



一种新的立体显示中的视差估计与虚拟视点绘制算法

摘要

针对虚拟视图绘制技术中空洞和伪边缘等问题,提出了一种新的立体匹配算法以获得稠密视差图完成虚拟视点合成.此方法只需要两张视图,对其采用 Mean shift 算法就可以进行视图分割.有效的虚拟视点绘制是基于良好的视差图与好的视图插值法实现的.通过权值多窗口立体匹配法获得初始视差图,再利用交叉检验法对不可信匹配点和闭塞遮挡区域进行滤除,很好地解决了视差图中遮挡区域引起的误匹配问题.最后,对获得的稠密视差图进行视图插值和相应的噪声滤除完成了多点位置的虚拟视图绘制.针对 Middlebury dataset 提供的“Tsukuba”等测试图像进行相关实验,实验结果表明,该算法具有良好的视图绘制效果.

关键词

立体匹配;视图插值;交叉检验;稠密视差图;虚拟视点绘制

中图分类号 TP391

文献标志码 A

收稿日期 2016-03-20

资助项目 江苏省高校信息与通信工程优势学科建设项目;江苏省气象探测与信息处理重点实验室项目(KDXS1403);江苏高校品牌专业建设工程资助项目(PPZY2015B134)

作者简介

梅永,女,博士,研究方向为图形图像处理、信号处理等.yongmei2002@126.com

1 南京信息工程大学 江苏省气象探测与信息处理重点实验室,南京,210044

2 南京信息工程大学 电子与信息工程学院,南京,210044

0 引言

在近年来的显示技术领域,3D 视频显示技术已经引起了人们高度的重视,它在多媒体中有着广泛的应用,比如 3D 游戏、电影、视频会议、3DTV 和医疗图像等.随着人们对 3D 视频显示技术需求的不断增加,MPEG(动态图像专家组)对于 3D 虚拟视频技术(3DAV)的标准化有着重大的贡献^[1].3D 视频显示的信息是由存储左右视图像素对的视差值的视差图来表征的.在视差图的基础上,对应不同相机位置处的虚拟视点都可以被合成,因此,视差估计和虚拟视点绘制是 3D 视频显示中最关键的两个部分.

虚拟视点合成就是从已知的左右视图来得到虚拟的视图,这个过程中最主要的是通过相关的立体匹配过程找到对应的像素对.在机器视觉中一个经典的问题——立体匹配法,它仍然活跃在近些年的相关领域中.考虑到视图对来源于同一个场景的事实,立体匹配的关键是找到两图之间的视差值.目前提出的大量算法可以被分为两大类,具体可参见文献[2].

视差估计中,立体匹配是核心过程.受到 Singh 等^[3]提出的立体匹配方法的启示,本文在立体匹配过程中采用对灰度图像和彩色图像都使用的 Mean shift 分割^[4]和权重多窗口^[5]相结合的立体匹配法,通过在同一场景中左右两幅视图获得的深度信息和原始图信息进行视图插值获得虚拟视图,可以完成已知左右视图之间任意位置的图像绘制.本文提出的方法具有良好的图像效果,轮廓清晰分明,可以实现已知双视图之间任意位置视图绘制,具有良好的立体效果.

1 立体摄像机系统原理

本论文算法是在假设两个摄像机平行放置前提下实现的.立体摄像机系统几何原理如图 1 所示.本文用到两个已知的摄像机 C_L 和 C_R ,两相机的坐标相互平行,它们的图像平面是共面的. B 表示基线距离, F 表示相机的焦距.在此模型中,3D 视图中某一点 P 在两幅已知图像中都可见,但该点在坐标上有所不同,它们在坐标中的差异就被称为视差值.视差估计应该满足极限约束,因为两相机的平行设置,那么垂直向量之间没有差距,目的是为了沿着极线方向只有一维变量在变化,这个设置对于本文十分关键.

假设场景中一点 P 在左视图和右视图平面内对应点分别是 P_L

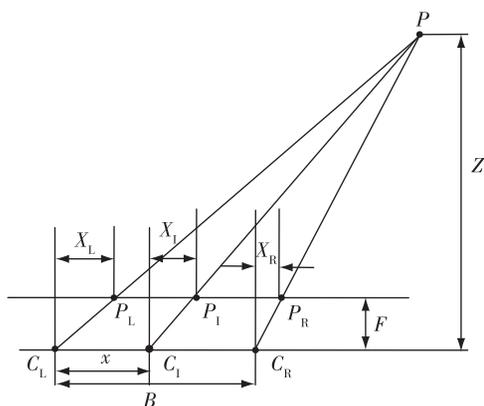


图1 立体摄像机系统

Fig. 1 Stereo camera system

和 P_R , 如果把左视图视为参考图, 那么从左到右的视差值可以表示为

$$d_{LR} = X_R - X_L = -\frac{BF}{Z}, \quad (1)$$

其中, Z 表示点 P 的深度.

