 Daubechies 小波变换, 除了在最低尺度下的 a_1 元素以外, 可以去除 V_i^Y 中的所有 a_1 元素。

[0073] 等价性质也适用于公式 (14)–(17) 的变换。即, 在本发明一个实施例中, 对于在 $2^b \times 2^b$ 、 $2^{b-1} \times 2^{b-1}$ 、 $2^{b-2} \times 2^{b-2}$ 等尺度 (例如, 64×64 、 32×32 、 16×16 等) 下的该变换, 除了在最低尺度下的 x_1 元素以外, 可以去除 V_i^Y 中的所有 x_1 元素。同样的性质也适用于公式 (I)–(IV) 的变换。

[0074] 把同样的描述符提取过程应用于 F_i 的 C_b 和 C_r 通道给出了描述符 $V_{i(64 \times 64)}^{Cb} \dots V_{i(2 \times 2)}^{Cb}$ 以及 $V_{i(64 \times 64)}^{Cr} \dots V_{i(2 \times 2)}^{Cr}$ 。所有这些特定尺度特定通道描述符都捕获了在多尺度下并在多个通道上的图像空间敏感色彩内容和互相关信息。

[0075] 在以上描述中, 假设 F_i 在 YC_bC_r 色彩空间中并且将所有通道用于对描述符的提取。另选实现可以在对描述符的计算之间在新色彩空间中表示该图像, 并且 / 或者在对其描述

符的计算中仅使用色彩平面的一子集。

[0076] 在以上描述中,在相同尺度下计算在 F_i 的不同色彩通道上的描述符。在另选实现中,不必如此。例如,本发明另选实施例在比 C_b 和 / 或 C_r 通道描述符更多的尺度并且 / 或者不同的尺度下计算 Y 通道描述符。这种细调将得到对特定图像或视频材料的更佳适应,并且可以改进性能。

[0077] 在以上描述中,将相同的邻域窗口和描述符提取过程用于 F_i 的所有色彩通道。在另选实现中,不必如此。另选实施例可以针对各通道使用不同的邻域参数和 / 或描述符提取过程。例如,本发明一另选实施例针对 C_b 和 C_r 通道使用快速描述符提取过程而针对 Y 通道使用较慢但是更强大的描述符提取过程。这种细调将得到对特定图像或视频材料的更佳适应,并且可以改进系统性能。

[0078] 在本发明一个实施例中,将一些或所有描述符元素表示成带有整数和小数部分的实数。在本发明另一实施例中,例如通过截去或舍入其小数部分把一些或所有描述符元素表示成整数形式。在本发明又一实施例中,通过比例换算把一些或所有描述符元素表示成整数形式。例如,公式 (1)-(4) 得到其中小数部分只有三个值的实数,即,.25、.50 以及 .75。因此,将这些描述符元素乘以值 4 将得到没有任何信息损失的整数表示。在本发明再一实施例中,通过任何方法把一些或所有描述符元素表示成整数形式,然后把它们量化为较短(例如 3 位或 4 位长的)值。

[0079] 因此,在本发明一个实施例中,图像 F_i 由其描述符 $V_{i(C1SCq)}^{C1} \dots V_{i(C1SC1)}^{C1} \dots V_{i(C2SCq)}^{C2} \dots V_{i(C2SC1)}^{C2} \dots V_{i(CpSCq)}^{Cp} \dots V_{i(CpSC1)}^{Cp}$ 来表示,其中, $C1$ 、 $C2$ 、... Cp 是 F_i 的第一、第二以及第 p 色彩通道, $C1SCq$ 是信道 $C1$ 的最高描述符尺度, $C1SC1$ 是信道 $C1$ 的最低描述符尺度, $C2SCq$ 是信道 $C2$ 的最高描述符尺度, $C2SC1$ 是信道 $C2$ 的最低描述符尺度,依此类推。所有这些特定尺度特定通道描述符都在多尺度(优选地,一个以上)下并在多个通道(优选地,一个以上)上捕获图像的空间敏感色彩内容和互相关信息。

[0080] 如现有技术中的普通方法那样地将图像 F_i 的所有这些描述符融合(即,连接)成 V_i 的单个描述符以与描述符 V_j 或 F_j 相比较,在比较速度方面和对 F_i 与 F_j 之间的关系的理解方面会导致非最优性能。

[0081] 本发明一个实施例的特征在于:将特定尺度特定通道描述符融合成捕获不同尺度的至少两个描述符。例如,对于在 6 个不同尺度下的特定尺度特定通道描述符,本发明一个实施例生成 6 个相同尺度下的特定尺度特定通道描述符。另一实施例生成两个面向尺度多通道描述符,一个针对低尺度,一个针对高尺度,而再一实施例生成 3 个面向尺度多通道描述符,一个针对低尺度,一个针对中间尺度,而一个针对高尺度。在本发明一优选实施例中,确实多通道上计算描述符并将它们融合成面向尺度描述符,但是一另选实施例只在单个通道上计算描述符并把它们融合成面向尺度描述符。

[0082] 图 4 示出了按照以上段落内容的本发明一个实施例。在步骤 4100 中,如前所述地提取特定尺度特定通道描述符。在步骤 4250 中,将这些描述符融合成一组面向尺度多通道描述符 $V_{i(SCQ)} \dots V_{i(SCA)}$,其中 SCA 和 SCQ 分别是最低和最高描述符尺度。在步骤 4300 中,将当前尺度 C_SC 设置为最低尺度 SCA 。在步骤 4400 中,仅将当前尺度 C_SC 描述符 $V_{i(C_SC)}$ 和 $V_{j(C_SC)}$ 用于对图像 F_i 与 F_j 之间的相似性测度 $D_{j,i(C_SC)}$ 的计算。在步骤 4500 中,将该相似性测度用于对两幅图像的相似性或不相似性进行临时判定。在步骤 4600 中,考查相似性

测度和 / 或临时判定以确定是否希望在较高尺度下进行处理。如果确定不希望进行这种处理,则过程进行到步骤 4900,在步骤 4900 处,对两幅图像的相似性或不相似性执行最终判定。否则,过程进行到步骤 4700,在步骤 4700 处,检查是否还剩有待处理的较高尺度的任何描述符。如果还有较高尺度描述符,则过程进行到步骤 4800,在步骤 4800 处,增加当前尺度 C_SC ,然后进行到步骤 4400,在步骤 4400 处,在新尺度下计算新相似性测度,依此类推。另一方面,如果没有较高尺度描述符,则过程进行到步骤 4900,在步骤 4900 处,对两幅图像的相似性或不相似性执行最终判定。

[0083] 如图 4 所示的这种面向尺度序列配置的目的是尽可能快地并以尽可能少的计算来确定两幅图像之间的相似性或不相似性。因此,与把所有可用描述符融合成单个描述符的方案相比,这种方案可以提高系统速度。

[0084] 下面更接近地考查该过程的各个步骤。

[0085] 在步骤 4400 中,可以把许多可选方法用于计算相似性测度。例如,在本发明一个实施例中,可以将由下式给出的两个描述符 V_i 与 V_j (各有 n 个元素) 之间的距离 $L1$ 计算为相似性测度:

$$[0086] \quad L1_{j,i} = \sum_{l=0}^n |V_j(l) - V_i(l)| \quad (20)$$

[0087] 这里所概述的本发明实施例涉及把所述描述符用于计算图像之间的一个或更多个相似性测度 (其包括至少两个值)。

[0088] 因此,用于计算描述符距离的公式 (20) 的 $L1$ 距离只涉及本发明一个可能实现。一另选实现把公式 (20) 替换成以下公式 (21),该公式 (21) 计算 $L2$ 距离:

$$[0089] \quad L2_{j,i} = \sqrt{\sum_{l=0}^n (V_j(l) - V_i(l))^2} \quad (21)$$

[0090] 在相关文献中存在很多种这样的距离度量。

[0091] 此外,一另选实现把公式 (20) 替换成以下公式 (22),该公式 (22) 计算加权 $L1$ 距离:

$$[0092] \quad L1_{j,i} = \sum_{l=0}^n (w(l) \cdot |V_j(l) - V_i(l)|) \quad (22)$$

[0093] 其中, $w(l)$ 是第 l 个描述符元素的加权系数。此外,也可以与公式 (21) 的距离 $L2$ 或文献中提供的任何其他可用距离量度一起使用这种加权。

[0094] 有很多种用于处理相似性测度的可选方法。在本发明一个实施例中,在步骤 4500 中将一阈函数作用于相似性测度 $D_{j,i}$ 。即,对于阈值 thr ,如果:

$$[0095] \quad D_{j,i} \geq thr \quad (23)$$

[0096] 则认为这两幅图像不相似并在步骤 4500 中作出临时判定“不相似”。据此,步骤 4600 确定不必进行其他处理,并在步骤 4900 中认定 (finalise) 该临时判定。另一方面,如果在步骤 4500 中不满足关系 (23),则在该步骤中作出临时判定“相似”。据此,步骤 4600 确定还希望进行其他处理。如果存在较高尺度下的描述符,那么过程返回步骤 4400。否则,在步骤 4900 中认定该临时判定“相似”。

