

# 一种改进的基于图割的立体匹配算法\*

刘亚竹<sup>1</sup>, 李 逵<sup>1</sup>, 狄红卫<sup>1,2</sup>

(1.暨南大学 光电工程研究所, 广东 广州 510632;

2.暨南大学 光电信息与传感技术广东普通高校重点实验室, 广东 广州 510632)

**摘要:** 为了提高立体匹配算法的精确度, 提出了一种结合图割与 Mean Shift 图像分割的立体匹配算法。首先用 Mean Shift 算法对参考图像进行图像分割, 得到分割后的标记图, 然后将分割信息结合到图割算法的能量函数中, 最后用改进的能量函数和图割算法得出致密的视差图。实验结果表明, 提出的算法具有更高的精确度和更好的边缘特征。

**关键词:** 立体匹配; 图割; Mean Shift; 视差

中图分类号: TP391

文献标识码: A

文章编号: 1674-7720(2012)07-0043-03

## An improved stereo matching algorithm based on graph cuts

Liu Yazhu<sup>1</sup>, Li Kui<sup>1</sup>, Di Hongwei<sup>1,2</sup>

(1. Institute of Opto-Electronic Engineering, Jinan University, Guangzhou 510632, China;

2. Key Laboratory of Opto-Electronic Information and Sensing Technologies of Guangdong Higher Education Institutes, Jinan University, Guangzhou 510632, China)

**Abstract:** In order to improve the accuracy of the stereo matching algorithm, a stereo matching algorithm based on graph cuts and Mean Shift image segmentation is proposed. Firstly, the reference image is segmented using Mean Shift algorithm and the label image has been got. Then the segmentation information is integrated into energy function of graph cuts algorithm. Finally, a dense disparity map is yielded using the improved energy function and graph cuts algorithm. Experimental results show that the proposed method has better accuracy and better edge feature.

**Key words:** stereo matching; graph cuts; Mean Shift; disparity

立体匹配是计算机视觉中的一个研究热点, 是立体视觉中最关键的步骤之一。目前常见的立体匹配算法分为局部算法和全局算法。局部算法从局部(如以像素为中心的窗口)得出视差, 这类算法运算量小, 但精确度不高。全局算法将能量最小化思想引入到立体匹配中, 首先定义一个能量评价函数, 然后通过全局优化算法实现能量最小化, 进而确定最终的匹配结果。这类算法精确度较高, 但运算量大。ROY 和 COX 首次将图割方法引入到立体匹配算法中<sup>[1]</sup>, 用图割法实现能量最小; BOYKOV Y 等提出了  $\alpha$ - $\beta$  交换和  $\alpha$  扩张两种有效的基于图割的算法<sup>[2]</sup>, 通过多次对以两种标记建立的图求最小割, 达到了求全局最小能量的目的, 而且简化了算法复杂度, 提高了运算速度; KOLMOGOROV 改进了网络的构造方

法<sup>[3]</sup>, 提出了基于二值变量的图的构造, 考虑了多种情况的邻接边, 提高了算法速度, 并达到了很好的精确度。由于图割算法具有处理组合优化问题的良好性能, 使视差图的质量得到了明显的提高<sup>[4]</sup>。BLEYER 等提出基于图像分割的立体匹配算法框架<sup>[5]</sup>, 图像分割充分利用彩色图像的颜色信息将图像分割成不同区域, 有效地克服了视差图边界模糊问题, 较好地处理了大的低纹理区域。

本文将图割算法与 Mean Shift 图像分割算法相结合, 并应用到立体匹配中, 得到了一种效果良好的立体匹配算法。首先用 Mean Shift 算法对图像进行分割, 然后将分割信息加入到能量函数中, 再根据能量函数构建图, 最后用扩张算法迭代地求出最小割。实验结果表明, 该算法结合了两种算法的优点, 实现了更高的匹配精确度。

\* 基金项目: 广州市科技计划重点项目 (2011J4300037)

1 基于二值变量的图的构造

一个双终端图  $G=\{V, E\}$  包含两个终端, 源  $s$  和汇  $t$ , 顶点集合  $V=\{v_1, v_2, \dots, v_n, s, t\}$ 。  $G$  上的每一个切割都切断一条终端边 (与终端相连的边), 将图分为  $S, T$  两部分, 因此每个切割可以用  $n$  个二值变量  $x_1, x_2, \dots, x_n$  来表示, 每个二值变量对应一个图中的一个节点, 切割的代价等同于图  $G$  的能量  $E(x_1, \dots, x_n)$ 。 令:

$$x_i = \begin{cases} 0, & v_i \in s \\ 1, & v_i \in t \end{cases} \quad (1)$$

可将能量函数表示为二值变量的形式:

$$E(x_1, x_2, \dots, x_n) = \sum_i E^i(x_i) + \sum_{i < j} E^{ij}(x_i, x_j) \quad (2)$$

