



(12)发明专利

(10)授权公告号 CN 103051891 B

(45)授权公告日 2018.08.28

(21)申请号 201210387202.7

G06T 7/11(2017.01)

(22)申请日 2012.10.12

(56)对比文件

(65)同一申请的已公布的文献号

申请公布号 CN 103051891 A

CN 101990100 A, 2011.03.23,

US 2010124274 A1, 2010.05.20,

US 2004088726 A1, 2004.05.06,

CN 101218605 A, 2008.07.09,

(43)申请公布日 2013.04.17

(30)优先权数据

11306322.6 2011.10.12 EP

AGARWAL G.A fast algorithm to find the region-of-interest in the compressed mpeg domain.《Proceedings of the 2003 international conference on multimedia and expo》.2003,第2卷全文.

(73)专利权人 汤姆森特许公司

地址 法国伊西莱穆利诺

Aniruddha Sinha.Region-Of-Interest based compressed domain video transcoding scheme.《IEEE International Conference on Montreal》.2004,第3卷正文第2,3节;附图1,2.

(72)发明人 F.厄班 C.查玛雷特 C.切万斯

(74)专利代理机构 北京市柳沈律师事务所

11105

代理人 吕晓章

审查员 张洪沛

(51)Int.Cl.

H04N 19/85(2014.01)

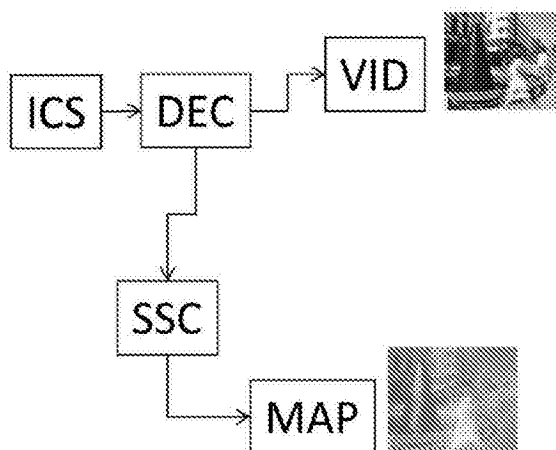
权利要求书2页 说明书4页 附图6页

(54)发明名称

确定数据流内分块预测编码的视频帧的块的显著值的方法和装置

(57)摘要

本发明涉及数据流内分块预测编码的视频的显著性确定的领域。提出了一种方法,该方法包括使用用于确定块的变换残差的编码成本以及使用所确定的编码成本来确定显著性映射图的处理手段。变换的块残差的编码成本取决于块中描述内容的生动性以及块被预测的程度有多好,因此是很好的显著性指示。



1. 一种用于确定数据流内分块预测编码的视频帧的块的显著值的方法,所述数据流包括所述视频帧的块的编码残差,所述方法包括使用处理部件来在对所述数据流进行解析时,在所述数据流中,确定与块相关的至少一个编码残差的编码成本,无需对所述与块相关的至少一个编码残差进行解码;并且使用所确定的编码成本确定所述显著值。

2. 如权利要求1所述的方法,其中当对所述块进行帧内预测编码时,所述与块相关的至少一个编码残差是所述数据流中的块的编码残差,并且所述确定包括通过使用rho-domain模型确定所述编码成本。

3. 如权利要求1所述的方法,其中当对所述块进行帧间预测或双向预测编码时,所述与块相关的至少一个编码残差是用于所述数据流中的所述块的帧间预测或双向预测的一个或多个参考块的至少一个编码残差。

4. 如权利要求1所述的方法,其中使用处理部件包括,利用所述块的大小来加权所确定的所述与块相关的至少一个编码残差的编码成本,以便当块被分解成较小的块时增加所述显著值。

5. 如权利要求1所述的方法,包括当所述块被帧间预测编码或双向预测编码时,使用所述块的至少一个运动矢量的编码成本确定所述显著值。

6. 如权利要求1所述的方法,其中使用处理部件包括规格化所确定的所述与块相关的至少一个编码残差的编码成本并使用所述规格化的编码成本确定所述显著值。

7. 如权利要求1所述的方法,其中使用处理部件包括当所述块被帧间预测编码或双向预测编码时,确定所述块是否以直接/跳跃模式被编码,并且其中在所述块是以直接/跳跃模式编码的情形下,使用衰减值来确定所述显著值。