[0097] 可以扩展这里所概述的本发明多个实施例以使用所述相似性测度 (其包括至少两个值) 来确定图像之间的相似性或不相似性。因此,在本发明一另选实施例中,如果:

[0098] $D_{j,i} \geq \text{thr}$ (24)

[0099] 则认为这两幅图像相似并在步骤 4500 中作出临时判定“相似”。据此,步骤 4600 确定不必进行其他处理,并在步骤 4900 中认定该临时判定。另一方面,如果在步骤 4500 中不满足关系 (24),则在该步骤中作出临时判定“不相似”。据此,步骤 4600 确定还希望进行其他处理。如果存在较高尺度下的描述符,那么过程返回步骤 4400。否则,在步骤 4900 中认定该临时判定“不相似”。

[0100] 本发明再一实施例使用两个阈值 thr1 和 thr2,其中 $\text{thr1} < \text{thr2}$ 。这样,

[0101] 如果:

[0102] $D_{j,i} \geq \text{thr1}$ (25)

[0103] 则认为这两幅图像相似并在步骤 4500 中作出临时判定“相似”。据此,步骤 4600 确定不必进行其他处理,并在步骤 4900 中认定该临时判定。另一方面,如果在步骤 4500 中不满足关系 (25),则在该步骤中作出临时判定“不相似”。然后,在步骤 4600 中,如果:

[0104] $D_{j,i} \geq \text{thr2}$ (26)

[0105] 则步骤 4600 确定不必进行其他处理,并在步骤 4900 中认定该临时判定。另一方面,步骤 4600 确定还希望进行其他处理。如果存在较高尺度下的描述符,那么过程返回步骤 4400。否则,在步骤 4900 中认定该临时判定“不相似”,或者,另选地,在步骤 4900 中生成新判定“可能”,其表示在已处理完所有描述符尺度但是无法达到确信的判定。

[0106] 在步骤 4900 中所作出的判定的语义取决于具体实现。

[0107] 在本发明一个实施例中, F_i 和 F_j 是时间有序图像。例如,它们是视频 F_z 的帧,目标是查明是否在 F_j 与 F_i 之间存在帧不连续性,如拍摄转场或照明效果等。然后,在步骤 4900 处的判定是: F_j 与 F_i 之间相似性测度是否表示这种帧不连续性。这种系统将用于分割视频。

[0108] 在本发明另一实施例中, F_j 和 F_i 是没有时间顺序的独立图像。例如, F_j 是一组图像 F_z 中的一幅图像而 F_i 是 F_z 内部或外部的另一幅图像,目标是找出 F_z 中与 F_i 相似的那些图像。然后,步骤 4900 的判定是: F_j 与 F_i 之间相似性测度是否表示图像是相似的。这种系统将用于通过查询进行图像检索或用于在数据库中对图像进行分类。

[0109] 在本发明一个实施例中,步骤 4400 的描述符 V_i 和 V_j 中的一个或两个不表示图像而表示图像组。

[0110] 例如,在本发明一个实施例中, V_j 是图像 $F_{j1}, F_{j2}, \dots, F_{jg}$ (它们形成了图像族 F_j) 的描述符 $V_{j1}, V_{j2}, \dots, V_{jg}$ 的函数 (如平均或中值)。例如, F_j 可以是表示汽车的图像族。对于图像 F_i ,确定描述符 V_i 与 V_j 之间的相似性不是揭示 F_i 是否与特定图像相似,而是揭示其是否与特定图像类或图像族相似。一组图像 F_z 可以包括任何数量个这种族,并且这种处理将揭示给定图像与哪个特定族最相配。此外,不仅可以由 V_j 表示族,而且可以通过其他统计量 (如其图像描述符的协方差阵) 表示族,该统计量使得可以在步骤 4400 中计算更复杂的相似性测度,如 Mahalanobis 距离或最大似然测度。

[0111] 在本发明的用于分割视频的另一实施例中, V_j 是帧 $F_{j1}, F_{j2}, \dots, F_{jg}$ 的描述符 $V_{j1}, V_{j2}, \dots, V_{jg}$ 的函数 (如平均或中值)。这些帧可能全都在帧 F_i 之前,或者可能全都在其之后,或者一些在其之前而一些在其之后。

[0112] 可以扩展这里所概述的本发明多个实施例以使用所述相似性测度 (其包括至少两个值) 确定图像之间的相似性或不相似性。

[0113] 因此,在本发明的用于分割视频的前述多个实施例中,在步骤 4500 中,把两个帧 F_i 与 F_j 之间的相似性测度用于检测帧不连续性。本发明一另选实施例不仅使用该相似性测度而且使用帧 F_i 和 F_j 附近的多个帧之间的相似性测度。

[0114] 例如,在本发明一个实施例中, $j = i-1$, 并且对帧不连续性的检测依赖于相似性测度 $D_{i+k-1, i+k}$, $\forall k \in [-t_1, t_2]$ 。在一优选实施例中, $t_1 = t_2 = t$ 。因此,帧 F_{i+k} 形成了以帧 F_i 为中心的时间窗口。然后,当且仅当以下两式成立时检测到 F_{i-1} 与 F_i 之间的帧不连续性:

$$[0115] \quad D_{i-1, i} \geq D_{i+k-1, i+k} \quad \forall k \in [-t, t], k \neq 0 \text{ 并且} \quad (27) \quad \frac{D_{i-1, i}}{D_{i+p-1, i+p}} \geq thr$$

$$\text{其中 } D_{i+p-1, i+p} \geq D_{i+k-1, i+k} \quad \forall k, p \in [-t, t], k, p \neq 0 \quad (28)$$

[0116] 根据公式 (27) 和 (28),在以 $D_{i-1, i}$ 为中心的大小为 $2t+1$ 的时间窗口内考虑各个相似性测度 $D_{i-1, i}$ 。然后,当且仅当以下条件成立时在帧 F_{i-1} 与 F_i 之间检测到不连续性: $D_{i-1, i}$ 是该时间窗口内的最大距离并且比该时间窗口内的第二大距离至少大 thr 倍。这种时间适应方案旨在减少由于帧之间的高运动级而产生的误检测数量。存在几个用于处理这种时间序列以检测帧不连续性的这种方案。

[0117] 在本发明一优选实施例中,用于处理面向尺度描述符的任何阈值和控制参数也是面向尺度的。

[0118] 在一另选实施例中,根据用户指令或使用该过程的自动系统的指令,在处理完最高尺度描述符之间结束对面向尺度描述符的串行处理。例如在实时在线应用中和在系统不能赶上数据速率时,可以如此操作以进一步提高系统速度。

[0119] 图 5 示出了本发明另一实施例。在步骤 5100 中,如前所述地提取特定尺度特定通道描述符。在步骤 5250 中,将这些描述符融合成一组面向尺度多通道描述符 $V_{i(SCQ)} \dots V_{i(SCA)}$, 其中 SCA 和 SCQ 分别是最低和最高描述符尺度。在步骤 5410、... 5490 中,把面向尺度描述符对 $V_{i(SCA)}$ 和 $V_{j(SCA)}$ 、... $V_{i(SCQ)}$ 和 $V_{j(SCQ)}$ 用于计算图像 F_i 与 F_j 之间的相似性测度 $D_{j, i(SCA)} \dots D_{j, i(SCQ)}$ 。因此,计算出多个相似性测度,每个描述符尺度具有一个。在步骤 5510、... 5590 中,把每个相似性测度用于对两幅图像的相似性或不相似性进行的临时判定。因此,执行多次临时判定,每个描述符尺度一次。然后,在步骤 5900 中,将该多次临时判定用于对两幅图像的相似性或不相似性进行的最终判定。

[0120] 如图 5 所示的这种面向尺度并行结构的目的是允许通过处理不同尺度下的相似性判定更加彻底地探查图像 F_i 与 F_j 之间的关系。因此,这种系统不仅可以提供最终相似性判定,而且可以提供附加信息,如“这些图像在很粗的级别上相似,但是在细节上很不相似”。与之对照的是,其中把特定尺度特定通道描述符融合成单个多尺度多通道描述符的系统只能提供两幅图像的相似性判定。

[0121] 因此,在本发明的用于从一组图像中检索出与给定图像 F_i 相似的图像的一个实施例中,如图 5 所示的方案允许以分级方式呈现检索结果。首先是在粗级上与 F_i 相似的一组图 F_A (即,根据低尺度描述符相似),然后是在更细微级上与 F_i 相似的一组图 F_B (其为 F_A 的子集),依此类推。

[0122] 在本发明的用于检测帧不连续性和视频分割的另一实施例中,图 5 所示的方案允许在不同粒度(granularity)上进行分割,即粗、中间、细等粒度。

[0123] 在图 5 中,步骤 5100 类似于图 4 的步骤 4100,并且先前针对步骤 4100 所描述的所有方法、扩展以及修改都同样在此适用。

[0124] 类似地,步骤 5250 类似于图 4 的步骤 4250,并且先前针对步骤 4250 所描述的所有方法、扩展以及修改都同样在此适用。

[0125] 类似地,各步骤 5410、... 5490 类似于图 4 的步骤 4400,并且先前针对步骤 4400 所描述的所有方法、扩展以及修改都同样在此适用。

