光电技术及应用

DOI：10.16818/j.issn1001－5868.2021033005

基于引导滤波及视差图融合的立体匹配

陆明军*，叶　兵

（合肥工业大学 电子科学与应用物理学院，合肥 230009）

摘　要： 立体匹配是双目视觉领域的重要研究方向。为在保证图片纹理区域匹配精度的同时降低弱纹理区域的误匹配率，提出一种基于引导滤波及视差图融合的立体匹配方法。首先，根据图像颜色相似性将图片划分为纹理较丰富区域和弱纹理区域。接着，分别采用不同参数的引导滤波进行代价聚合及视差计算，得到两张视差图。然后依据纹理区域划分的结果对获得的两张视差图进行融合。最后，通过左右一致性检测、加权中值滤波等视差优化步骤得到最终视差图。对 Middlebury 测试平台上标准图像对的实验结果表明，该方法在 6 组弱纹理图像上的平均误匹配率为 9.67%，较传统引导滤波立体匹配算法具有更高的匹配精度。

关键词： 双目立体视觉；立体匹配；引导滤波；视差图融合

中图分类号： TP391　**文章编号：** 1001－5868(2021)06－0931－05

Stereo Matching Based on Guided Filtering and Disparity Map Fusion

LU Mingjun，YE Bing

（School of Electronic Science and Applied Physics，Hefei University of Technol.，Hefei 230009，CHN）

Abstract： Stereo matching is an important research direction in the field of binocular vision. In order to ensure the matching accuracy of image texture regions and reduce the mismatch ratio of weak texture regions. A stereo matching method is proposed based on guided filtering and disparity map fusion. First，the image is divided into areas with rich texture and areas with weak texture according to the similarity of image colors. Secondly，two disparity maps are obtained by cost aggregation and disparity calculation by using guided filtering with different parameters. Then，the two obtained disparity maps are merged according to the result of texture region division. Finally，the final disparity map is obtained through disparity optimization steps such as left and right consistency detection and weighted median filtering. Experiments performed on standard image pairs on the Middlebury test platform show that the average mismatch rate of this method on 6 sets of weak texture images is 9.67%，the matching accuracy is much higher than that of the traditional guided filter stereo matching algorithm.

Key words： binocular stereo vision；stereo matching；guided filtering；disparity map fusion

0　引言

作为计算机视觉领域的重要组成部分，双目立体视觉被广泛应用于三维重建[1]、生物医学[2]、无人驾驶[3]等领域。立体匹配是双目视觉的关键步骤，从左右视角的图像中找到相互匹配的像素点，计算像素点的视差值进而可获得对应点的深度信息。通过对典型立体匹配算法的研究和总结，Scharstein 等[4]将立体匹配分为匹配代价计算、代价聚合、视差计算、视差精化等四个步骤。

代价聚合对立体匹配的结果起关键作用，近年来，基于滤波器的代价聚合方法引起广泛关注。

收稿日期：2021－03－30.

基金项目：安徽省自然科学基金项目(J2014AKZR0032).

* **通信作者：**陆明军

Yoon 等[5]将双边滤波器用于代价聚合，取得了不错的效果，但其缺点是计算复杂度较高，不利于实际应用。Hosni 等[6]将计算复杂度与窗口大小无关的引导滤波用于代价聚合，在获得更好的边缘保持效果的同时大大提升了计算效率。但上述基于滤波的立体匹配方法均未考虑不同纹理对匹配结果的影响，使其在弱纹理图像上表现较差。

基于以上分析，本文提出基于引导滤波及视差图融合的立体匹配方法，通过纹理划分对纹理较丰富区域及弱纹理区域区别对待，从而得到两张不同的视差图，并依据纹理划分的结果对两张图进行融合，使其在保证纹理较丰富区域匹配精度的同时在弱纹理区域也获得了更为精确的视差图。