若假设在左右摄像机之间有某个虚拟摄像机 C_I 并且与左摄像机的距离为 x , 那么从左视图到中间任一点虚拟视图的视差值可以表示为

$$d_{LI} = X_I - X_L = -\frac{xF}{Z} = \left(\frac{x}{B}\right) d_{LR} = \alpha d_{LR}, \quad (2)$$

其中, $\alpha = \frac{x}{B}$, $0 < \alpha < 1$.

由式(1)和式(2)可得出如下关系式:

$$X_I = (1 - \alpha)X_L + \alpha X_R = X_L + \alpha d_{LR}. \quad (3)$$

因此, 如果已知中间虚拟摄像机的坐标位置, 能够确定原始图像像素点和它们的视差值, 那么点 P 在虚拟视图中的坐标位置即可由式(3)确定.

2 视差估计

本算法的设计流程如图2所示.

虚拟视点合成就是从已知的左右视图来得到虚拟的视图. 已知左视图和右视图, 一种方法是直接进行视差图的虚拟视点绘制. 本文的方法是先对左右视图进行区域分割, 在图像基础上利用基于像素间的颜色相似性的权值多窗口匹配法, 在匹配过程中引入匹配代价函数并利用 Winner-Take-All 思想优化视差值, 从而获得较精确的立体图像对的初始视差图, 再对初始视图存在的误匹配点进行剔除, 利用 Mean shift 分割的图像对视差图进行平滑操作, 最后对前面获得的稠密视差图进行正向视图插值, 并对

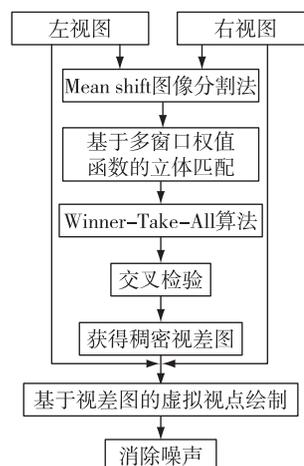


图2 算法设计流程

Fig. 2 Algorithm flow chart

正向映射插值出现的空洞现象进行填补, 以及对于视差估计过程中出现的噪声进行滤除, 最终完成已知左右视图之间任意位置的图像绘制.

2.1 Mean shift 图像分割

Comaniciu 等^[6]提出的 Mean shift 图像分割算法具有计算量小和图像边缘信息保留良好的优点. 本文鉴于该方法的优点, 将其融合应用到视差估计算法中. Mean shift 算法可以简述为: 将已知参考视图按照颜色的相似性分成若干区域并将不同颜色区域标号, 但需要假设视差值较大的像素点视为颜色不连续的边界.

假设 x_i 和 z_i ($i = 1, 2, 3, \dots, n$) 分别表示 d 维原始图像点和收敛点, L_i 为图像分割后的第 i 个区域标号, 那么 Mean shift 图像分割流程如下:

1) 将图像转换为合适的色度空间. 灰度图可直接操作, 彩色图需转换为 $L \times U \times V$ 色度空间进行操作.

2) 对 Mean shift 进行平滑操作:

① 令 $j = 1, y_{i,1} = x_i$;

② 计算当前像素中心点 $y_{i,j+1}$;

③ 计算 Mean shift 向量 $m_{h,c}(y_{i,j}) = y_{i,j+1} - y_{i,j}$;

④ 重复上述的操作直到找到收敛点 $y = y_{i,c}, z_i =$

$(x_i^s, y_{i,c}^r)$.

3) 将各像素的收敛点 z_i 按照一定的规则进行合并, 具体条件是空间域小于 h_s , 颜色域小于 h_r .

4) 完成所需图像的分割算法.

2.2 基于多窗口权值的立体匹配法

本文采用局部匹配算法^[7], 该算法假设匹配窗

口内所有像素在一个相近深度平面,这些像素点具有相似的视差值.该方法的基本思想是在深度不连续区域和同一区域内能得到精确的匹配结果,要为每个像素自适应地选取合适的匹配窗口进行立体匹配.一般初始视差图是采用基于窗口匹配法获得的.本文借鉴文献[8]提出的基于窗口权值函数匹配法,它在深度值不连续处具有良好的预留信息的优点,但在图像区域有过平滑现象出现,因此本文在此基础上采用多窗口匹配的方法,消除此缺点.