[0126] 类似地,各步骤 5510、... 5590 类似于图 4 的步骤 4500,并且先前针对步骤 4500 所描述的所有方法、扩展以及修改都同样在此适用。

[0127] 在本发明一个实施例中,步骤 5510、... 5590 的临时判定是二元判定,例如,“1”表示“相似”,“0”表示“不相似”。

[0128] 在本发明另一实施例中,步骤 5510、... 5590 的初始判定不是二元的,而是包括至少三个值。在此情况下,术语“似然测度”比术语“初始判定”更适于描述处理步骤 5510、... 5590 的结果。例如,在本发明一个实施例中,步骤 5510、... 5590 中的每个步骤都使用三个阈值。如果相似性测度低于第一阈值,则生成表示“不相似”的相似性似然测度“0”。如果相似性测度高于第一阈值但是低于第二阈值(其高于第一阈值),则生成表示“可能不相似”的相似性似然测度“1”。如果相似性测度高于第二阈值但是低于第三阈值(其高于第二阈值),则生成表示“可能相似”的相似性似然测度“2”。如果相似性测度高于第三阈值,则生成表示“相似”的相似性似然测度“3”。在本发明另一实施例中,不通过阈值比较而通过对将相似性测度归一化到固定范围(例如,0 与 1 之间)生成似然测度。

[0129] 在把二元临时判定或似然测度表示成 $Z_{j,i(SCA)} \dots Z_{j,i(SCQ)}$ 的情况下,在本发明一个实施例中,在步骤 5900 中,当且仅当下式成立时作出 F_i 与 F_j 相似的最终判定:

$$[0130] \quad W_{SCA} \cdot Z_{j,i(SCA)} + \dots + W_{SCQ} \cdot Z_{j,i(SCQ)} \geq \text{thr}_f \quad (29)$$

[0131] 其中 thr_f 是阈值, $W_{SCA} \dots W_{SCQ}$ 是被赋予各面向尺度判定或似然测度的控制重要性的权重。如果希望,通过设置 $W_{SCA} = \dots = W_{SCQ} = 1$ 可以取消这些权重。

[0132] 在本发明另一实施例(其中似然测度包括由步骤 5510、... 5590 生成的至少三个值)中,当且仅当满足公式(29)的条件时认为这两幅图像相似,此外,没有一个或不超过给定数量个独立似然测度低于新的第二阈值。

[0133] 在本发明一优选实施例中,用于处理面向尺度描述符的任何阈值和控制参数也是面向尺度的。

[0134] 在对一组描述符(如在本实施例中描述的描述符)的并行处理中,可以使用许多这样的方案以达到明确的判定。

[0135] 本发明一个实施例的特征在于:将特定尺度特定通道描述符融合成在不同色彩通道上的至少两个描述符。例如,对于在 Y 、 C_b 以及 C_r 通道上的特定尺度特定通道描述符,本发明的一个实施例生成三个多尺度面向通道描述符,而另一实施例生成两个多尺度面向通道描述符,一个是 Y 通道的,一个是 C_b 和 C_r 通道的。在本发明一优选实施例中,确实多尺度上计算描述符并将它们融合成面向尺度描述符,但是一另选实施例只在单个尺度上计算描述符并把它们融合成面向通道描述符。

[0136] 图 6 示出了按照以上段落内容的本发明一个实施例。在步骤 6100 中,如前所述地提取特定尺度特定通道描述符。在步骤 6250 中,将这些描述符融合成一组多尺度面向通道

描述符 $V_{i,i}^{CCA}, \dots, V_{i,i}^{CCQ}$, 其中 CCA 和 CCQ 是色彩通道或色彩通道组。例如, 在本发明一个实施例中, CCA 表示 Y 通道, CCB 表示通道对 $C_b C_r$ 。在本发明另一实施例中, CCA 表示 RGB 的 R 通道, CCB 表示 G 通道, CCC 表示 B 通道。在步骤 6410、... 6490 中, 把面向通道描述符对 $V_{i,i}^{CCA}$ 和 $V_{j,i}^{CCA}, \dots, V_{j,i}^{CCQ}$ 和 $V_{j,i}^{CCQ}$ 用于计算图像 F_i 与 F_j 之间的相似性测度 $D_{j,i}^{CCA}, \dots, D_{j,i}^{CCQ}$ 。因此, 计算出多个相似性测度, 每个面向通道描述符具有一个。在步骤 6510、... 6590 中, 把每个相似性测度用于对两幅图像的相似性或不相似性进行的临时判定。因此, 执行多次临时判定, 每个面向通道描述符尺度一次。然后, 在步骤 6900 中, 将该多次临时判定用于对两幅图像的相似性或不相似性进行的最终判定。

[0137] 如图 6 所示的这种面向通道并行结构的目的是允许通过处理不同色彩通道上的相似性判定更加彻底地探查图像 F_i 与 F_j 之间的关系, 捕获关于图像的不同信息。因此, 这种系统不仅可以提供最终相似性判定, 而且可以提供附加信息, 如“这些图像在色度通道上相似, 但是在亮度信道上不相似”。与之对照的是, 其中把特定尺度特定通道描述符融合成单个多尺度多通道描述符的系统只能提供两幅图像的相似性判定。

[0138] 因此, 在本发明的用于检测帧不连续性 (如拍摄转场或照明效果等) 的一个实施例中, 如图 6 所示的系统不仅可靠地检测帧不连续性, 而且将拍摄转场与其他照明效果区分开来。例如, 如果发现两个帧在色度和亮度两方面都很不相同, 这就表示拍摄转场。另一方面, 如果发现两个帧在亮度方面很不相同但是在色度方面差别很小, 这就表示整体照明效果, 而非拍摄转场。

[0139] 因此, 在本发明的用于从一组图像中检索出与给定图像 F_i 相似的图像的一另选实施例中, 如图 6 所示的方案允许检索出与查询图像语义上相似的图像, 例如具有相同的风景或城市风光但是在一天中不同时间和在不同照明条件下拍摄的图像。

[0140] 在图 6 中, 步骤 6100 类似于图 5 的步骤 5100, 并且先前针对步骤 5100 所描述的所有方法、扩展以及修改都同样在此适用。

[0141] 类似地, 步骤 6250 类似于图 5 的步骤 5250, 并且先前针对步骤 5250 所描述的所有方法、扩展以及修改都同样在此适用。

[0142] 类似地, 步骤 6410、... 6490 类似于图 5 的步骤 5410、... 5490, 并且先前针对步骤 5410、... 5490 所描述的所有方法、扩展以及修改都同样在此适用。

[0143] 类似地, 步骤 6510、... 6590 类似于图 5 的步骤 5510、... 5590, 并且先前针对步骤 5510、... 5590 所描述的所有方法、扩展以及修改都同样在此适用。

[0144] 类似地, 步骤 6900 类似于图 5 的步骤 5900, 并且先前针对步骤 5900 所描述的所有方法、扩展以及修改都同样在此适用。

[0145] 本发明另一实施例的特征在于: 最初把特定尺度特定通道描述符分成特定尺度特定通道面向特征语义描述符, 然后将它们融合成不同特征语义的至少两个描述符。例如, 在本发明一个实施例中, 对于在 6 个不同尺度下并在 3 个不同色彩通道上的特定尺度特定通道描述符, 最初共有 18 个描述符。将这 18 个描述符分成 36 个描述符, 其中一半只包含有色彩内容元素 (如公式 (1) 的 LL), 而另一半只包含有色彩互相关元素 (如 (2)-(4) 的 HL、LH 以及 HH)。然后, 把这 36 个描述符融合成两个描述符, 一个包含有在多尺度下并在不同通道上的所有色彩内容元素, 而另一个包含有在多尺度下并在不同通道上的所有色彩互相关元素。在本发明另一实施例中, 将每个初始描述符分成两个以上描述符。例如, 将每个初

始描述符分成 4 个描述符,LL、HL、LH 以及 HH 类型的元素中的每一个元素具有一个描述符。然后将这些描述符融合成 4 个最终描述符,每个最终描述符都仅包含有在不同尺度下和不同通道上的一种类型的元素,例如,仅 LL 元素或仅 HH 元素。在本发明一优选实施例中,确实在多尺度下和多通道上计算描述符,但是一另选实施例只在单个尺度上并且 / 或者只针对单个通道计算描述符。

[0146] 图 7 示出了按照以上段落内容的本发明一个实施例。在步骤 7100 中,如前所述地提取特定尺度特定通道描述符。在步骤 7200 中,根据如上所述的特征语义拆分这些描述符。在步骤 7250 中,将所得描述符融合成一组多尺度多通道面向特征语义描述符 $V_i^{FSA}, \dots, V_i^{FSQ}$, 其中 FSA、...FSQ 表示特征类型。例如,在本发明一个实施例中,FSA 表示 LL 元素,FSB 表示 HL、LH 以及 HH 元素。在本发明另一实施例中,FSA 表示 LL 元素,FSB 表示 HL 元素,FSC 表示 LH 元素,FSD 表示 HH 元素。在步骤 7410、... 7490 中,将面向特征语义描述符对 V_i^{FSA} 和 $V_j^{FSA}, \dots, V_i^{FSQ}$ 和 V_j^{FSQ} 用于计算图像 F_i 与 F_j 之间的相似性测度 $D_{j,i}^{FSA}, \dots, D_{j,i}^{FSQ}$ 。因此,计算出多个相似性测度,每个面向特征语义描述符具有一个。在步骤 7510、... 7590 中,把每个相似性测度用于对两幅图像的相似性或不相似性进行的临时判定。因此,执行多次临时判定,每个面向特征语义描述符尺度一次。然后,在步骤 7900 中,将该多次临时判定用于对两幅图像的相似性或不相似性进行的最终判定。

