

任意视点电视系统中深度数据的 自适应非均匀量化方法

刘晓仙, 常义林, 杨海涛, 元辉

(西安电子科技大学综合业务网理论及关键技术国家重点实验室, 710071, 西安)

摘要: 为了降低现有任意视点电视系统中深度数据量化所带来的量化失真, 提出一种自适应非均匀深度量化方法. 首先根据深度数据的统计特性, 选取有效深度子空间; 然后对有效深度子空间重新进行精细划分, 得到对应的自适应量化表; 最后使用该自适应量化表对原深度数据进行自适应量化. 实验结果表明, 与现有的均匀量化和普通非均匀量化方法相比, 使用所提出的方法可使深度数据的量化信噪比平均提高 20 dB. 因此, 自适应非均匀量化方法能够显著降低深度数据量化误差, 有助于在任意视点电视系统的接收端获得高精度深度数据, 从而得到主观效果更好的虚拟视图.

关键词: 任意视点电视; 深度数据; 深度子空间; 自适应非均匀量化

中图分类号: TN919.8 **文献标志码:** A **文章编号:** 0253-987X(2010)02-0082-06

Adaptive Non-Uniform Quantization of Depth Data in Free-Viewpoint Television System

LIU Xiaoxian, CHANG Yilin, YANG Haitao, YUAN Hui

(State Key Laboratory of Integrated Service Networks, Xidian University, Xi'an 710071, China)

Abstract: An adaptive non-uniform quantization method is proposed to reduce the distortion due to the depth data quantization in free-viewpoint television (FTV) system using the ordinary non-uniform quantization method. The proposed method determines the valid depth sub-spaces according to the statistical distribution characteristics of the depth data. Then, these depth sub-spaces are finely divided to form an adaptive quantization look-table. Finally, the original depth data are adaptively quantized using the quantization look-table. Experimental results and comparisons with the ordinary non-uniform depth data quantization method show that the proposed method improves the quantization signal-to-noise ratio of the depth data by an average of 20 dB. Thus the adaptive non-uniform quantization method can significantly reduce the quantization distortion of the depth data, and benefits the high-accurate reconstruction of the depth data and the better subject visual effect of the synthesized image at the receiver of the FTV system.

Keywords: free-viewpoint television; depth data; depth sub-space; adaptive non-uniform quantization

任意视点电视(FTV)系统于2002年在MPEG会议中提出^[1]之后, 得到工业界的普遍关注与大力支持^[2]. 2007年4月, MPEG组织建立了促进FTV标准化的FTV Ad Hoc组, 并于同年11月给出了

FTV系统的初步框架与要求^[3], 根据深度估计和虚拟视图合成的位置, 确定了FTV系统的实现方案^[4]. 为在FTV系统接收端合成虚拟视图, 实现用户的任意视点选择, 除视频数据外, 还需要获取与之

相应的深度数据^[5],这就使得高质量深度数据的获取、转换、传输、处理成为 FTV 系统中的关键技术之一.深度数据除可以直接通过深度摄像机拍摄得到之外^[6],还可以根据多视点视频数据采用立体匹配估计算法得到^[7-9].MPEG 组织所设计的 FTV 系统采用立体匹配估计算法获取深度数据^[10],并对深度数据进行相应的转换与量化,以利于深度数据的压缩、传输和存储.目前,FTV 系统使用普通非均匀量化方法对深度数据进行量化^[11].然而,普通非均匀量化方法没有考虑输入深度数据的统计特性,而造成深度数据的量化误差较大.因此,本文在普通非均匀量化方法的基础上,提出一种自适应非均匀深度量化方法,该方法依据统计到的深度数据分布特性自适应地更新量化表,有效降低了深度数据的量化误差.