1 算法描述

本文算法主要分为纹理区域划分、匹配代价计算、代价聚合、视差计算、视差融合、视差精化六步。整个流程如图 1 所示。

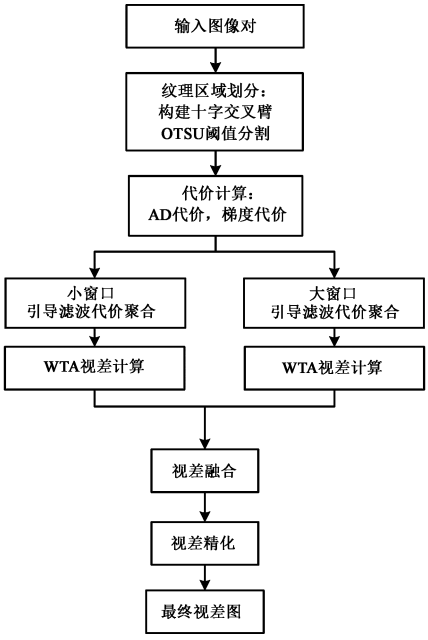


图 1 本文算法流程图

1.1 纹理区域划分

通过构建十字交叉臂与 OTSU(最大类间方差法)二值分割，将原始图片分为纹理较丰富区域和弱纹理区域。

十字支撑的自适应窗口被广泛应用于代价聚合[7-8]。自适应支撑窗口的最短臂长值能够反映像素点的区域特性，离边缘越近的点最短臂长越短[9]。故可对像素点构建十字交叉臂并根据对应的最短臂长值判断其周围纹理信息。本文采用 Mei 等[10]的方法来为图像的每个像素点构建十字交叉臂，十字交叉臂的构建规则如图 2 所示。从待构建像素点 p 出发进行十字臂的延展，当不满足设置的限定条件时，停止延展，限制条件为

$$\begin{cases} L \leqslant L_2 \\ \max\limits_{c\in R,G,B} | I_c(p) - I_c(t) | \leqslant \tau_1 \\ \max\limits_{c\in R,G,B} | I_c(p) - I_c(t) | \leqslant \tau_2, \text{if} \quad L_1 \leqslant L \leqslant L_2 \\ \max\limits_{c\in R,G,B} | I_c(t) - I_c(t^*) | \leqslant \tau_3 \end{cases} \tag{1}$$

式中，L_2 为允许的最大臂长，τ_1，τ_2 为中心点 p 与臂上点之间颜色差的阈值容限，其中，τ_2 的取值比 τ_1 小，目的是在臂长较长时降低阈值容限，避免引入不相关的点，同时为保证臂展跨过边界，使用第四个条件来限制臂上相邻两点的颜色差值范围。

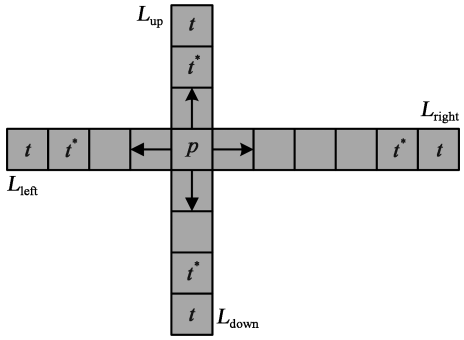


图 2 十字交叉臂

通过求取每点四个臂长的最短臂长值，可得到一张最短臂长图。对于获得的最短臂长图，使用 OTSU 阈值分割算法对其分割后得到二值图像，整个过程如图 3 所示。图 3(a)为 Middlebury 测试平台提供的数据集中 Teddy 的左图；图 3(b)为所构建的最短臂长图，为便于显示，图 3(b)中所有像素点的灰度值取为 8 倍最短臂长值；图 3(c)为对最短臂长图经过分割后得到的二值图。可以看出，通过阈值分割，将原图像中弱纹理区域和纹理较丰富区域做了很好地区分。



(a) 数据集中图 (b) 最短臂长图 (c) 二值图

图 3 纹理区域划分

1.2 匹配代价计算

代价计算用于衡量左右图像对应像素间的相似度，传统的基于 AD 与梯度融合的代价函数可表示为

$$\begin{cases} C_1(p,d)=\min\left[\dfrac{1}{3}\sum_{c\in R,G,B} I_c(p)-I'_c(p_d),\tau_c\right] \\ C_2(p,d)=\min[\nabla_x I(p)-\nabla_x I'(p_d),\tau_g] \\ C(p,d)=\alpha\cdot C_1(p,d)+(1-\alpha)\cdot C_2(p,d) \end{cases} \tag{2}$$

式中，α 为固定权重系数，$C_1(p,d)$ 为 AD 代价，$C_2(p,d)$ 为梯度代价，$C(p,d)$ 为融合后的总代价，τ_c，τ_g 为截断阈值。

受文献[9]的启发，本文根据像素点的最短臂长值，自适应调整 AD 与梯度的权重。具体为

$$\begin{cases} \beta=1-\exp\left(-\dfrac{1}{h_{\min}}\right) \\ C(p,d)=\beta\alpha C_1(p,d)+(1-\beta\alpha)C_2(p,d) \end{cases} \tag{3}$$