由于图具有可加性, 对于这样的能量形式可以对每项分别建图, 然后将子图合并在一起。 首先, 对于  $E^i$ , 它只依赖于一个变量  $x_i$ 。 如果  $E^i(0) > E^i(1)$ , 那么就增加边  $(s, v_i)$ , 边的权值为  $E^i(1) - E^i(0)$ , 如图 1 所示; 否则就增加边  $(v_i, t)$ , 边的权值为  $E^i(0) - E^i(1)$ , 如图 2 所示。 这种情况下构造出来的图都表示  $E^i$ , 但常量不同, 前者的常量为  $E^i(0)$ , 后者的常量为  $E^i(1)$ 。

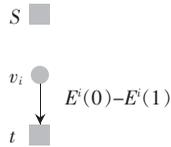


图 1  $E^i(0) > E^i(1)$  所构成的  $E^i$  图

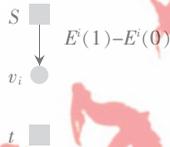


图 2  $E^i(0) < E^i(1)$  所构成的  $E^i$  图

其次, 对于  $E^{ij}$ , 它依赖于两个变量  $x_i$  和  $x_j$ 。 由于  $x_i, x_j$  为二值变量,  $E^{ij}$  可表示为表 1 所示的 4 种形式。

表 1  $E^{ij}$  的 4 种形式

$E^{ij}(0,0)$	$E^{ij}(0,1)$
$E^{ij}(1,0)$	$E^{ij}(1,1)$

令  $E^{ij}(0,0)=A, E^{ij}(0,1)=B, E^{ij}(1,0)=C, E^{ij}(1,1)=D$ , 表 1 又可转化为:

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = A + \begin{bmatrix} 0 & 0 \\ C-A & C-A \end{bmatrix} + \begin{bmatrix} 0 & D-C \\ 0 & D-C \end{bmatrix} + \begin{bmatrix} 0 & B+C-A-D \\ 0 & 0 \end{bmatrix} \quad (3)$$

在式 (3) 中, 第 1 项为常函数, 因此不需要增加边; 第 2 和第 3 项只依赖于一个变量  $x_i$  或  $x_j$ , 因此可用前面  $E^i$  的构造方法; 第 4 项与  $x_i, x_j$  两个变量都有关, 需要在图中增加边  $(v_i, v_j)$ , 边的权值为  $B+C-A-D$ 。  $E_{i,j}$  构成的图如图 3 所示。 这种方法没有增加辅助节点, 因而产生的图比较小, 有利于更快地计算最小割。

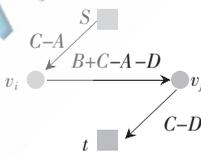


图 3  $E^{ij}$  图

2 构造能量函数

本文构造的能量函数形式为:

$$E(f) = E_{\text{data}}(f) + E_{\text{smooth}}(f) + E_{\text{vis}}(f) \quad (4)$$

其中,  $E_{\text{data}}(f)$  为数据项, 用来约束对应匹配像素具有一致的强度;  $E_{\text{smooth}}(f)$  为平滑项, 用来约束邻域像素具有平滑的视差;  $E_{\text{vis}}(f)$  为可见项, 用来约束匹配像素不违反可

见条件。

(1) 数据项  $E_{\text{data}}(f)$  为 RGB 三原色的匹配代价之和<sup>[8]</sup>, 它可以降低因图像采样的离散化而产生的误差。

$$C(p, q) = \begin{cases} I_{R_{\text{min}}} - I_{Lp}, & I_{Lp} < I_{R_{\text{min}}} \\ I_{Lp} - I_{R_{\text{max}}}, & I_{Lp} > I_{R_{\text{max}}} \\ 0, & \text{其他} \end{cases} \quad (5)$$

其中,  $(p, q)$  为左右视图中匹配的两像素,  $I$  代表亮度,  $C$  代表匹配代价。

$$I_{R_{\text{min}}} = \min\left\{\frac{I_{R_{\text{left}}} + I_{R_{\text{left}}}}{2}, \frac{I_{R_{\text{right}}} + I_{R_{\text{right}}}}{2}, \frac{I_{R_{\text{up}}} + I_{R_{\text{up}}}}{2}, \frac{I_{R_{\text{down}}} + I_{R_{\text{down}}}}{2}, I_{R_{\text{d}}}\right\} \quad (6)$$

$$I_{R_{\text{max}}} = \max\left\{\frac{I_{R_{\text{left}}} + I_{R_{\text{left}}}}{2}, \frac{I_{R_{\text{right}}} + I_{R_{\text{right}}}}{2}, \frac{I_{R_{\text{up}}} + I_{R_{\text{up}}}}{2}, \frac{I_{R_{\text{down}}} + I_{R_{\text{down}}}}{2}, I_{R_{\text{d}}}\right\} \quad (7)$$