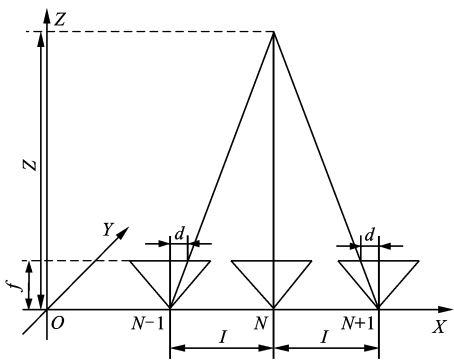
1 深度数据的普通非均匀量化方法^[11]

在 MPEG 组织提出的 FTV 系统方案中,使用立体匹配估计算法得到的是视差值^[10],因此需要将视差值转换为相应的深度值.图 1 描述了视差值与深度值之间的关系.由图 1 所示的三角形比例关系得 $f/d=Z/I$,即若已知摄像机间距、摄像机镜头的焦距,则深度值为

$$Z = \frac{If}{d} \tag{1}$$

普通非均匀量化方法^[11]将式(1)计算得到的深度值 Z 规整至 $[0,255]$ 范围内进行 8 b 非均匀量化,其中 0 表示最远深度值,255 表示最近深度值,则深度值 Z 的量化值为

$$D = \lfloor 255 \left(\frac{1}{Z} - \frac{1}{Z_{\text{far}}} \right) / \left(\frac{1}{Z_{\text{near}}} - \frac{1}{Z_{\text{far}}} \right) + 0.5 \rfloor \tag{2}$$



N :摄像机数; P :空间中一点; I :摄像机间距; f :摄像机镜头的焦距; Z : P 点沿摄像机 Z 轴到摄像机的真实距离,即 P 点的深度值; d : P 点的视差值

图 1 视差值与深度值的关系图

式中:符号 $\lfloor \alpha \rfloor$ 表示小于等于 α 的最大整数值; Z_{near} 与 Z_{far} 分别表示三维场景中的物体距离摄像机最近距离平面与最远距离平面的深度值.为了阐述方便,本文将 $1/Z_{\text{near}}$ 与 $1/Z_{\text{far}}$ 之间的空间称为深度空间.

在 FTV 系统接收端,根据式(2),可从接收到的量化值得到深度值的恢复值

$$Z_r = 1 / \left(\frac{D}{255} \left(\frac{1}{Z_{\text{near}}} - \frac{1}{Z_{\text{far}}} \right) + \frac{1}{Z_{\text{far}}} \right) \tag{3}$$

普通非均匀量化方法^[11]是在均匀量化方法^[10]的基础上对深度值 Z 进行非均匀量化得到的.均匀量化方法对 Z_{near} 与 Z_{far} 之间的深度值规整至 $[0,255]$ 范围内进行均匀量化,则深度值 Z 的量化值为

$$D = \lfloor 255 \frac{Z_{\text{far}} - Z}{Z_{\text{far}} - Z_{\text{near}}} + 0.5 \rfloor \tag{4}$$

同理,根据式(4),可从接收到的均匀量化后的量化值得到深度值的恢复值

$$Z_r = Z_{\text{far}} - D(Z_{\text{far}} - Z_{\text{near}}) / 255 \tag{5}$$

由于深度值 Z 的分布是非均匀的,与均匀量化方法相比,普通非均匀量化方法可得到更为精确的深度数据^[11],因此 MPEG 组织采用普通非均匀量化方法对深度数据进行量化.然而,从式(2)可以看出,普通非均匀量化方法实际是对深度空间进行均匀划分,即对 $1/Z$ 均匀量化,而 $1/Z$ 的分布并不是均匀的,因此普通非均匀量化方法会引入较大的量化误差.

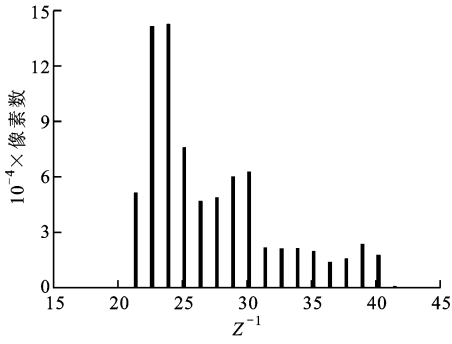
2 自适应非均匀深度量化方法

目前,MPEG 组织设计的 FTV 系统依据式(2)与式(3)对深度数据进行量化与反量化,并没有考虑输入深度数据的分布特性.因此,本文在普通非均匀量化方法的基础上,提出一种自适应非均匀深度量化方法.该方法依据统计到的深度数据分布特性设计自适应量化表,以减少深度数据的量化误差.