8. 一种用于确定数据流内分块预测编码的视频帧的块的显著值的装置,所述数据流包括所述视频帧的块的编码残差,所述装置包括处理部件,所述处理部件适配于:

在对所述数据流进行解析时,在所述数据流中,确定与块相关的至少一个编码残差的编码成本,无需对所述与块相关的至少一个编码残差进行解码;并且使用所确定的编码成本确定所述显著值。

9. 如权利要求8所述的装置,其中当对所述块进行帧内预测编码时,所述与块相关的至少一个编码残差是所述数据流中的块的编码残差,并且所述确定包括通过使用rho-domain模型确定所述编码成本。

10. 如权利要求8所述的装置,其中当对所述块进行帧间预测或双向预测编码时,所述与块相关的至少一个编码残差是用于所述数据流中的所述块的帧间预测或双向预测的一个或多个参考块的至少一个编码残差。

11. 如权利要求8所述的装置,其中所述处理部件适配于利用所述块的大小来加权所确定的所述与块相关的至少一个编码残差的编码成本,以便当块被分解成较小的块时增加所述显著值。

12. 如权利要求8所述的装置,其中所述处理部件适配于当所述块被帧间预测编码或双向预测编码时,使用所述块的至少一个运动矢量的编码成本来确定所述显著值。

13. 如权利要求8所述的装置,其中所述处理部件适配于规格化所确定的所述与块相关的至少一个编码残差的编码成本,并且适配于使用规格化的编码成本来确定所述显著值。

14. 如权利要求8所述的装置,其中所述处理部件适配于当所述块被帧间预测编码或双

向预测编码时,确定所述块是否以直接/跳跃模式被编码,并且其中在所述块是以直接/跳跃模式被编码的情形下,还使用衰减值来确定所述显著值。

确定数据流内分块预测编码的视频帧的块的显著值的方法和装置

[0001] 优先权申请的交叉引用

[0002] 本申请要求于2011年10月12日提交的标题为“Saliency Value Determination of Predictively Encoded Video Streams”的欧洲专利申请第11306322.6号的优先权。

技术领域

[0003] 本发明涉及确定视频显著性(saliency)的领域。

背景技术

[0004] 在视频图像帧中检测更感兴趣的位置或引人注意的特征,也被称作显著性特征,在现实世界中许多应用。例如,它可以被应用于诸如导航协助、机器人控制、监视系统、对象检测和辨识,以及场景理解之类的计算机视觉任务中。这些预测还可以应用到其它的领域,包括广告设计、图像和视频压缩、图像和视频重用(repurposing)、图像数据库查询和动画凝视(gaze animation)。

[0005] 某些现有技术的视觉注意计算模型根据源数据的低级特征,诸如色彩、亮度、对比度、方向、运动,以及输入图像或视频信号的其它的统计分析来计算显著映射图。

[0006] 例如,Bruce、NDB和Tsotsos,JK在“Saliency based on information maximization”(Advances in neural information processing systems.p.155-162, 2006)中提出了一种最大化从场景中取样的信息的自底向上的外显注意(overt attention)模型。Itti L.、Koch C.和Niebur E.在“Model of saliency-based visual attention for rapid scene analysis”(IEEE Trans Pattern Anal Mach Intell.20(11):1254-9,1998)中展示了一种受到早期灵长类动物(primate)视觉系统行为和神经结构启发的视觉注意系统。该系统通过以高效计算的方式快速选择将被详细分析的显著位置而分解场景理解的复杂问题。

[0007] Fabrice U.等人在文献“Medium Spatial Frequencies,a Strong Predictor of Saliency”(Cognitive Computation.Volume 3,Number 1,37-47,2011)中发现中频完全允许对注意进行最佳预测,发现在人造街道场景中使用中到高频并且在自然景观场景中使用低到中频时,固定位置更具有可预测性。

发明内容

[0008] 发明人认识到现有技术的用于压缩编码的视频材料的显著性确定方法和装置需要解码这些材料,尽管这些材料通常基于空间变换、空间和时间预测以及运动信息以下面的方式而被压缩:保留感兴趣的位置的显著特征和信息,并且因此已经包含在解码中丢失的某些显著性信息。