式中，β 为权重系数 α 的调整系数，$h_{\min}$ 为像素点对应的最短臂长值。像素点距离图像边缘越远，对应的最短臂长越大，β 的取值越小，从而 AD 代价项权重越小，梯度代价项的权重越大。因此，在弱纹理区域，通过权重调整系数 β 的作用，使得表现较差的 AD 代价项权重变小，更多地使用梯度代价项来进行代价计算。

1.3 引导滤波代价聚合

引导滤波器是一种很好的线性滤波器，其规则是假定在一个窗口内，滤波器的输出与输入满足线性关系：

$$q_i=\alpha_k I_i+b_k,\quad \forall i\in\omega_k \tag{4}$$

式中，$\boldsymbol{I}$ 为引导图像，$\boldsymbol{q}$ 为滤波器的输出代价矩阵，I_i 和 q_i 为像素点 i 处的对应值。a_k 与 b_k 为窗口中心为 k 时的局部窗口的线性系数。假设 $\boldsymbol{p}$ 为初始代价矩阵，则通过最优化求解可得：

$$a_k=\frac{\frac{1}{|\omega|}\sum_{i\in\omega_k} I_i p_i-\mu_k\bar{p}_k}{\sigma_k^2+\varepsilon} \tag{5}$$

$$b_k=\bar{p}_k-a_k\mu_k \tag{6}$$

其中，μ 和 σ^2 分别表示 $\boldsymbol{I}$ 在局部窗口中的均值和方差，ε 为正则化参数，用以调整 a_k 的值，防止其过大。因任意一个像素均存在于好几个窗口中，故通过求取均值的方式可得到任意像素的滤波系数。从而，任一像素点的代价滤波结果可表示为

$$\begin{cases} q(i)=\bar{a}_i\cdot I(i)+\bar{b}_i \\ \bar{a}_i=\dfrac{1}{|\omega|}\sum_{k\in\omega_i} a_k \\ \bar{b}_i=\dfrac{1}{|\omega|}\sum_{k\in\omega_i} b_k \end{cases} \tag{7}$$

对式(4)求导可知：输出代价矩阵的梯度完全由引导图像的梯度信息确定，a_k 的值决定了图像的边缘保持效果，对于弱纹理区域需要更大的 ε 来平滑图像。同时，匹配代价聚合是假定窗口内的像素的视差是平滑的，对于图像弱纹理区域需要更大的局部窗口来包含足够多的纹理信息。本文采用两个不同大小的窗口和不同的 ε 分别对初始代价进行引导滤波代价聚合，以满足不同纹理区域的要求。参数的设定规则为

$$\begin{cases} r_2=\mathrm{Floor}(r_1+\lambda_1 t) \\ \varepsilon_2=\varepsilon_1\cdot(1+\lambda_2 t) \end{cases} \tag{8}$$

其中，r_1 和 ε_1 为小窗口引导滤波的窗口大小及正则化系数，r_2 和 ε_2 为大窗口引导滤波的窗口大小及正则化系数，t 为上文中 OTSU 分割得到的分割阈值，λ_1，λ_2 为固定系数。

1.4 视差计算

对于代价聚合后的两组匹配代价，分别采用 WTA(Winner-Take-All)算法对待匹配像素点 p 选择最小匹配代价所对应视差值作为初始视差，表达式为

$$d_p=\underset{d\in D}{\mathrm{argmin}}[C(p,d)] \tag{9}$$

式中，D 代表视差搜索范围内所有可能的视差值。

1.5 视差融合

通过视差计算得到两张初始视差图，根据图像划分的结果对两张视差图进行融合。融合规则为

$$F_d=Disp_1\cdot T_1+Disp_2\cdot T_2 \tag{10}$$

式中，F_d 表示融合后的视差图，$Disp_1$ 为采用小窗口的引导滤波所得视差图，$Disp_2$ 为采用大窗口的引导滤波所得视差图。T_1 和 T_2 分别表示 OTSU 二值分割后的结果，T_1 在纹理较丰富区域取 1，弱纹理区域取 0，T_2 在弱纹理区域取 1，纹理较丰富区域取 0，即小窗口引导滤波所得的视差图为纹理较丰富区域提供视差值，大窗口引导滤波所得的视差图为弱纹理区域提供视差值。

1.6 视差精化

由于遮挡等因素，融合后的视差图通常还存在异常匹配点，需要通过视差精化对其进行处理。首先通过左右一致性检测找出异常匹配点，然后从异

常匹配点分别向左和向右搜索第一个可靠匹配点所对应的视差值中较小的一个作为异常匹配点的视差值，最后经过加权中值滤波得到最终视差图。

2 实验结果分析

2.1 算法性能验证

为验证本文算法的有效性，采用传统引导滤波算法(GF)与本文算法对 Middlebury 测试平台上的四组标准图像以及 2006 数据集中的六幅弱纹理图像进行实验，实验均采用相同的后处理过程。实验中的参数设置如表 1 所示。