[0147] 如图 7 所示的这种面向特征语义并行结构的目的是允许通过处理基于不同特征类型的相似性判定更加彻底地探查图像 F_i 与 F_j 之间的关系,捕获关于图像的不同信息。因此,这种系统不仅可以提供最终相似性判定,而且可以提供附加信息,如“这些图像在局部边沿方面相似,但是在色彩内容上不相似”。与之对照的是,其中把特定尺度特定通道描述符融合成单个多尺度多通道描述符的系统只能提供两幅图像的相似性判定。

[0148] 因此,在本发明的用于检测帧不连续性(如拍摄转场或照明效果等)的一个实施例中,如图 7 所示的系统不仅可靠地检测帧不连续性,而且将拍摄转场与其他照明效果区分开来。例如,如果发现两个帧在色彩内容和色彩互相关两方面都很不相同,这就表示拍摄转场。另一方面,如果发现两个帧在色彩内容方面很不相同但是在色彩互相关方面差别很小,这就表示整体照明效果,而非拍摄转场。

[0149] 在本发明的用于从一组图像中检索出与给定图像 F_i 相似的图像的一另选实施例中,如图 7 所示的方案允许检索出与查询图像语义上相似的图像,例如具有相同的场景但是在不同照明条件下或使用不同设备导致的不同对比级下拍摄的图像。

[0150] 在图 7 中,步骤 7100 类似于图 6 的步骤 6100,并且先前针对步骤 6100 所描述的所有方法、扩展以及修改都同样在此适用。

[0151] 在图 7 中,步骤 7200 类似于图 6 的步骤 6200,并且先前针对步骤 6200 所描述的所有方法、扩展以及修改都同样在此适用。

[0152] 类似地,步骤 7250 类似于图 6 的步骤 6250,并且先前针对步骤 6250 所描述的所有方法、扩展以及修改都同样在此适用。

[0153] 类似地,步骤 7410、... 7490 类似于图 6 的步骤 6410、... 6490,并且先前针对步骤 6410、... 6490 所描述的所有方法、扩展以及修改都同样在此适用。

[0154] 类似地,步骤 7510、... 7590 类似于图 6 的步骤 6510、... 6590,并且先前针对步骤 6510、... 6590 所描述的所有方法、扩展以及修改都同样在此适用。

[0155] 类似地,步骤 7900 类似于图 6 的步骤 6900,并且先前针对步骤 6900 所描述的所有方法、扩展以及修改都同样在此适用。

[0156] 本发明另一实施例的特征在于:最初把特定尺度特定通道描述符分成特定尺度特定通道面向区域描述符,然后将它们融合成在图像中覆盖不同区域的至少两个描述符。例如,在本发明一个实施例中,对于在 5 个不同尺度 64×64 、 32×32 、... 4×4 下并在 3 个不同色彩通道上的特定尺度特定通道描述符,最初共有 15 个描述符。将这 15 个描述符分成 4 组 15 个描述符,每个组都只包含与图像的特定象限对应的那些元素,即,一组用于左上象限,一组用于右上象限,一组用于左下象限,一组用于右下象限。然后把这些描述符融合成 4 个描述符,一个包含有图像的左上象限的在多尺度下并在不同通道上的所有描述符元素,而另一个包含有图像的右上象限,等等。在本发明另一实施例中,与图像的垂直片段对应地形成多个面向区域描述符。在本发明还一实施例中,与图像的水平片段对应地形成多个面向区域描述符。在本发明一优选实施例中,确实多尺度下和多通道上计算描述符,但是另选实施例可以只在单个尺度上并且 / 或者只针对单个通道计算描述符。

[0157] 图 8 示出了按照以上段落内容的本发明一个实施例。在步骤 8100 中,如前所述地提取特定尺度特定通道描述符。在步骤 8200 中,根据如上所述的图像区域拆分这些描述符。在步骤 8250 中,将所得描述符融合成一组多尺度多通道面向区域描述符 $V_i^{RSA} \dots V_i^{RSQ}$, 其中 REA、... REQ 表示图像区域。例如,在本发明一个实施例中,REA 表示图像的左上象限,REB 表示图像的右上象限,REC 表示图像的左下象限,RED 表示图像的右下象限。在步骤 8410、... 8490 中,将面向区域描述符对 V_i^{REA} 和 V_j^{REA} 、... V_i^{REQ} 和 V_j^{REQ} 用于计算图像 F_i 与 F_j 之间的相似性测度 $D_{j,i}^{REA} \dots D_{j,i}^{REQ}$ 。因此,计算出多个相似性测度,每个面向区域描述符具有一个相似性测度。在步骤 8510、... 8590 中,把每个相似性测度用于对两幅图像的相似性或不相似性进行的临时判定。因此,执行多次临时判定,每个面向区域描述符尺度一次。然后,在步骤 8900 中,将该多次临时判定用于对两幅图像的相似性或不相似性进行的最终判定。

[0158] 如图 8 所示的这种面向区域并行结构的目的是允许通过处理由于图像的不同区域引起的相似性判定更加彻底地探查图像 F_i 与 F_j 之间的关系。因此,这种系统可以提供最终相似性判定,提供类似“这些图像在某些区域中很相似,但是在其他区域中不相似”的信息,还可以识别不同图像的对应区域之间的关系。与之对照的是,其中把特定尺度特定通道描述符融合成单个多尺度多通道描述符的系统只能提供两幅图像的相似性判定。

[0159] 因此,在本发明的用于从一组图像中检索出与给定图像 F_i 相似的图像的一个实施例中,如图 8 所示的方案允许检索出与查询图像部分地相似的图像,例如这样的图像,即,其具有相同的场景但是带有由于人突然走入相机光场并遮挡了该场景的其余部分所造成的阻挡。

[0160] 类似地,在本发明的用于检测帧不连续性的一个实施例中,如图 8 所示的系统不仅可靠地检测帧不连续性,而且将拍摄转场与其他内摄干扰(比如,人员迅速靠近相机并短暂地遮挡了视场)区分开来。

[0161] 在图 8 中,步骤 8100 类似于图 7 的步骤 7100,并且先前针对步骤 7100 所描述的所有方法、扩展以及修改都同样在此适用。

[0162] 类似地,步骤 8200 类似于图 7 的步骤 7200,并且先前针对步骤 7200 所描述的所有

方法、扩展以及修改都同样在此适用。

[0163] 类似地,步骤 8250 类似于图 7 的步骤 7250,并且先前针对步骤 7250 所描述的所有方法、扩展以及修改都同样在此适用。

[0164] 类似地,步骤 8410、... 8490 类似于图 7 的步骤 7410、... 7490,并且先前针对步骤 7410、... 7490 所描述的所有方法、扩展以及修改都同样在此适用。

[0165] 类似地,步骤 8510、... 8590 类似于图 7 的步骤 7510、... 7590,并且先前针对步骤 7510、... 7590 所描述的所有方法、扩展以及修改都同样在此适用。

[0166] 类似地,步骤 8900 类似于图 7 的步骤 7900,并且先前针对步骤 7900 所描述的所有方法、扩展以及修改都同样在此适用。

[0167] 应当清楚地理解,在本发明优选实施例中,在许多尺度下和在多个通道上提取描述符。然而,在本发明另选实施例中不必如此。例如,为了实现面向尺度框架,可以针对单个色彩通道并且 / 或者针对单个图像区域提取描述符。类似地,为了实现面向通道框架,可以针对单个尺度并且 / 或者针对单个图像区域提取描述符。类似地,为了实现面向特征语义框架,可以针对单个尺度并且 / 或者针对单个通道并且 / 或者针对单个图像区域提取描述符。类似地,为了实现面向区域框架,可以针对单个尺度并且 / 或者针对单个通道提取描述符。

[0168] 应当清楚地理解,这里所描述的串行和并行结构并非仅限于特定描述符拆分或面向性。仅以例示性的目的将特定结构应用于特定描述符面向。例如,前述实施例考虑了仅用于面向尺度描述符的串行结构。本发明另选实施例使用其他面向性的描述符的串行结构。例如,本发明一另选实施例按串行结构处理面向通道描述符,以快速过滤掉在多个通道中的一个通道中与给定图像不相似的图像,从而不必处理其他通道。本发明又一另选实施例根据与以上原理相同的原理按串行结构处理面向特征语义描述符。本发明再一另选实施例根据与以上原理相同的原理按串行结构处理面向区域描述符。