[0009] 因此,发明人提出从压缩视频中提取显著性信息以得出较低计算成本的显著性模型。计算成本的降低基于编码引起的可用数据的重用。

[0010] 也就是说,发明人提出了一种用于确定数据流内被分块预测编码的视频帧的块的显著值的、根据权利要求1的方法和根据权利要求2的装置。所述方法包括使用用于确定块的变换残差的编码成本以及使用所确定的编码成本用于确定显著值的处理手段。

[0011] 变换的块残差的编码成本取决于块中描述内容的生动性以及块被预测的程度有多好。因此,编码成本是很好的显著性指示。

[0012] 在一个实施例中,对块进行帧内预测编码,并且确定编码成本包括确定使用rho-domain模型。

[0013] 在另一个实施例中,对块进行帧间预测编码,并且确定编码成本包括确定用于所述块的帧间预测的参考块的变换残差的编码成本。

[0014] 在另一个实施例中,使用块的大小来加权所确定的参考块的编码成本。

[0015] 在另一个实施例中,块的运动矢量的编码成本被进一步用于确定显著值。

[0016] 在另一个实施例中,所确定的编码成本被规格化,规格化的编码成本被用于确定显著值。

[0017] 在块以直接/跳跃模式被编码的情况下,衰减值可以被进一步用于确定显著值。

[0018] 更多的有利的实施例的特征在从属权利要求中指定。

附图说明

[0019] 附图图示了本发明示例性实施例,在下面的描述中,对这些实施例进行了更详细地说明。对示例性实施例的说明仅为了阐释本发明,而不是为了限制本发明的公开内容或权利要求限定的范围。

[0020] 在附图中:

[0021] 图1描述了使用现有技术推导显著映射图的示例性流程图;

[0022] 图2描述了通过从压缩的视频流中推导空间显著映射图而从该流中推导显著映射图的第一实施例的示例性流程图;

[0023] 图3描述了通过从压缩的视频流中推导时间显著映射图而从该流中推导显著映射图的第二实施例的示例性流程图;

[0024] 图4描述了通过从压缩的视频流中推导空间显著映射图和时间显著映射图并融合推导的映射图而从该流中推导显著映射图的第三实施例的示例性流程图;

[0025] 图5描述了从压缩的视频流中推导空间显著映射图的示例性流程图;

[0026] 图6描述了从压缩的视频流中推导时间显著映射图的示例性流程图;以及

[0027] 图7描述了融合空间显著映射图和时间显著映射图的示例性流程图。

具体实施方式

[0028] 本发明可以在任何包括相应地适用的处理装置的电子装置上实现。本发明在需要基于显著性的应用的低功率装置上尤其有用但不限于此。例如,本发明可以在机顶盒、图形输入板(tablet)、网关、电视机、移动视频电话、个人计算机、数字摄像机或车载娱乐系统中实现。

[0029] 当前发明公开并利用了下述事实:编码的流已经包含可以被用来以非常小的额外计算成本推导显著映射图的信息。该信息可以通过视频解码器在全解码过程中提取。或者,

可以实现仅对视频流进行解析而不对其完全解码的部分解码器。

[0030] 在图2所示的第一示例性实施例中,计算显著映射图MAP只包括空间显著映射图计算SSC。

[0031] 在图3所示的第二示例性实施例中,计算显著映射图MAP只包括时间显著映射图计算TSC。

[0032] 在图4所示的第三示例性实施例中,计算显著映射图MAP包括空间显著映射图计算SSC、时间显著映射图计算TSC以及融合FUS计算的空间显著映射图和计算的时间显著映射图。

[0033] 第一、第二和第三示例性实施例中计算的空间和/或时间显著映射图是根据输入的压缩流ICS中的可用信息计算的,并未对在输入的压缩流ICS中编码的视频VID完全解码DEC。

[0034] 本发明不限于特定编码方案。可以使用任何预测编码方案,例如H.264/MPEG-4AVC,MPEG-2或其它的,来压缩输入的压缩流ICS。

[0035] 在不同的示例性实施例中,空间显著映射图计算SCC是基于编码成本估计的。Z.He在文献“p-domain rate-distortion analysis and rate control for visual coding and communication”(Santa Barbara,PhD-Thesis,University of California,2001)中描述了块变换的非零变换系数的数目与该块的编码成本成比例。图5示例性示出的空间显著映射图计算SCC利用了这个事实并为帧内编码的块分配了使用这些块的编码成本确定的显著值,编码成本使用rho-domain模型确定,如He所述。