使用 Tsukuba，Venus，Teddy，Cones 四组标准图像对来验证算法在纹理较丰富图像上的效果。实验结果如表 2 所示，其中，Nocc，All，Disc 分别表示非遮挡区域、所有区域和视差不连续区域。实验中的误差阈值设置为 1，即将实验所得结果与真实视差比较，将误差超过一个像素的点认定为误匹配点。从表中数据可以看出，较传统引导滤波立体匹配算法，本文算法在四组标准图像对上的平均误匹配率为 5.59%，较传统 GF 算法有所降低，说明本文算法在纹理较丰富图像上保留了传统引导滤波立体匹配算法的良好性能。

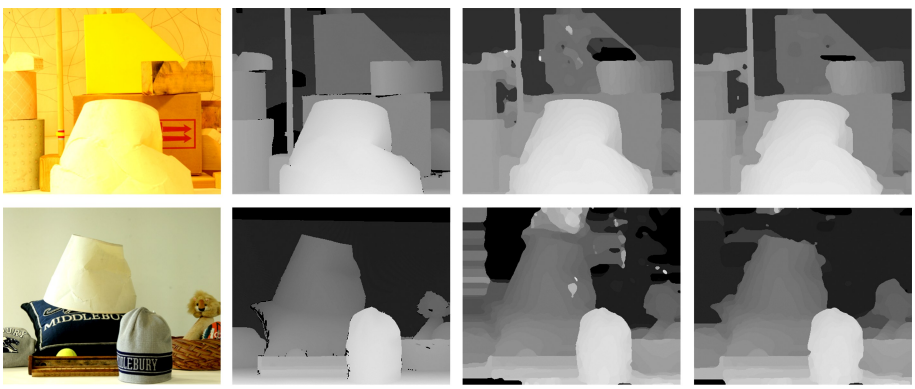
表 1 实验参数设置

参数	L_1	L_2	τ_1	τ_2	τ_3	α	τ_c	τ_g	ε_1	r_1	λ_1	λ_2
数值	17	34	20	6	6	0.11	0.027 45	0.007 84	0.000 1	9	0.6	0.3

表 2 不同算法的误匹配率(%)

算法	Tsukuba			Venus			Teddy			Cones			Avg
	Nocc	All	Disc	Nocc	All	Disc	Nocc	All	Disc	Nocc	All	Disc	
GF	1.86	2.17	7.83	0.18	0.45	2.43	6.97	12.53	16.40	2.73	8.53	7.79	5.82
Ours	1.85	2.15	7.83	0.17	0.45	2.39	6.72	11.39	14.57	2.84	8.69	8.09	5.59

图 4 为 GF 算法与本文算法在 Lampshade1，Lampshade2，Midd1，Midd2 两组弱纹理图像上所得视差图的对比，图 4(a)从上到下依次为 Lampshade1，Lampshade2，Midd1，Midd2 的原图像，图 4(b)为真实视差图，图 4(c)为 GF 算法所得视差图，图 4(d)为本文算法所得视差图。可以看出，传统 GF 算法在弱纹理区域出现大块空洞，而本文算法获得了较准确的视差值。



(a) 原图 (b) 真实视差图 (c) GF 算法视差图 (d) 本文算法视差图

图 4 不同方法所得视差图

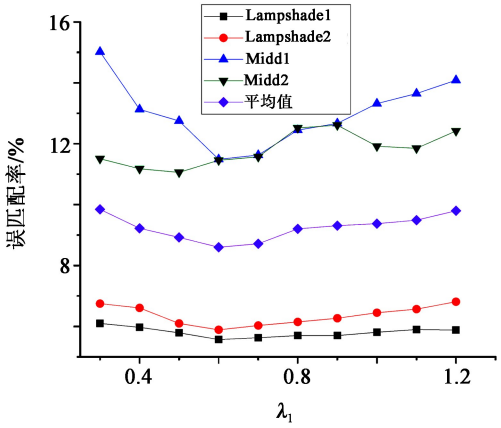
表 3 列出了两种算法在六组弱纹理图像上的误匹配率。可以看出，本文算法的平均误匹配率仅为 9.67%，而传统的引导滤波立体匹配算法的平均误匹配率为 25.2%。可见，本文算法在弱纹理图像上表现更好。

表 3 不同算法的误匹配率(%)