本文介绍的自适应匹配法中相似性测度函数采用的是绝对差之和 SAD(Sum of Absolute Difference)和图像梯度信息 GRAD 结合的匹配代价函数 $C(x,y)$:

$$C_{\text{SAD}}(x,y) = \sum_{c \in \{r,g,b\}} |I_c(x) - I_c(y)|, \quad (4)$$

$$C_{\text{GRAD}}(x,y) = \sum_{c \in \{r,g,b\}} |\nabla_x I_c(x) - \nabla_x I_c(y)|^2 + \sum_{c \in \{r,g,b\}} |\nabla_y I_c(x) - \nabla_y I_c(y)|^2, \quad (5)$$

$$C(x,y) = (1-w) \times C_{\text{GRAD}}(x,y) + w \times C_{\text{SAD}}(x,y), \quad (6)$$

其中 $I_c(x)$ 表示当前像素点彩色值, x 表示当前点坐标, $I_c(y)$ 表示当前点周围领域彩色值, y 表示领域坐标, ∇_x 和 ∇_y 分别代表图像横向和纵向梯度, w 是介于 0 和 1 之间的权值.

设计 RGB 图像的代价函数:

1) 颜色相似性权值公式

$$f_s(\Delta c_{pq}) = \exp\left(-\frac{\Delta c_{pq}}{\gamma_c}\right), \quad (7)$$

其中, Δc_{pq} 为两像素点的颜色差异, $\Delta c_{pq} = |R_p - R_q| + |G_p - G_q| + |B_p - B_q|$, γ_c 为评价颜色相似性参数,此处取 3.

2) 参考视图和目标图中多窗口权值计算公式

$$w'(p,q) = \begin{cases} 1, \\ \exp\left(-\frac{\Delta c_{pq}}{\gamma_c}\right), \end{cases} \quad q \in S_p, \quad (8)$$

其中, S_p 是点 p 的移动区域.

3) 图形分割加权累计误差能量函数

$$E'(p,\bar{p}_d) = \frac{\sum_{q \in N_p, q_d \in N_{p,d}} w'(p,q)w'(\bar{p}_d,\bar{q}_d)C(q,\bar{q}_d)}{\sum_{q \in N_p, q_d \in N_{p,d}} w'(p,q)w'(\bar{p}_d,\bar{q}_d)}, \quad (9)$$

其中, \bar{p}_d 和 \bar{q}_d 分别是目标图像中 p, q 在视差为 d 时的对应点, $N_{p,d}$ 表示 \bar{p}_d 支持窗内的其余任意像素,

$C(q,\bar{q}_d)$ 表示 q 和 \bar{q}_d 的绝对灰度差.

本文算法采用 WTA(Winner-Takes-All)和误差能量函数相结合的方法,在视差值范围 $[d_{\min}, d_{\max}]$ 内,计算代价函数值为最小时的视差值作为该像素的视差值,以此得到初始视差图.

2.3 交叉检验

在初始视差图的基础上,本文采用双向交叉检验法滤除不可靠的视差值,以确保视差值的连续性.文献[9]中提出两幅视图间的像素对 p 和 q 存在弱连续性限制,这种限制也适用于闭塞区域.通过计算左视图视差值与右视图视差值,检测出闭塞区域.如果违背弱连续性限制,即:

$$d_{L \rightarrow R}(x,y) > d_{R \rightarrow L}(x - d_{L \rightarrow R}(x,y),y), \quad (10)$$

则将此视为闭塞区域,用 $d_{R \rightarrow L}(x - d_{L \rightarrow R}(x,y),y)$ 代替 $d_{L \rightarrow R}(x,y)$.本文在此基础上加入能量误差函数的限制条件,使得视差值更加精确,能更好判断出闭塞区域并加以处理.如果 $e_{R \rightarrow L}(x,y) \leq e_{L \rightarrow R}(x,y)$,则为正确视差值,否则将拒绝该视差值, $d_{R \rightarrow L}(x,y) = \infty$.

3 虚拟视点绘制

3.1 视图插值

在立体匹配法获得的稠密视差图基础上,本文采用线性视图插值法将左右视图像素对的颜色信息插值到虚拟视点位置完成虚拟视点绘制.