[0036] 由于在大多数时候只有相对显著性是重要的,因此显著映射图可以被规格化。

[0037] 除了编码成本,块大小也可以被用来确定显著值。较小的块通常与对象的边缘相关,从而也与感兴趣区域相关。宏块的成本映射图随着被分解成较小的块的数目而增加。例如,在子块分解的情形中每个块的成本值都会加倍。

[0038] 对于使用帧间预测或双向预测编码的块,可以从流中提取运动信息,并继而将其用于为一幅或多幅用于帧间预测或双向预测的参考图像确定的空间显著映射图的运动补偿。

[0039] 如图6示例性示出的,时间显著映射图计算TSC是基于运动信息的。因此,它仅被确定以用于帧间预测或双向预测帧。在帧间或双向预测的帧内,帧内编码的宏块表示未被覆盖的或者显示不能通过帧间预测或双向预测很好地预测的这种大运动的区域。在一个示例性实施例中,二进制帧内编码块映射图ICM被用来确定时间显著映射图。例如,在二进制帧内编码块映射图中,每个内部块取值1。

[0040] 由于表示明显的、引人注意的运动的运动矢量不能被很好地预测从而需要明显更多的比特以用于编码,因此运动矢量编码成本映射图MCM被进一步用来确定时间显著映射图。运动矢量编码成本映射图MCM和帧内编码块映射图ICM被规格化并相加。被分配给合成映射图中的块的时间显著值可以被衰减以用于那些在跳跃(SKIP)或直接(DIRECT)模式中编码的块。例如,使用因子0.5对跳跃或直接模式编码的块的编码成本进行加权,使用其它模式编码的块的编码成本保持不变。

[0041] 融合FUS通过空间显著性计算SSC和时间显著性计算TSC产生的显著映射图可以是简单的相加。或者,如图7示例性示出的,可以在相加之前使用权重a,b对空间显著映射图和

时间显著映射图进行加权。权重a取决于帧中的帧内编码的块的相对数量,权重b取决于帧中的帧间或双向预测的块(P或B)的相对数量。融合FUS也可以使用利用取决于比特率变化和编码类型的权重c加权的之前的帧的之前的显著映射图。

[0042] 发明人的实验显示,a、b和c的下列示例值产生了好的结果:

$$[0043] \quad a = \frac{1}{12} + \frac{\text{number_of_I_MB}}{4 \times \text{number_of_MB}},$$

$$[0044] \quad b = \frac{1}{12} + \frac{\text{number_of_P_MB} + \text{number_of_B_MB}}{4 \times \text{number_of_MB}}$$

$$[0045] \quad c = \frac{1}{12} + \frac{f(\text{bitRate}, \text{type})}{4}$$

[0046] 其中,

$$[0047] \quad f(\text{bitRate}, \text{type}) = \frac{1}{2} + \Delta \text{bitRate} \text{ 用于双向预测帧 (B-帧)}$$

$$[0048] \quad f(\text{bitRate}, \text{type}) = \frac{1}{4} + \Delta \text{bitRate} \text{ 用于帧间预测帧 (P-帧)}$$

$$[0049] \quad f(\text{bitRate}, \text{type}) = \frac{1}{8} + \Delta \text{bitRate} \text{ 用于帧内预测帧 (I-帧)}。$$