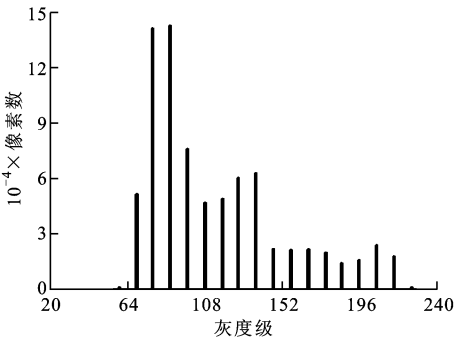
2.1 方法原理

由统计量化前的深度数据倒数 $1/Z$ 的分布特性,得到如图 2 所示的 $1/Z$ 的直方图.从图 2 可以看出, $1/Z$ 的分布并不是均匀的.根据最优量化器的设计原理,在给定量化的阶数的情况下,为了降低量化失真,应对信源分布概率密度高的区域细量化,而对概率密度低的区域粗量化.由于普通非均匀量化方法^[11]采用式(2)对非均匀分布的深度空间进行均匀量化,因此会引入较大的量化失真.

采用普通非均匀量化方法将图 2 所示的 $1/Z$

图2 量化前的 $1/Z$ 统计直方图

规整至 $[0, 255]$ 范围内进行量化,得到如图3所示的统计直方图,可以看出存在未被利用的量化区间.基于此,本文提出一种自适应非均匀量化方法,仅对那些有深度对象的深度区间进行有效表示和量化,剔除那些无深度对象的深度区间,从而提高有深度对象的深度区间的量化性能.先根据深度数据的分布特性,找出那些有深度数据落入的全部子空间,标记为有效深度子空间,然后剔除无深度数据的子空间,重新按顺序连结有效深度子空间,构成有效深度子空间集合,并对其进行重新划分,得到精细的量化区间,从而有效降低深度数据的量化误差.

图3 采用普通非均匀量化方法对图2所示的 $1/Z$ 进行量化后的统计直方图

2.2 FTV 系统发送端自适应非均匀量化方法的实现

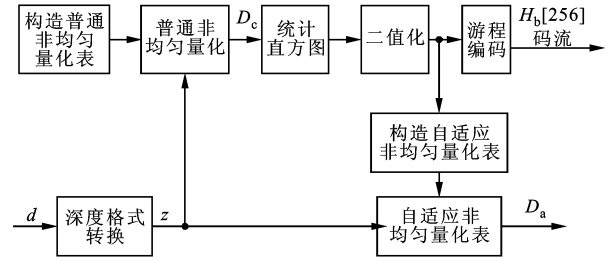
FTV 系统发送端自适应非均匀量化方法的流程如图4所示.

(1)利用式(3)描述的反量化过程求出每一个量化输出值 i 对应的量化区间的下界与上界

$$\left. \begin{aligned} Z_{cl}[i] &= 1 / \left(\frac{i+0.5}{255} \left(\frac{1}{Z_{near}} - \frac{1}{Z_{far}} \right) + \frac{1}{Z_{far}} \right) \\ Z_{cu}[i] &= 1 / \left(\frac{i-0.5}{255} \left(\frac{1}{Z_{near}} - \frac{1}{Z_{far}} \right) + \frac{1}{Z_{far}} \right) \end{aligned} \right\} \quad (6)$$

$$0 \leq i \pm 0.5 \leq 255$$

同时由式(1)将立体匹配计算得到的视差值 d 转换为深度值 Z .接着,对输入的深度值 Z 进行普通非均



D_c :采用普通非均匀量化方法的量化输出值; D_a :采用自适应非均匀量化方法的量化输出值; $H_b[256]$:普通非均匀量化表中的有效量化区间

图4 FTV 系统发送端自适应量化方法流程

匀量化.如果当前输入深度值 Z 落入普通非均匀量化表中某个量化输出值 i 对应的量化区间内,则 Z 的量化值为

$$\left. \begin{aligned} D_c &= i, Z \in [Z_{cl}[i], Z_{cu}[i]] \\ 0 &\leq i \pm 0.5 \leq 255 \end{aligned} \right\} \quad (7)$$

(2)统计深度数据的分布特性.计算 D_c 的统计直方图 $H[256]$,并对 $H[256]$ 进行二值化,即凡有深度数据落入的量化区间标记为1,否则为0,得到普通非均匀量化表中的有效量化区间 $H_b[256]$.