[0169] 应当清楚地理解,系统可以基于多种描述符面向性。例如,在本发明一个实施例中,在对特定尺度特定通道描述符进行了初始计算之后,形成面向尺度面向特征语义描述符。例如,形成 4 个描述符,即,一个描述符用于色彩内容元素的低尺度,一个描述符用于色彩内容元素的高尺度,一个描述符用于色彩互相关元素的低尺度,一个描述符用于色彩互相关元素的高尺度。在本发明一个实施例中,如前所述地按并行结构处理这种描述符。在本发明另一实施例中,如前所述地按串行结构处理这种描述符。在本发明再一实施例中,使用如图 9 所示的混合结构。其中,SCA 和 SCQ 分别是最低和最高描述符尺度,FSA、... FSQ 表示特征类型。

[0170] 也可以将该混合结构用于单个描述符面向性。例如,一个另选实施例基于串行结构,通过该串行结构,在每个循环中,在嵌入并行结构中不处理单个面向尺度描述符,而是处理小数量个面向尺度描述符。因此,还应清楚地理解,无论本发明的不同实施例是基于单个描述符面向性还是基于多个描述符面向性,它们都可以基于串行结构或并行结构或混合结构。应当清楚地理解,出于例示性目的,图 9 仅示出了一个可能的混合结构,其中将并行结构嵌入串行结构。另选实施例使用其他混合结构,例如其中把一个或更多个串行结构嵌入并行结构的结构。

[0171] 如图 9 所示,最初按起始粗尺度处理图像,如果不能达到确信的判定,然后在下一

较细尺度下进行处理,依此类推,直到达到确信的判定或者用完了可用 / 允许的尺度为止。

[0172] 在一优选实施例中,对于每个尺度,每个图像具有 4 个描述符,并行处理这些描述符并且将其相应的结果组合起来以给出该尺度的判定。这类似于图 9,但是所使用的描述符是如下描述符:(i) 根据亮度提取的色彩内容;(ii) 根据亮度提取的色彩互相关;(iii) 根据色度提取的色彩内容;以及(iv) 根据色度提取的色彩互相关。

[0173] 在一另选实施例中,仅将两个色彩内容描述符(一个是亮度的,一个是色度的)用于起始尺度。对于所有其他尺度,仅使用两个色彩互相关描述符。

[0174] 此外,在前述实施例中,对于对帧不连续性的检测,考虑对所有连续帧对的处理。另选实施例可以更稀疏地处理帧。一另选实施例仅处理每个第四帧,而再一实施例仅处理每个第 16 帧。这种改进进一步提高了视频分割系统的速度。

[0175] 此外,上述描述涉及对未压缩数字视频(其中由其像素值表示每个帧)的处理。本发明也可以直接应用于 MPEG 压缩视频而不必进行解压。

[0176] 对于本领域的技术人员来说,MPEG 视频的特性是公知的,在书 Watkinson, J., “The MPEG Handbook :MPEG-1, MPEG-2, MPEG-4”, Focal Press, Oxford 2001, ISBN :0-240-51656-7 中对其进行了专业描述。

[0177] 在本发明的用于 MPEG 压缩视频的一个实施例中,仅处理视频的 I 帧。更具体来说,仅处理 I 帧的 DC 表示,即,对于 8×8 像素的各块,仅采用其 DCT 变换的 DC 分量。实际上,在通过 8×8 块平均对其进行压缩之前, I 帧的该 DC 表示仅仅是原始帧的再采样版。由于不必进行解压就可以直接从压缩视频流取回 DC 分量,显然可以把本发明实施例直接应用于对 MPEG 压缩视频的 I 帧的处理。

[0178] 例如可以在带有合适的软件和 / 或硬件修改的计算机系统中实现本发明。例如,可以通过具有以下装置的计算机或类似设备实现本发明:控制或处理装置,如处理器或控制装置;包括图像存储装置的数据存储装置,如存储器、磁存储器、CD、DVD 等;数据输出装置,如显示器或监视器或打印机;数据输入装置,如键盘;以及图像输入装置,如扫描仪;或这些组件与附加组件的任何组合。可以按软件和 / 或硬件形式或者按特定应用设备或特定应用模块(如集成电路)提供本发明的多个方面。在可以从其他组件(例如,在因特网上)远程提供根据本发明实施例的设备中的系统组件。

[0179] 图 10 示出了具有计算机系统形成的合适的设备的框图,其中,控制单元 10100 连接到图像存储装置 10200、描述符存储装置 10300 以及显示器 10400。该图像存储装置 10200 存储有与待根据前述实施例表示或比较的图像对应的数据。该数据可以是与单幅图像、多幅图像、图像族、图像序列或整个图像数据库对应的数据。描述符存储装置 10300(其可以是数据库)用于存储所得到的所有描述符,包括由控制单元 10100 计算出的子描述符。在显示器 10400 上把控制单元的操作结果呈现给用户。

[0180] 术语“尺度”与“分辨率”在这里是可互换使用的。图像的区域可以意指整幅图像或图像的一部分。

[0181] 术语“描述符”表示对图像或图像的一部分所获得的表示,可以将其用于判定进行过程。术语“子描述符”表示对图像或图像的一部分所获得的表示,其为一描述符相关联。然而,本领域的技术人员会明白,这两个术语的区别只是一种方便,并且在所描述的方法中每个术语都表示特定含意。特定实施例的各个子描述符可以充当不同实施例中的描述符,

反之亦然。因此,在具体实施例中仅使用术语“描述符”。

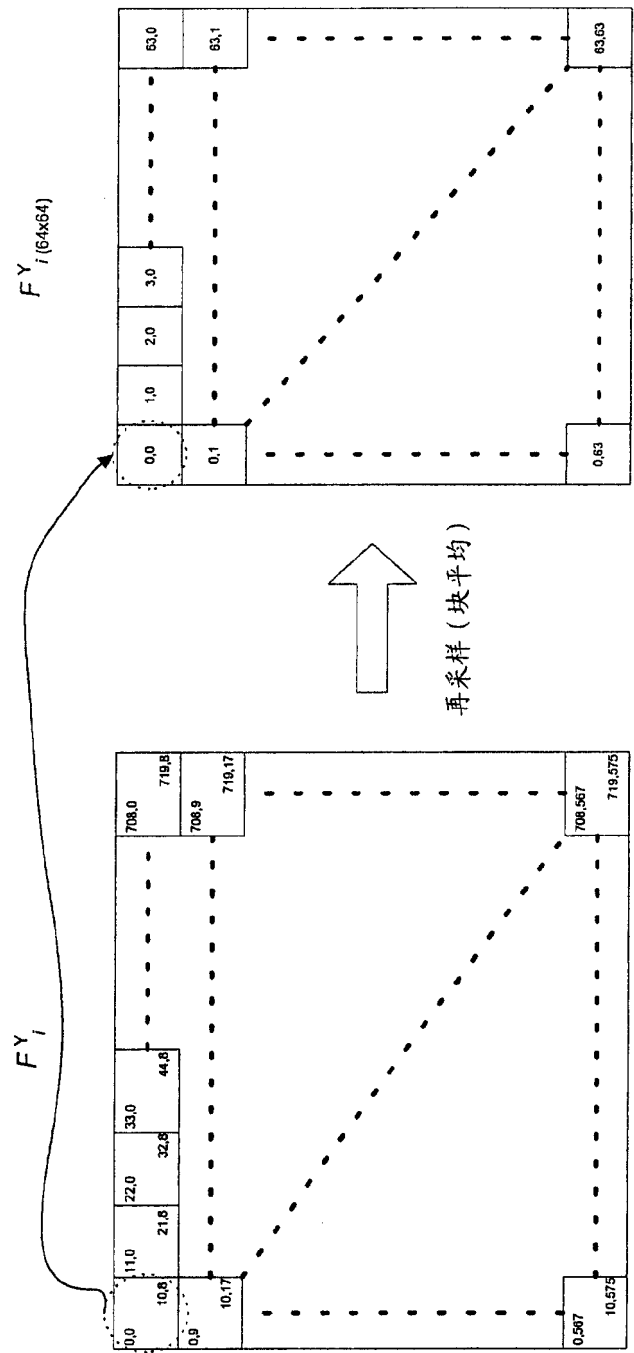
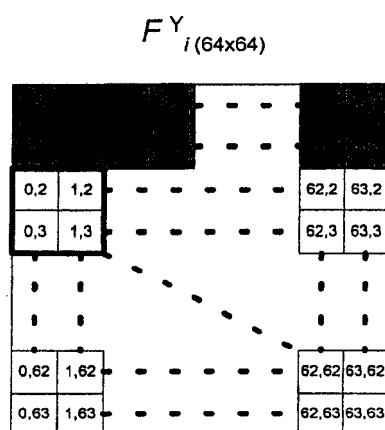
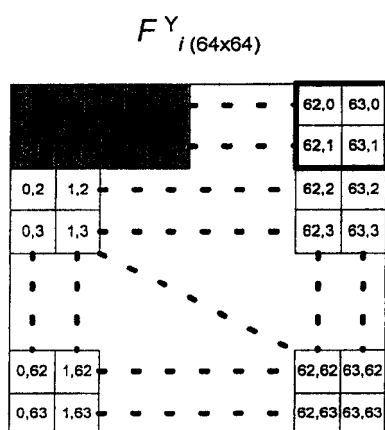
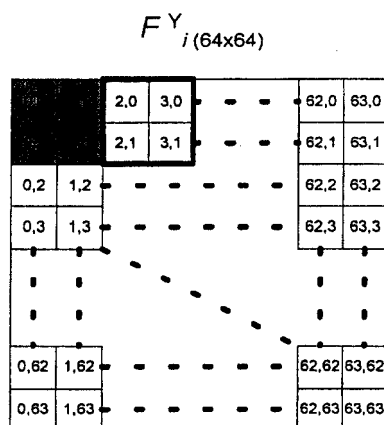
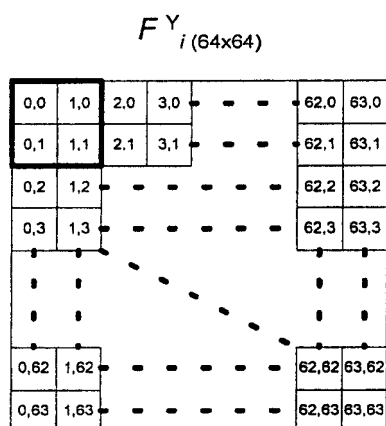