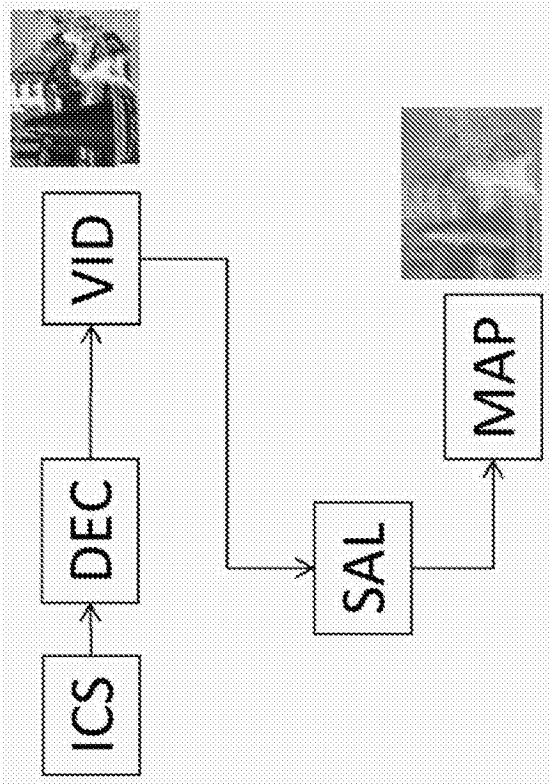


图1

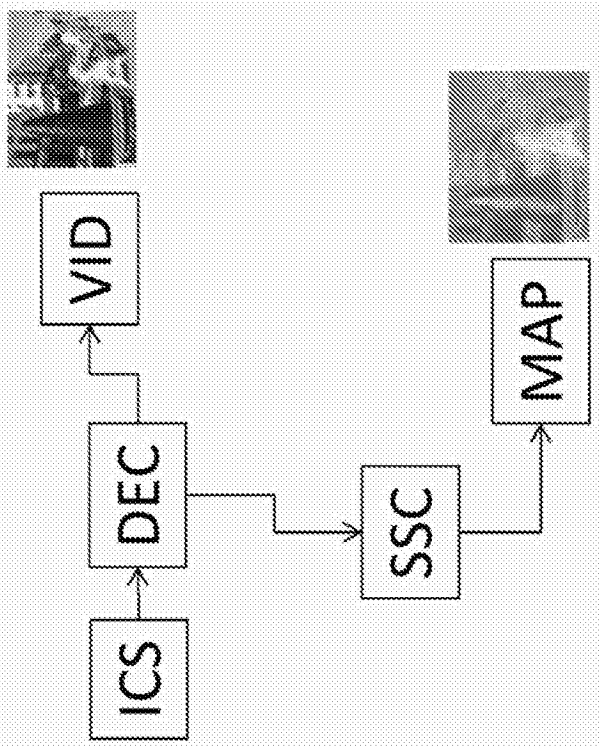