(3)根据 $H_b[256]$ 及普通非均匀量化表构造自适应量化表.若普通非均匀量化表中有 K 个有效量化区间,则需要将 K 个有效量化区间重新映射为256份,得到自适应量化区间.对 K 个有效量化区间进行映射的直接方法是用 K 整除256,而该方法的缺陷在于所得到的最后一个映射量化区间的长度可能会比较大,从而增加量化误差.为减少量化区间划分引入的误差,算法采用自适应量化区间划分,即对每一个有效量化区间划分结束后,重新调整普通非均匀量化表中剩余的有效量化区间数量与自适应量化表中剩余待分配的量化区间数量,迭代 $(K-1)$ 次结束.

以普通非均匀量化表中第 k 个有效量化区间 $[Z_{cl}[k], Z_{cu}[k]]$ 为例,其自适应量化区间的计算步骤如下.

(1)确定普通非均匀量化表中剩余待分配的有效量化区间数量 $N_c = K - k + 1$.

(2)确定自适应量化表中剩余待分配的量化区间数量 N_a .

(3)确定本次分配中将使用的自适应量化表中的量化区间数量

$$N_e = \lfloor \frac{N_a}{N_c} \rfloor \quad (8)$$

(4) 确定本次分配中将使用的自适应量化表中的量化区间序号

$$s = 256 - N_a + m, \quad 0 \leq m \leq N_e - 1 \quad (9)$$

(5) 根据式(6),将其中的量化区间数量 255 用 N_e 替代, Z_{near} 和 Z_{far} 分别用普通非均匀量化表中第 k 个有效量化区间的下界 $Z_{\text{cl}}[k]$ 与上界 $Z_{\text{cu}}[k]$ 替代,从而得到本次分配所有量化区间的下界与上界

$$\left. \begin{aligned} Z_{\text{al}}[s] &= 1 / \left(\frac{m+0.5}{N_e} \left(\frac{1}{Z_{\text{cl}}[k]} - \frac{1}{Z_{\text{cu}}[k]} \right) + \frac{1}{Z_{\text{cu}}[k]} \right) \\ Z_{\text{au}}[s] &= 1 / \left(\frac{m-0.5}{N_e} \left(\frac{1}{Z_{\text{cl}}[k]} - \frac{1}{Z_{\text{cu}}[k]} \right) + \frac{1}{Z_{\text{cu}}[k]} \right) \end{aligned} \right\} \quad 0 \leq m \pm 0.5 \leq N_e - 1 \quad (10)$$

在计算得到所有自适应量化区间之后,对输入的深度值 Z 进行自适应非均匀量化. 如果 Z 落入自适应非均匀量化表中某个量化输出值 i 对应的量化区间内,则 Z 的量化值为

$$\left. \begin{aligned} D_a &= i, \quad Z \in [Z_{\text{al}}[i], Z_{\text{au}}[i]] \\ 0 &\leq i \leq 255 \end{aligned} \right\} \quad (11)$$

同时,为了在 FTV 系统接收端构造同样的自适应量化表,需将 1、0 子区间位置经游程编码后传输到接收端.

2.3 FTV 系统接收端自适应非均匀量化方法的实现

FTV 系统接收端自适应非均匀量化方法的实现步骤如下,首先对 1、0 子区间位置进行游程解码,再使用与 FTV 发送端相同的方法由 1、0 子区间位置构造自适应量化表,进而根据式(4),将其中的 Z_{near} 和 Z_{far} 分别用 $Z_{\text{al}}[i]$ 与 $Z_{\text{au}}[i]$ 替代. 由于相邻 2 个量化输出值距离为 1,因此自适应非均匀反量化表中每一个量化输出值 i 所对应的反量化值为

$$\left. \begin{aligned} I_a[i] &= 1 / \left(0.5 \left(\frac{1}{Z_{\text{al}}[i]} - \frac{1}{Z_{\text{au}}[i]} \right) + \frac{1}{Z_{\text{au}}[i]} \right) \\ 0 &\leq i \leq 255 \end{aligned} \right\} \quad (12)$$

最后,对深度数据的自适应量化输出值 D_a 进行反量化. 根据式(12),将 D_a 作为输入,深度值的恢复值为

$$Z_r = I_a[D_a] \quad (13)$$

3 实验结果与分析

为了验证本文方法的性能,以 MPEG 组织提供的深度估计参考软件 DE2.2 和虚拟视图合成参考

软件 VS2.3 为实验平台,采用 MPEG 组织提供的 4 组多视点视频序列 Book-arrival、Leaving-laptop、Dog 与 Pantomime 进行实验^[12],对比测试均匀量化方法、普通非均匀量化方法以及自适应非均匀量化方法所得到的深度数据的量化信噪比,同时比较 3 种量化方法对 FTV 系统接收端合成的虚拟视图的主观视觉效果的影响.