X_1 表示虚拟位置视图, β 表示虚拟视图位置的参数 $(\beta = \left| \frac{C_1 - C_L}{C_R - C_L} \right|, 0 < \beta < 1)$, $p(x_L, y_1)$ 表示像素对 $\{p(x_L, y), q(x_R, y)\}$ 在虚拟视图 X_1 中对应的像素点,那么虚拟视图中像素点位置可表示为

$$x_1 = (1-\beta) \times x_L + \beta \times x_R, \quad (11)$$

$p(x_1, y_1)$ 的像素值为

$$V_s = (1-\beta) \times V_p + \beta \times V_q, \quad (12)$$

其中 V_p 和 V_q 分别表示原始图像像素对的像素值.

3.2 滤除噪声

采用正向视图插值会出现空洞现象,主要是对象之间的遮挡和区域拉伸原因引起的视图中一些边界和隐藏点无法重现,这些点在视差图中视差值为 0.本文采用方法是确定空洞点的位置,利用左视图在虚拟位置的虚拟视图对空洞点位置进行颜色信息填充.

$$I_{\text{IL}}(X_1, Y) = I_{\text{IL}}(X_L + (1-\alpha) \times d^{\text{IL}}, Y) = I_L(X_L, Y), \\ I_1(X_1, Y) = I_{\text{IL}}(X_1, Y) + I_{\text{IR}}(X_1, Y). \quad (13)$$

其中, $X_L + (1 - \alpha) \times d^{ll}$ 表示虚拟视点位置与左视图位置关系, $I_{ll}(X_L, Y)$ 表示由左视图变换到中间虚拟视点像素坐标为 (X_L, Y) 处的像素值, d^{ll} 表示左右视图的视差值, $I_L(X_L, Y)$ 表示左视图像素坐标为 (X_L, Y) 处的像素值, $I_I(X_L, Y)$ 表示最终虚拟视图在像素坐标为 (X_L, Y) 处的像素值, $I_{rr}(X_L, Y)$ 表示由右视图变换到中间虚拟视点像素坐标为 (X_L, Y) 处的像素值.

生成的虚拟视图必然会存在噪声,主要表现为灰度值很大、像素不连续等.为滤除噪声,本文采用相邻像素的颜色信息对存在噪声的像素点进行替代处理.

4 仿真结果与分析

本文所有仿真都是在 Matlab 2012b 软件环境下完成的,实验中的素材和标准视差图来源于 Middlebury dataset 提供的测试图像.图 3 是采用 Mean shift 算法的区域分割图像,可以发现对于不同的参数取值的分割效果明显不同.

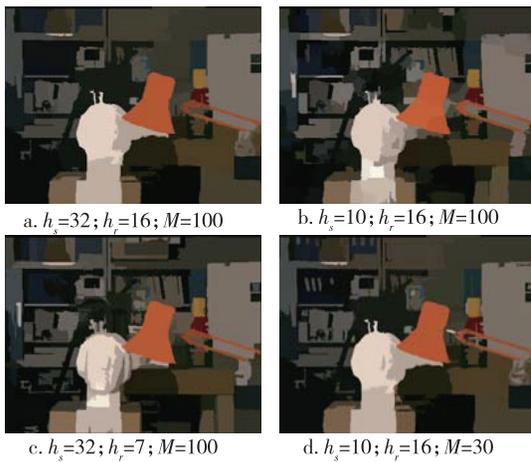


图 3 不同参数的 Mean shift 分割算法

Fig. 3 Mean shift segmentation results with different parameters

由图 3 可以看出:图像分割效果受到 3 个参数影响,其中 h_s 是控制区域的大小,区域越大进入空间区域的点越多,分割效果并不与该值呈现正比关系,此值一般不易过大; h_r 和 M 对分割效果有一定的影响, M 值越小景物细节保留得越清晰, h_r 值越小,色彩的细节保留越好.在分割后互不相干的几个区域,区域的划分已经消除了距离上的差异特征,匹配算法中的代价函数只需考虑相关窗口内像素的颜色相似性.该方法在绘制速度上有待进一步提高.

图 4 是左右参考视图与中间虚拟视图进行噪声滤除前后的效果.由图 4 可清晰地看出,不经过相关处理的虚拟绘制视图中一些区域有明显的像素断裂和空洞,尤其是在物体的边缘附近,而经过处理的绘制方法能有效地解决这些问题,景物色彩光滑连续,绘制效果更理想.图 5 给出了测试图像“Tsukuba”的原左右视图和不同位置处的虚拟合成视图.在高精度稠密视差图的基础上,绘制出了高质量的虚拟视图,表明本文中提出的立体匹配算法的精确度很高.