图 1



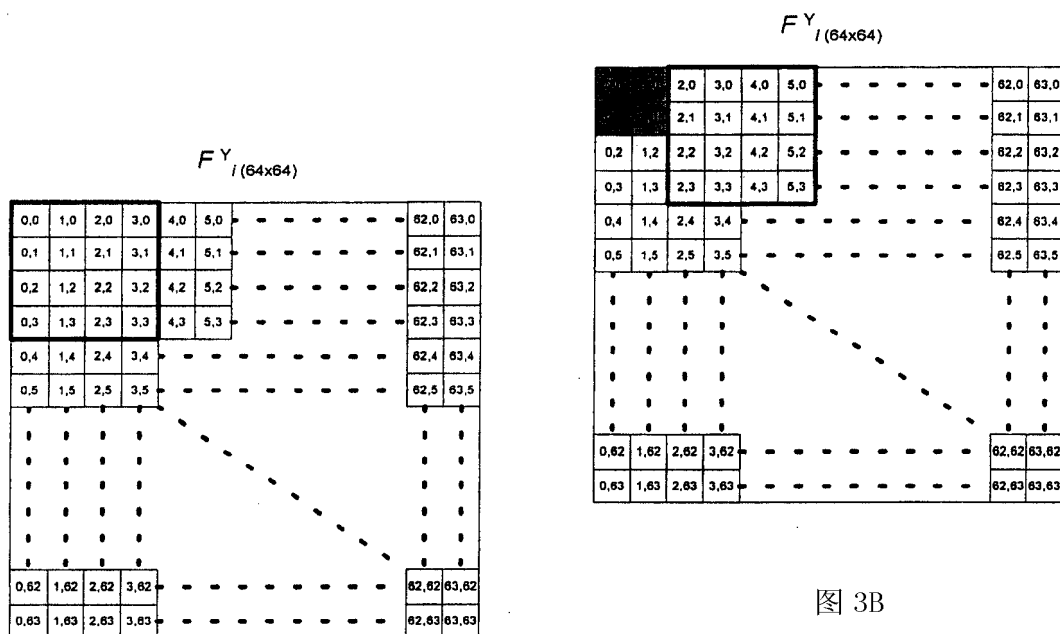


图 3A

图 3B

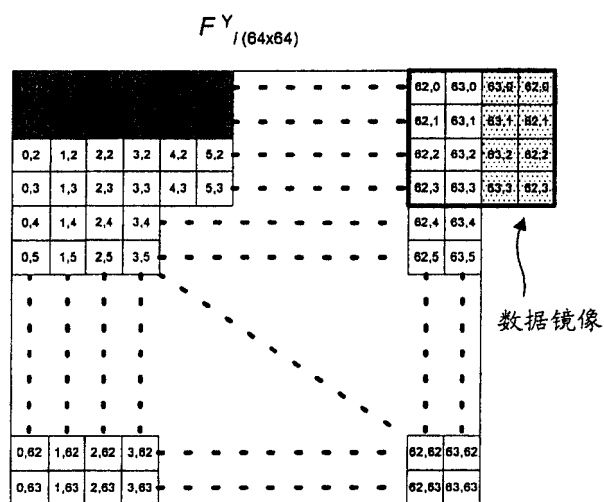


图 3C

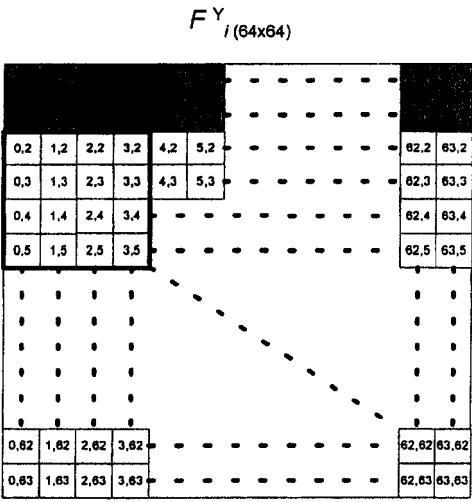


图 3D