首先,分别估计左视点视图与右视点视图的深度数据. 根据参考视图与待估计的中央视图的摄像机间距,实验中分别选取窄摄像机间距方式(即选取与待估计视图相邻的左右视图作为参考视图),与宽摄像机间距方式(即选取与待估计视图左右两边分别间隔一个视图的视图作为参考视图),进行深度估计. 深度估计的配置参数均根据 MPEG 组织所推荐的实验条件^[12]设定. 其次,根据式(1)将立体匹配计算得到的各像素的视差值 d 转换为深度值 Z ,分别采用 3 种量化方法对深度值 Z 进行量化. 在 FTV 系统接收端,分别利用式(5)、式(3)与式(13)计算经过 3 种量化方法所得到的深度值的恢复值 Z_r ,然后根据式(14)分别计算 3 种量化方法得到的深度数据的量化信噪比. 实验结果如表 1 和表 2 所示.

$$\left(\frac{S}{N_q} \right) = 10 \lg \frac{\left(\sum_{x=0}^{W-1} \sum_{y=0}^{h-1} |Z(x, y)|^2 \right) / (Wh)}{\left(\sum_{x=0}^{W-1} \sum_{y=0}^{h-1} |Z(x, y) - Z_r(x, y)|^2 \right) / (Wh)} \quad (14)$$

式中: W 表示图像的宽度; h 表示图像的高度.

从表 1 和表 2 可以看出,与其他 2 种方法相比,采用本文方法所得到的深度数据的量化信噪比平均可以提高 20 dB 左右. 由于均匀量化方法和普通非均匀量化方法均是对非均匀分布的信源进行均匀量化,因此会引入较大的量化误差,而本文的自适应非均匀量化方法可以充分利用深度数据的分布特性,对非均匀分布的信源进行自适应量化,从而有效提高了深度数据的量化信噪比.

最后,采用 VS 2.3 软件对 Book-arrival 序列和 Dog 序列进行虚拟视图合成,其主观视觉效果对比如图 5 和图 6 所示. 从图 5 与图 6 中圈出的地方可以看出,与均匀量化方法和普通非均匀量化方法相比,自适应非均匀量化方法均可以获得更好的主观视觉效果. 这是因为采用自适应量化方法可得到

量化信噪比较高的深度数据,而高精度的深度数据 重要的影响.
对于虚拟视图合成,特别是视图边缘区域的合成有

表 1 窄摄像机间距方式时左视点视图与右视点视图处的深度数据的量化信噪比

视频序列	量化信噪比/dB					
	左视点视图			右视点视图		
	均匀量化方法	普通非均匀量化方法	本文方法	均匀量化方法	普通非均匀量化方法	本文方法
Book_arrival	55.46	55.57	78.77	55.03	55.16	79.47
Leaving_laptop	54.91	55.43	78.54	54.46	54.81	79.22
Dog	64.83	65.34	94.33	65.70	66.60	95.54
Pantomime	52.37	56.37	84.05	53.41	55.55	77.58

表 2 宽摄像机间距方式时左视点视图与右视点视图处的深度数据的量化信噪比

视频序列	量化信噪比/dB					
	左视点视图			右视点视图		
	均匀量化方法	普通非均匀量化方法	本文方法	均匀量化方法	普通非均匀量化方法	本文方法
Book_arrival	56.33	56.91	72.56	55.10	55.62	73.43
Leaving_laptop	56.15	56.66	72.59	55.57	55.67	73.77
Dog	61.41	62.26	86.82	62.52	62.99	87.59
Pantomime	57.85	58.60	86.07	57.68	58.50	78.10



(a)原始视图 (b)均匀量化方法 (c)普通非均匀量化方法 (d)本文方法

图 5 采用 3 种量化方法对 Book_arrival 序列合成虚拟视图的主观视觉效果对比



(a)原始视图 (b)均匀量化方法 (c)普通非均匀量化方法 (d)本文方法

图 6 采用 3 种量化方法对 Dog 序列合成虚拟视图的主观视觉效果对比

4 结 论

为了降低深度量化过程中所引入的量化误差,提出一种基于深度数据分布特性的自适应非均匀量化方法.实验结果表明,与均匀量化方法以及普通非均匀量化方法相比,在FTV系统中采用自适应非均匀量化方法,深度值的量化信噪比平均提高了20 dB左右,极大降低了深度值的平均量化误差,因此可在FTV系统接收端得到高精度的深度数据,从而可以获得主观视觉效果更好的虚拟视图.