图2

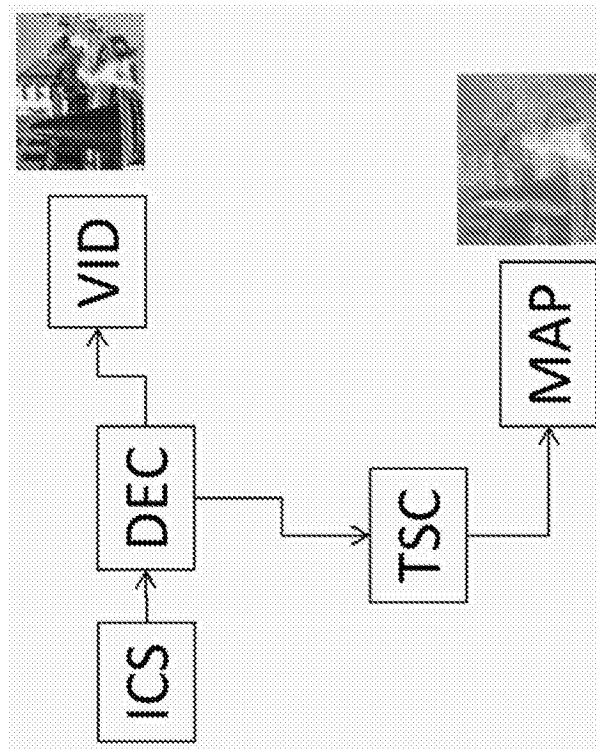


图3

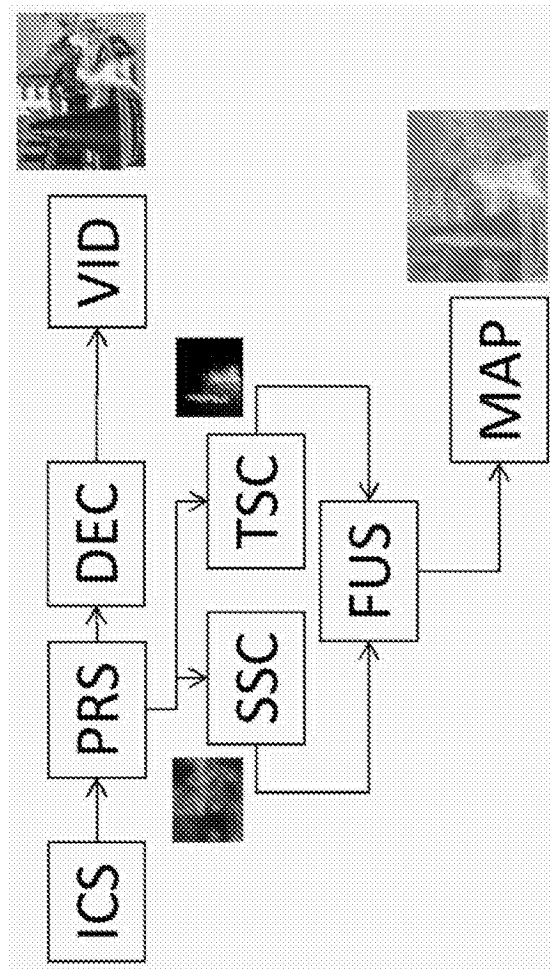