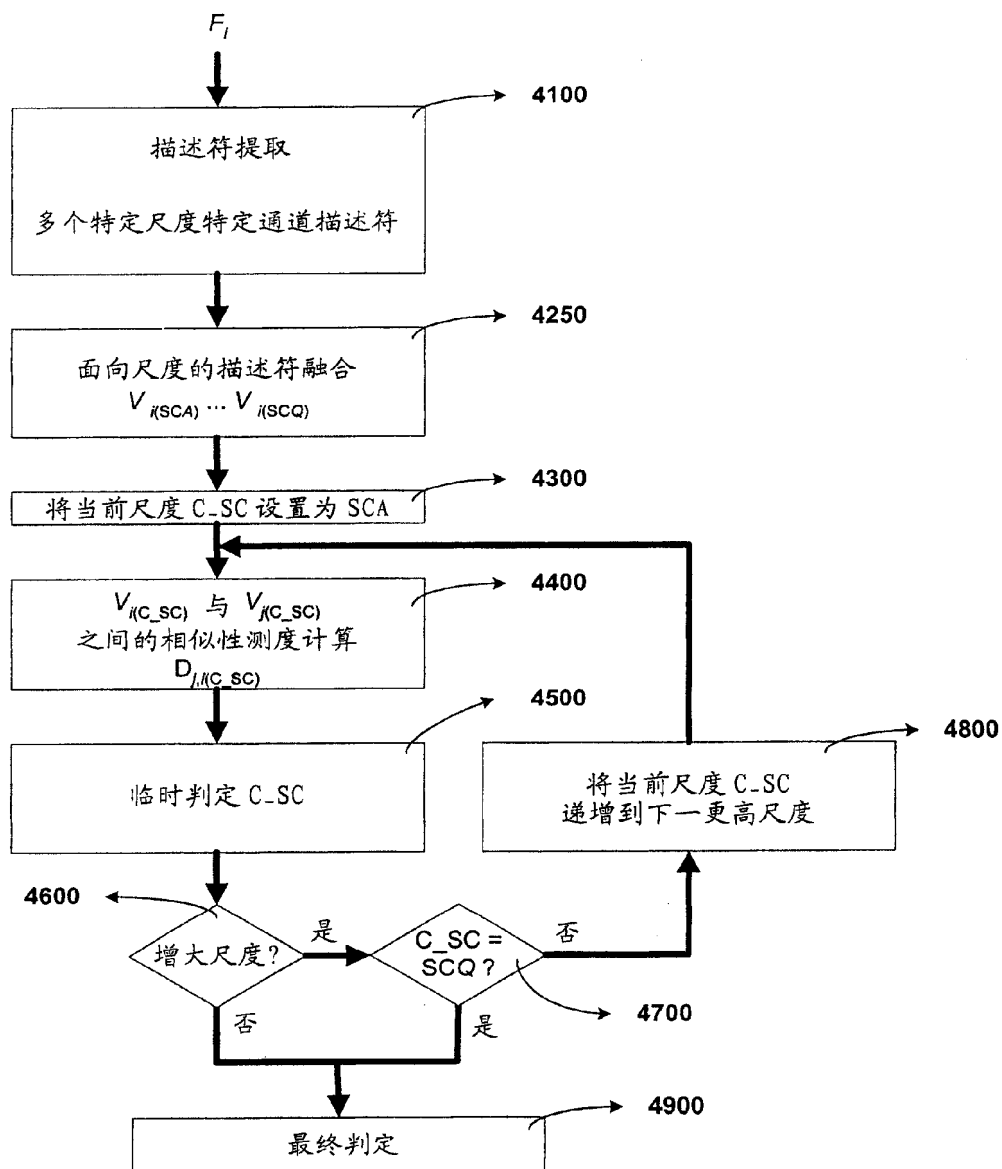


图 4

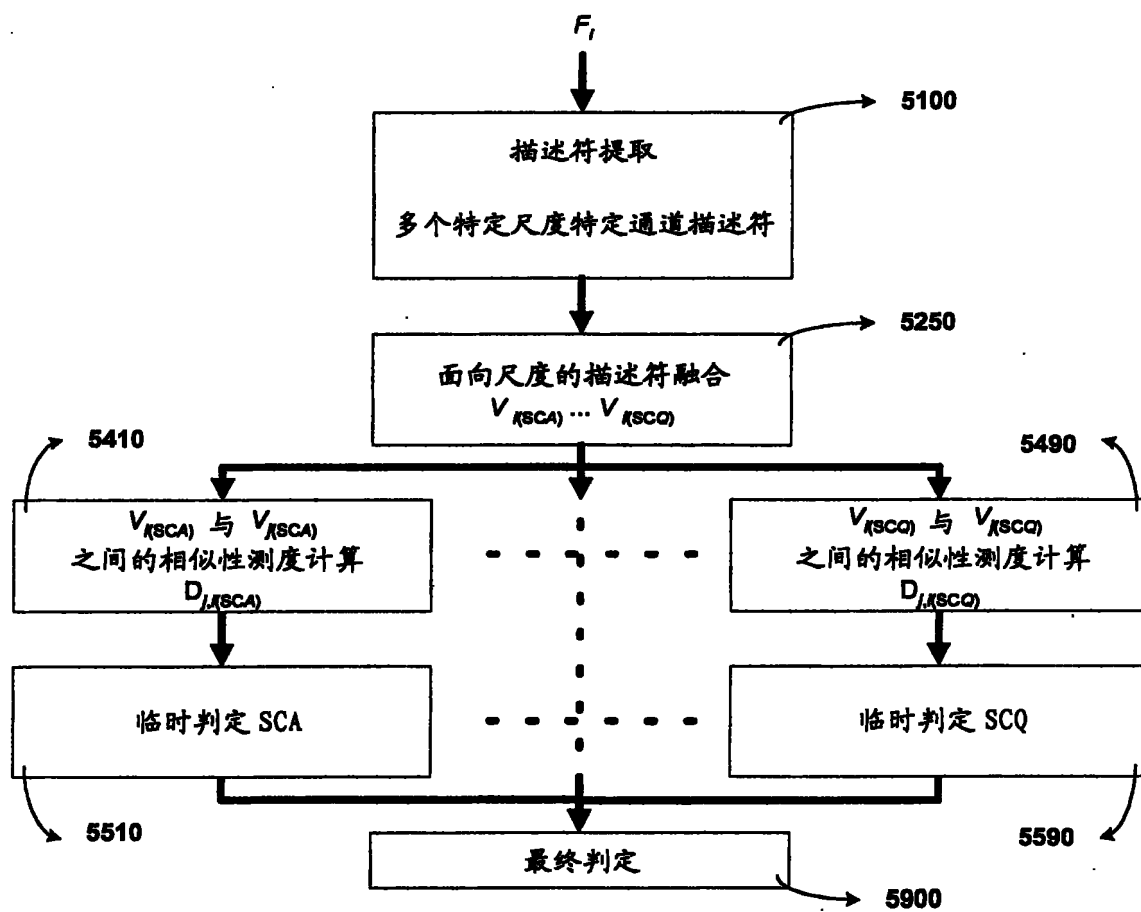


图 5

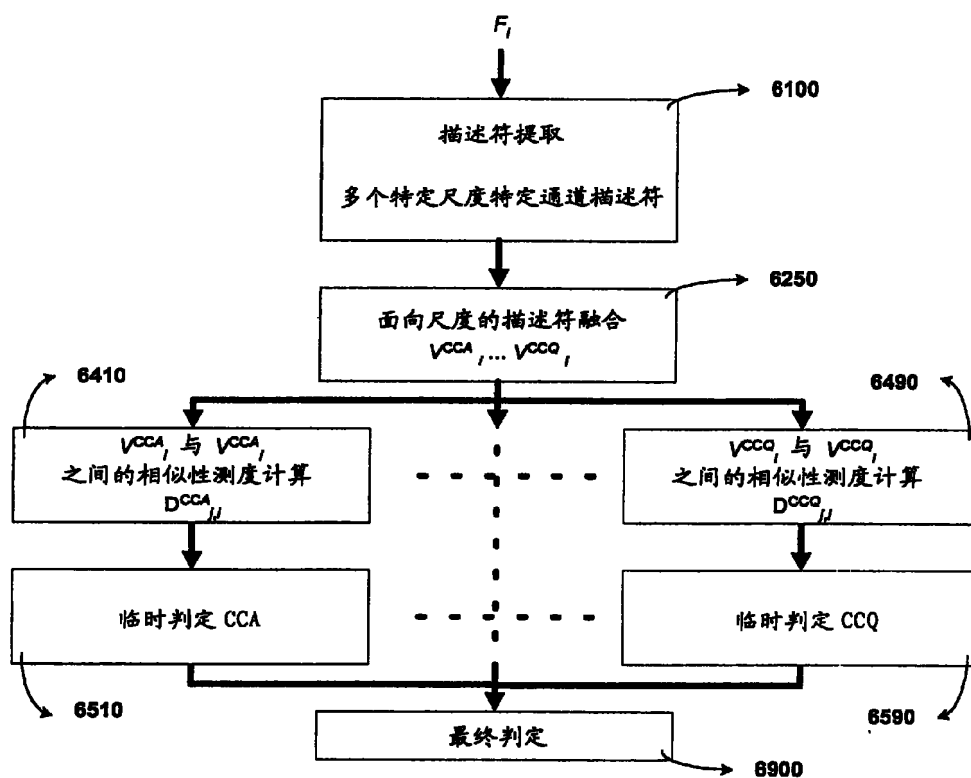


图 6

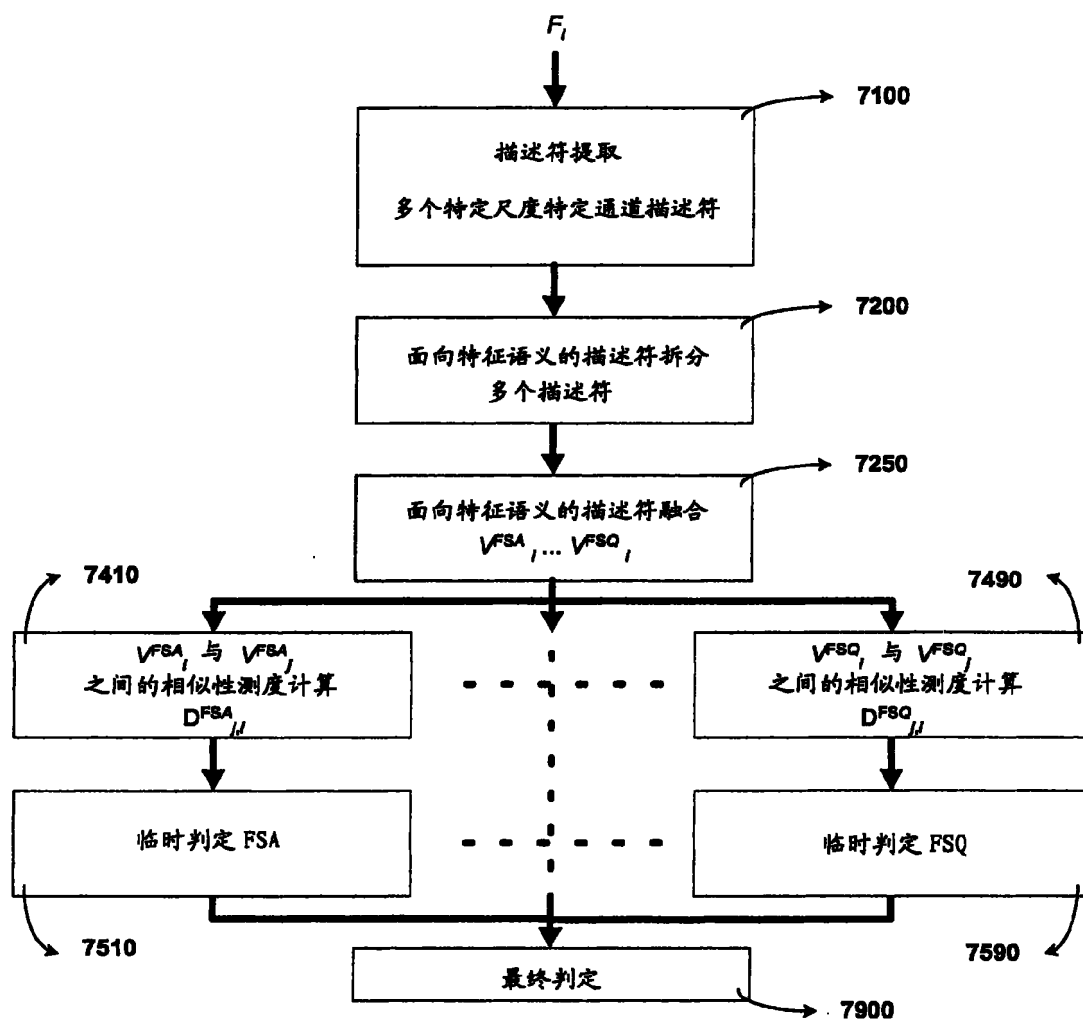