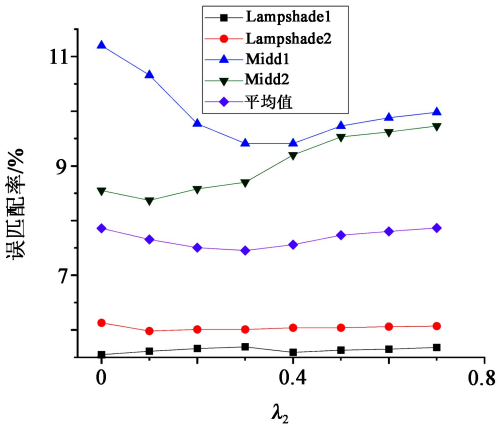
算法	Lampshade1	Lampshade2	Midd1	Midd2	Monopoly	Plastic	Avg
GF	8.24	18.74	39.30	31.22	19.34	34.39	25.20
Ours	5.69	6.01	9.41	8.70	5.93	22.33	9.67

2.2 参数分析

为验证实验参数 λ_1，λ_2 对实验结果的影响，分别采用不同的 λ_1，λ_2 对 Lampshade1，Lampshade2，Midd1，Midd2 四组弱纹理图像进行实验，实验结果如图 5 所示。可以看出，当 λ_1 取 0.6，λ_2 取 0.3 时，四组弱纹理图像的平均误匹配率最低。故本文将 λ_1 设置为 0.6，λ_2 设置为 0.3。



(a)　不同 λ_1 值对误匹配率的影响



(b)　不同 λ_2 值对误匹配率的影响

图 5　参数 λ_1，λ_2 对实验结果的影响

3　结论

本文提出了一种基于引导滤波及视差图融合的立体匹配方法。通过构建最短臂长图及阈值分割将原图分为两块区域分别进行针对性处理，利用视差图融合的方法将不同区域的结果进行融合。实验结果表明，相较传统引导滤波立体匹配算法，本文算法能更好地适应纹理不同的图像对立体匹配的影响，有效提高了弱纹理图像的匹配精度。

参考文献：

[1] Shen S H. Accurate multiple view 3D reconstruction using patch-based stereo for large-scale scenes[J]. IEEE Trans. on Image Processing, 2013, 22(5): 1901-1914.

[2] 苏　赫. 微创手术机器人图像系统设计方法与空间映射策略研究[D]. 天津：天津大学，2016.
Su H. Design method and space mapping of minimally invasive surgical robot image system [D]. Tianjin: Tianjin University, 2016.

[3] Xiao X W, Guo B X, Li D R, et al. Multi-view stereo matching based on self-adaptive patch and image grouping for multiple unmanned aerial vehicle imagery [J]. Remote Sensing, 2016. 8(2): 89-119.

[4] Scharstein D, Szeliski R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms[J]. Inter. J. of Computer Vision, 2002, 47(1-3): 7-42.

[5] Yoon K J, Kweon I S. Adaptive support-weight approach for correspondence search[J]. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2006, 28(4): 650-656.

[6] Hosni A, Rhemann C, Bleyer M, et al. Fast cost-volume filtering for visual correspondence and beyond [J]. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2013, 35(2): 504-511.

[7] 闫　利，王　芮，刘　华，等. 基于改进代价计算和自适应引导滤波的立体匹配[J]. 光学学报，2018，38(11)：257-267.
Yan L, Wang R, Liu H, et al. Stereo matching method based on improved cost computation and adaptive guided filter[J]. Acta Optica Sinica, 2018, 38(11): 257-267.

[8] 孔令寅，朱江平，应三丛. 基于引导图像和自适应支持域的立体匹配[J]. 光学学报，2020，40(9)：86-98.
Kong L Y, Zhu J P, Ying S C. Stereo matching based on guidance image and adaptive support region[J]. Acta Optica Sinica, 2020, 40(9): 86-98.

[9] 王云峰，吴　炜，余小亮，等. 基于自适应权重 AD-Census 变换的双目立体匹配[J]. 工程科学与技术，2018，50(4)：153-160.
Wang Y F, Wu W, Yu X L, et al. A stereo matching system with the adaptive weight AD-Census [J]. Adv. Engin. Sciences, 2018, 50(4): 153-160.

[10] Mei X, Sun X, Zhou M, et al. On building an accurate stereo matching system on graphics hardware[C]// IEEE Inter. Conf. on Computer Vision Workshops, 2011: 467-474.

作者简介：

陆明军(1991—)，男，安徽省合肥市人，硕士生，研究方向为激光技术与应用；

叶　兵(1961—)，男，安徽省合肥市人，博士，教授，硕士生导师，主要研究方向为激光技术与应用，以及电子电路设计。