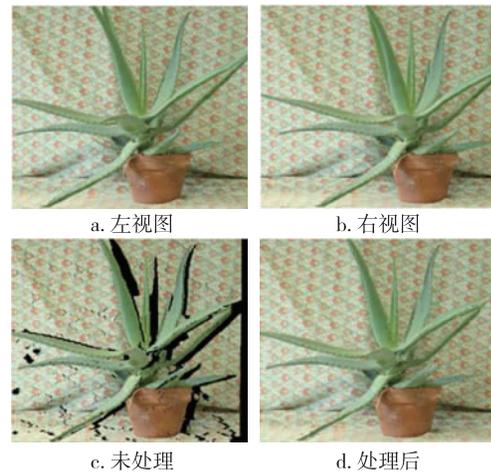


图 4 处理前后对比

Fig. 4 Original left view (a), right view (b) of test image, before (c) and after (d) noise filter

5 总结

本文提出了一种新的基于 Mean shift 图像区域分割和多窗口权值的立体匹配算法.有效的虚拟视点绘制是基于良好的视差图与好的视图插值法实现的.本文通过权值多窗口立体匹配法获得初始视差图,再利用交叉检验法对不可信匹配点和闭塞遮挡区域进行滤除,很好地解决了视差图中遮挡区域引起的误匹配问题.仿真结果表明本文提出的算法具有良好的绘制效果,绘制速度较快,而且对于灰度图和彩色图都适用.本文未讨论视点垂直方向的变化对虚拟视图的影响,对于更具有挑战性的自由立体显示技术^[9-10]有待进一步的研究.

参考文献

References

- [1] Lee C M G, Travis A R L, Lin R. Flat-panel autostereoscopic 3D display[J]. IET Optoelectronics, 2009, 2(1): 24-28
- [2] Cheong L F, Gao Z. Quasi-parallax for nearly parallel



a.左视图



b.右视图



c.合成虚拟视图

图5 测试图像

Fig.5 Original left view(a),right view(b), and reconstructed virtual image(c) of Tsukuba

- frontal eyes-A possible role of binocular overlap during rapid locomotion [J]. International Journal of Computer Vision,2103,101(1):45-63
- [3] Singh M K,Gupta R D,Snehmani, et al.Effect of sensor modelling methods on computation of 3D coordinates from Cartosat-1 stereo data [J]. Geocarto International, 2016, 31(5):506-526
- [4] Ghassabeh Y A.On the convergence of the mean shift algorithm in the one-dimensional space [J]. Pattern Recognition Letters,2014,34(12):1423-1427
- [5] Hong S W,Shim W M.When audiovisual correspondence disturbs visual processing [J]. Experimental Brain Research,2016,234(5):1325-1532
- [6] Comaniciu D, Meer P. Mean shift: A robust approach toward feature space analysis [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2002,24(5):604-619
- [7] Schmeing M, Jiang X Y. Faithful disocclusion filling in depth image based rendering using superpixel-based inpainting [J]. IEEE Transactions on Multimedia, 2015, 17(12):2160-2173
- [8] Verhaeghe P P, Van der Bracht K, Van de Putte B. Discrimination of tenants with a visual impairment on the housing market: Empirical evidence from correspondence tests [J]. Disability & Health Journal, 2015, 9(2):234-238
- [9] Gong M, Yang Y H. Fast stereo matching using reliability based dynamic programming and consistency constraints [C] // Proceedings of International Conference on Computer Vision, 2003:610
- [10] Jin C M, Hong J. Intermediate view synthesis for multi-view 3D displays using belief propagation-based stereo matching [C] // Third 2008 International Conference on Convergence and Hybrid Information Technology, 2008: 919-924

A stereo matching algorithm based on disparity estimation and virtual viewpoint rendering

MEI Yong^{1,2} CHEN Shuyuan² WANG Youquan²

1 Jiangsu Key Laboratory of Meteorological Observation and Signal Processing, Nanjing University of Information Science & Technology, Nanjing 210044

2 School of Electronic & Information Engineering, Nanjing University of Information Science & Technology, Nanjing 210044

Abstract Aiming at the problem of holes and false contour in virtual view rendering technology, a new stereo matching algorithm is presented in this paper to obtain dense disparity map to complete the virtual view synthesis. Only two images are required for this method, which are segmented using mean shift algorithm. A good view of the disparity map and an effective interpolation method are basis for successful virtual viewpoint drawing. The weighted multi-window stereo matching method is used to obtain the initial disparity map, and the cross checking is employed to filter false match points and occlusion parts. The obtained dense disparity map is then view interpolated and noise filtered to finish the multiple point virtual view drawing. Experimental results with test images like Tsukuba provided by Middlebury dataset show that the algorithm effectively improves the quality of virtual view images.

Key words stereo matching; view interpolation; cross checking; dense disparity map; virtual view draw