图 7

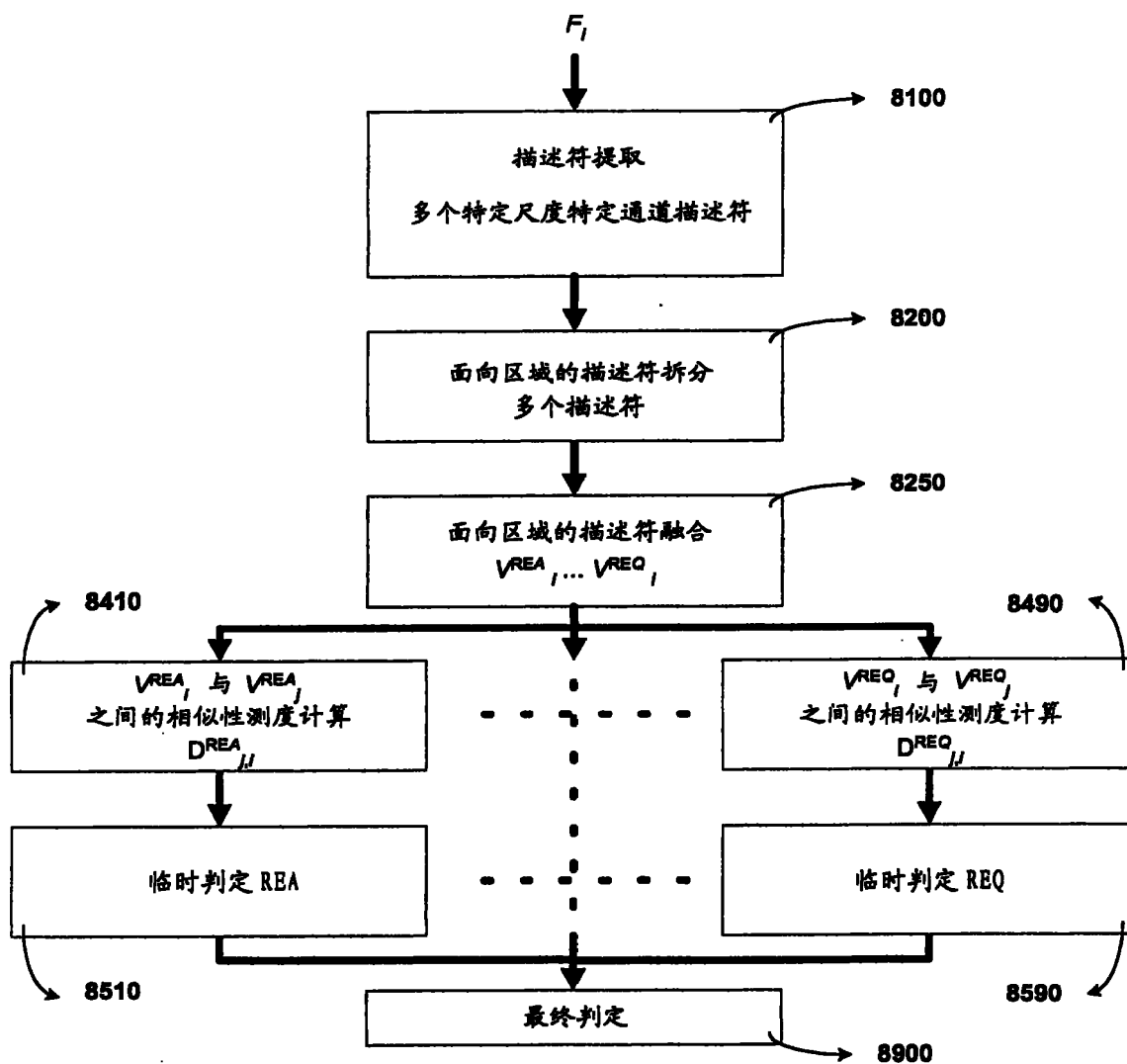


图 8

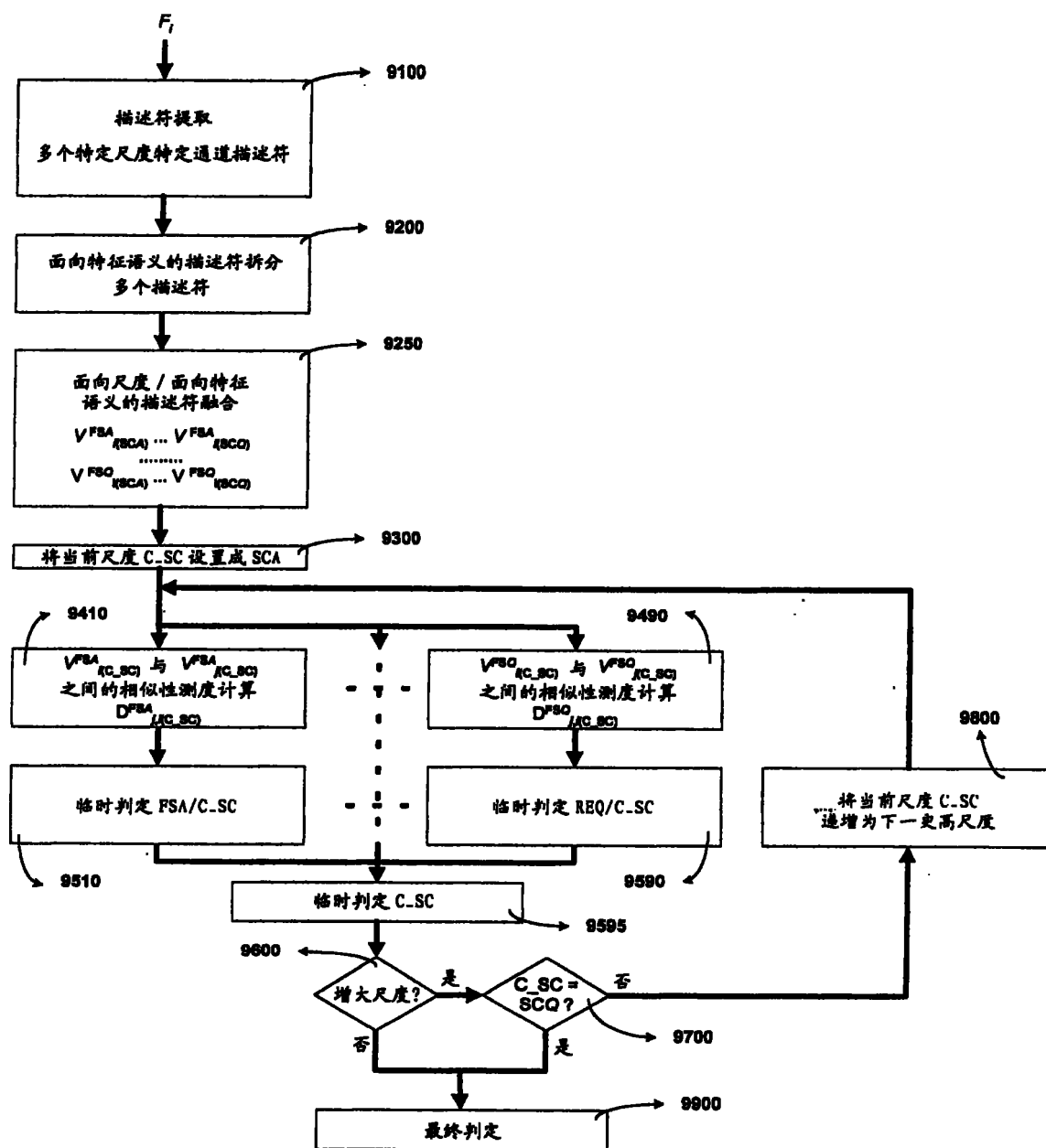


图 9

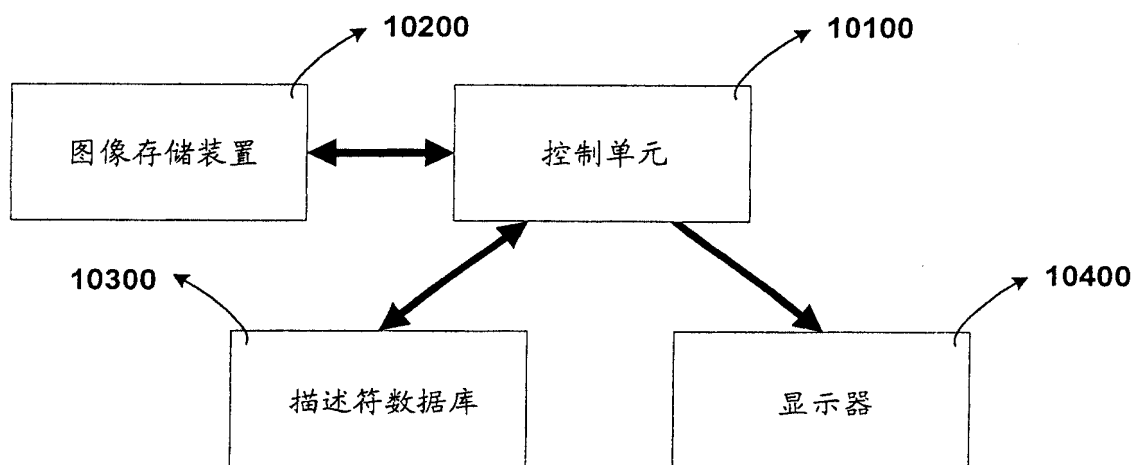


图 10