图4

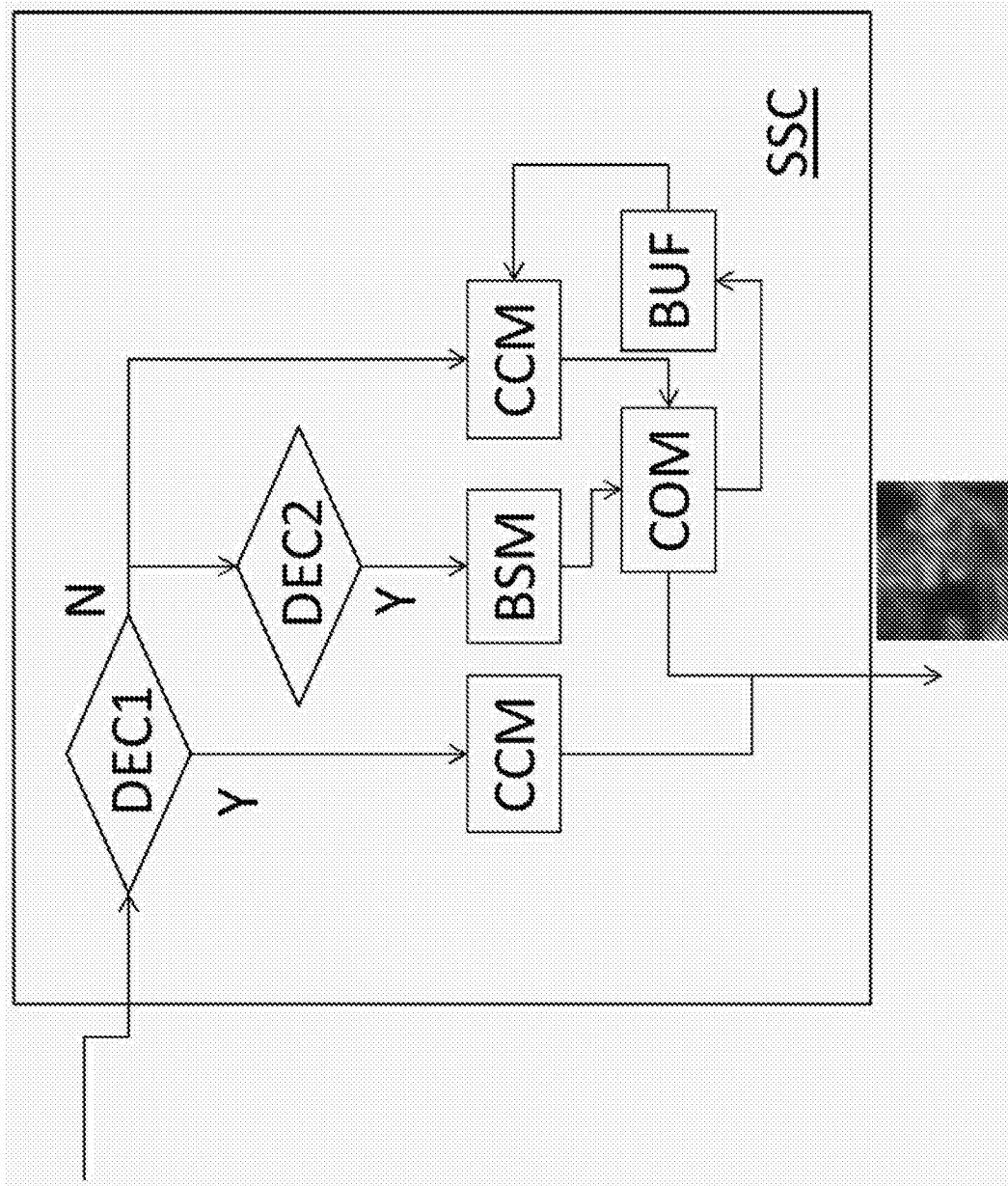


图5