参考文献:

- [1] MASAYUKI T, TOSHIKI F. FTV: free viewpoint television [EB/OL]. (2002-07-16)[2008-05-20]. <http://wg11.sc29.org/mpeg/docs/61.Klagenfurt/contrib/m8595.zip>.
- [2] Moving Picture Experts Group. Call for comments on 3DAV [EB/OL]. (2003-10-24)[2008-06-15]. <http://wg11.sc29.org/mpeg/docs/66.Brisbane/wg11/w6051.zip>.
- [3] Moving Picture Experts Group. Preliminary FTV model and requirements [EB/OL]. (2007-07-07)[2008-06-15]. <http://wg11.sc29.org/mpeg/docs/81.Lausanne/wg11/w9168.zip>.
- [4] MASAYUKI T, TOSHIKI F. Proposal on EE for FTV [EB/OL]. (2007-06-27)[2008-06-15]. <http://wg11.sc29.org/mpeg/docs/81.Lausanne/contrib/m14647.zip>.
- [5] FEHN C. Depth-image-based rendering (DIBR): the compression and transmission for a new approach on 3DTV [C]//Proceedings of Stereoscopic Displays and Virtual Reality Systems XI. Bellingham, WA, USA: SPIE, 2004: 93-104.
- [6] IDDAN G J, YAHAV G. 3D imaging in the studio [C]//Proceedings of Videometrics and Optical Methods for 3D Shape Measurements. Bellingham, WA, USA: SPIE, 2001: 48-55.
- [7] BLEYER M, GELAUTZ M. Graph-cut-based stereo matching using image segmentation with symmetrical treatment of occlusions [J]. Signal Processing: Image Communication, 2007, 22(2): 127-143.
- [8] INCE S, MARTINIAN E, YEA S, et al. Depth estimation for view synthesis in multiview video coding [C]//Proceedings of IEEE 3DTV Conference. Piscat-

away, NJ, USA: IEEE, 2007: 1-4.

- [9] LEE S B, OH K J, HO Y S. Segment based multi-view depth map estimation using belief propagation from dense multi-view video [C]//Proceedings of IEEE 3DTV Conference. Piscataway, NJ, USA: IEEE, 2008: 193-196.
- [10] MASAYUKI T, TOSHIKI F, KAZUYOSHI S. Multi-view depth map of Rena and Akko & Kayo [EB/OL]. (2007-10-17)[2008-06-15]. <http://wg11.sc29.org/mpeg/docs/82.Shenzhen/contrib/m14888.zip>.
- [11] MASAYUKI T, TOSHIKI F, KAZUYOSHI S. Improvement of depth map estimation and view synthesis [EB/OL]. (2008-01-09)[2008-06-15]. <http://wg11.sc29.org/mpeg/docs/83.Antalya/contrib/m15090.zip>.
- [12] Moving Picture Experts Group. Description of exploration experiments in 3D video coding [EB/OL]. (2008-07-28)[2008-08-10]. <http://wg11.sc29.org/meetings/85.Hannover/wg11/w9991.zip>.

[本刊相关文献链接]

视频通信中的选择性重传技术. 2009, 43(6): 282-287.

可分级视频编码覆盖网非一致性随机成员构建算法. 2009, 43(6): 1-4.

适合于小波视频编码的尺寸自适应图像组结构. 2009, 43(4): 52-56.

利用 MPEG-4 视频流的运动目标检测算法. 2008, 42(4): 454-457.

保证 H. 264 视频通信质量的实时传输协议载荷. 2007, 41(12): 1455-1459.

基于时域局部线性预测的视频镜头变换检测. 2006, 40(10): 1060-1064.

基于高斯混合模型的视频对象分割算法. 2006, 40(6): 724-728.

一种有效的 MPEG 视频运动注意力区域提取方法. 2005, 39(10): 1135-1138.

一种新的交通流视频检测方法. 2004, 38(4): 396-399.

摄像机运动识别及视频分割的组合检测新方法. 2003, 37(12): 1267-1271.

(编辑 刘杨)