其中, 下标 left、right、up 和 down 分别代表像素的 4 邻域。 为了减小前景膨胀效应, 对匹配代价采用截断处理:

$$C'(p, q) = \min\{C(p, q), T_{\text{cut}}\} \quad (8)$$

(2) 平滑项  $E_{\text{smooth}}(f)$  由图像分割信息得出。 首先用 Mean Shift 算法对左、右视图进行分割, 得到标记图  $lab$ , 若邻域像素视差一致, 则平滑代价为 0; 若视差不一致, 则  $p, q$  应属于不同的标记区域。 若标记一致, 则应加大“惩罚”。

$$E_{\text{smooth}}(p, q) = \begin{cases} 0, & f_p = f_q \\ \lambda_1, & f_p \neq f_q \text{ 且 } lab_p = lab_q \\ \lambda_2, & f_p \neq f_q \text{ 且 } lab_p \neq lab_q \end{cases} \quad (9)$$

式中,  $p, q$  为同一视图中的邻域像素,  $f_p, f_q$  分别为  $p, q$  点的视差,  $lab_p, lab_q$  为  $p, q$  点的标记。

(3) 若满足可见约束, 则可见项  $E_{\text{vis}}(f)$  为 0, 否则可见项为无穷大:

$$E_{\text{vis}}(f) = \begin{cases} \infty, & \{<p, f(p)>, <q, f(q)>\} \in I_{\text{vis}} \\ 0, & \text{其他} \end{cases} \quad (10)$$

3 建图与求最小割

本文采用  $\alpha$  扩张算法来计算图的最小割, 所有像素的视差分为  $\alpha$  和  $\bar{\alpha}$  两种情况来考虑, 正好对应双终端图的汇点和源点, 而图中的节点则可用二值变量来描述。 视差为  $\bar{\alpha}$  的像素参与构建图, 用上面描述的建图方式和提出的能量函数建图, 通过求网络最小割完成像素视差向  $\alpha$  的转变, 即  $\alpha$  扩张。  $\alpha$  的取值范围为  $0 \sim k$ ,  $k$  为最大视差。  $\alpha$  取完所有视差值即完成一次迭代, 其中取值顺序可以是随机的。 每完成一次迭代计算一次能量, 直到能量不再变小, 算法终止。

4 实验结果与分析

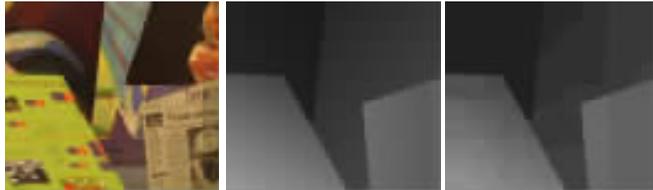
根据以上算法进行了仿真实验, 实验对象采用 Middlebury 立体图像库中的 Tsukuba 和 Venus 图像<sup>[6]</sup>, 实验结果分别如图 4、图 5 所示。

表 2 为误匹配像素百分比 (PBM) 值的比较<sup>[9]</sup>。 从表 2 可以看出, 由于本文算法在图割算法的基础上结合了 Mean Shift 图像分割, 在能量函数的平滑项中加入了图像分割的信息, 而图像分割的信息能反映出视差跳变的边缘, 因此该算法取得了更好的结果。



(a)左图 (b)真实视差图 (c)本文算法得到的视差图

图4 Tsukuba 图像匹配结果



(a)左图 (b)真实视差图 (c)本文算法得到的视差图

图5 Venus 图像匹配结果

表2 误匹配像素百分比(PBM)比较

算法	Tsukuba			Venus		
	无遮挡区域	总的错误	不连续区域	无遮挡区域	总的错误	不连续区域
GC+occ	1.19	2.01	6.24	1.64	2.19	6.75
Multi Cam GC	1.27	1.99	6.48	2.79	3.13	3.60
Segm+visib	1.30	1.57	6.92	0.79	1.06	6.76
本文算法	0.70	1.18	6.67	0.057	0.128	2.21

本文提出了一种基于图割与 Mean Shift 图像分割算法的立体匹配算法,在图割算法的基础上构造了新的能量函数,融入经 Mean Shift 图像分割后的信息,根据新的能量函数构建基于二值变量的图,用  $\alpha$  扩张算法迭代地求出最终视差。实验结果表明,该算法有效提高了算法的精确度,且视差图的边缘特征更好。

#### 参考文献

- [1] ROY S, COX I. A maximum-flow formulation of the N-camera stereo correspondence problem [C]. Proceedings of International Conference Computer Vision, 1998: 492-499.
- [2] BOYKOV Y, VEKSLER O, ZABIH R. Fast approximate energy minimization via graph cuts[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001, 23(11): 1222-1239.
- [3] KOLMOGOROV V, ZABIH R. Multi-camera scene reconstruction via graph cuts [A]. European Conference on Computer Vision, 2002,3:82-96.
- [4] BOYKOV Y, KOLMOGOROV V. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004,26(9):1124-1137.
- [5] 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.
- [6] <http://vision.middlebury.edu/stereo/>. 2011-01-01.

(收稿日期:2011-01-06)

#### 作者简介:

刘亚竹,女,1985年生,硕士研究生,主要研究方向:图像与视频处理。

李逵,男,1978年生,硕士,主要研究方向:嵌入式系统。

狄红卫,男,1969年生,博士,教授,硕士生导师,主要研究方向:信号与信息处理。