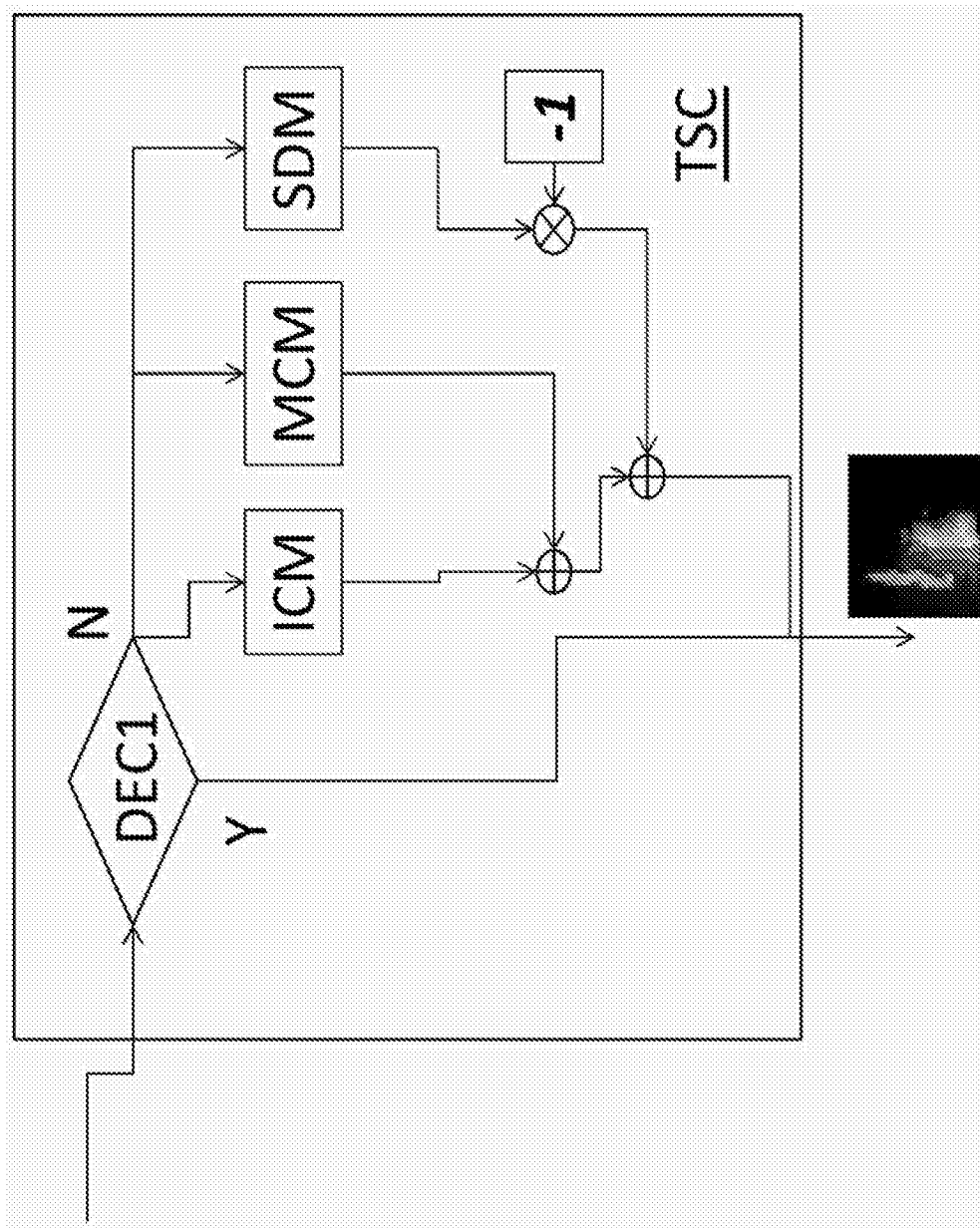


图6

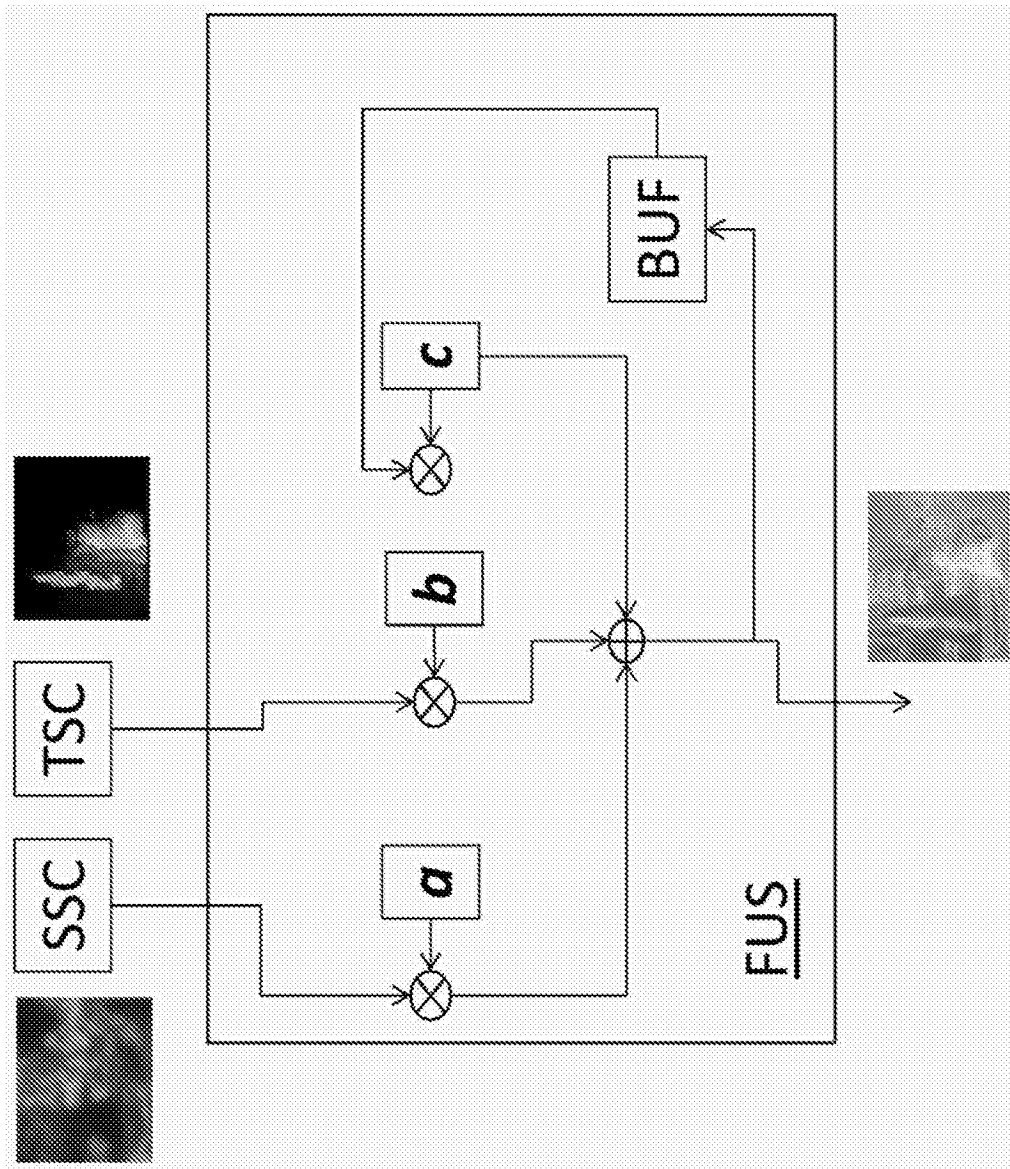


图7