

利用块几何约束及视差概率的立体匹配算法^{*}

翟振刚^{1,2+}, 陆耀¹, 赵红^{1,3}

¹(北京理工大学 计算机科学技术学院 智能信息技术北京市重点实验室,北京 100081)

²(防空兵指挥学院,河南 郑州 450052)

³(河北大学 数学与计算机科学技术学院,河北 保定 071002)

Stereo Matching Algorithm Using Segment Geometric Constraint and Disparity Probability

ZHAI Zhen-Gang^{1,2+}, LU Yao¹, ZHAO Hong^{1,3}

¹(Beijing Key Laboratory of Intelligent Information Technology, School of Computer Science and Technology, Beijing Institute of Technology, Beijing 100081, China)

²(Air Defense Forces Command Academy, Zhengzhou 450052, China)

³(College of Mathematics and Computer Science, Hebei University, Baoding 071002, China)

+ Corresponding author: E-mail: zz_gang@bit.edu.cn

Zhai ZG, Lu Y, Zhao H. Stereo matching algorithm using segment geometric constraint and disparity probability. *Journal of Software*, 2010,21(11):2985–2998. <http://www.jos.org.cn/1000-9825/3695.htm>

Abstract: To handle the slanted or curved surfaces in stereo matching, the adjacent segment geometric constraint and the statistical information of disparity are adopted in a proposed algorithm. The segment geometric constraint is incorporated in a new global energy function to acquire the optimal plane of each segment. The statistical information of disparity is adopted to find the reliable disparity pixel and a reliable segment. The constraints between the segments and pixels are used to estimate the disparity of the unreliable pixels. Experiments are performed on the classical images with a larger disparity range, more slanted surfaces, and surface with less texture and show the effectiveness of the proposed algorithm.

Key words: stereo matching; disparity; slanted surface; curved surface; geometric constraint; statistical information

摘要: 为了解决倾斜表面或曲面的匹配问题,提出了一种基于图像分割块之间的几何约束和视差值的概率分布信息的视差估算方法.在一个全局能量函数中增加了图像分割块之间的几何约束项,通过计算匹配能量得到分割块的最优视差平面.为了确定可信像素和可信分割块,利用了视差的概率分布信息.同时,利用了分割块之间的几何约束和分割块内像素之间的约束来估计不可信像素点的视差值.用包含大视差范围、更多倾斜表面、曲面和弱纹理表面等典型图像对所建议的算法进行测试,实验结果表明,该方法对于存在倾斜表面和曲面的立体视差计算是有效的.

关键词: 立体匹配;视差;倾斜表面;曲面;几何约束;统计信息

中图分类号: TP391 文献标识码: A

* Supported by the Beijing Key Discipline Construction Program of China (北京市重点学科建设规划项目)

Received 2008-11-26; Revised 2009-04-27; Accepted 2009-07-17

立体匹配是计算机立体视觉中一个重要的研究方向,近几年来有许多文献^[1-4]都在研究这个问题.根据所采用优化方法的不同,立体匹配算法可以分为两大类:全局匹配算法^[5-8]和局部匹配算法^[9,10].虽然在文献[11]中提到了许多算法,但是在弱纹理区域^[12]、深度不连续^[13]和遮挡区域^[14,15]情况下,立体匹配算法仍面临着很多难题.此外,倾斜平面和曲面的匹配在研究中很少涉及,许多算法基本上采用了正面平行表面模型假设.即在进行匹配时,认为某一分割块内像素的视差值是一个恒定值.而现实中许多场景却不符合这种假设,所以采用当前的算法,在对具有许多倾斜表面和曲面的场景进行立体匹配时会产生误差,容易产生阶梯波形的视差图.即使文献[16]采用了截取平滑项,也不能很好地解决倾斜平面立体匹配问题,如图 1(c)所示.

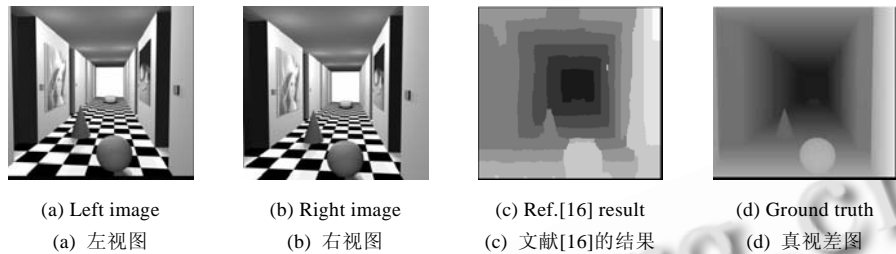


Fig.1 Illumination of slanted plane in stereo matching

图 1 立体匹配中倾斜平面问题

为了提高倾斜表面和曲面的匹配精度,文献[17]利用分段线性曲面模型(piecewise linear surface model)表示空间几何结构.根据视差空间的梯度信息,把采用 Mean Shift 方法得到的图像分割块进一步细分,然后排除像素点的匹配奇异点,采用平面拟合方法得到视差值.该方法是把大的曲面分割成非常小的曲面小块,近似认为这些小块曲面具有同一视差值,即小块内的所有像素具有相同的视差值.

在文献[18]中,为了解决倾斜表面问题,Ogale 和 Aloimonos 提出了基于扫描线的立体匹配算法(scan-line based matching algorithm).在文献[12,19,20]中,Li 和 Zucker 针对倾斜表面和曲面立体匹配问题提出了解决问题的方法.文献[19]能量函数中,融合了视差的一阶导数信息来解决倾斜表面的立体匹配问题;文献[12]利用了相邻像素点要满足几何一致性约束上下文几何信息(contextual geometric information);文献[20]在能量函数中加入像素的几何约束.文献[12,19,20]都是采用像素的上下文几何约束(pixel-wise contextual geometric constraint)信息,而本文采用分割块的上下文几何约束(segment contextual geometric constraint)信息和视差概率分布信息.在确定不可信像素的视差值时,采用的处理方法和文献[12,19,20]的方法都不相同.本文采用了两级约束:分割块之间约束和分割块内像素间约束,而文献[12,19,20]只利用了像素之间的几何关系.在文献[20]中,采用线性截断代价函数(truncated linear cost function for smooth term)作为平滑项,也未能很好地解决倾斜平面的匹配问题.

本文利用分割块的上下文几何约束信息,即把邻域分割块之间的几何一致性约束强加给每一个分割块,将其作为全局能量函数中的一个数据项进行处理.为了消除奇异点(outlier),利用视差值概率分布信息判断可信像素点和可信分割块.然后利用全局能量函数,通过迭代的方法把可信分割块的视差信息传递给不可信分割块.本文主要有两个贡献:(1) 采用了图像分割块上下文的几何约束;(2) 使用视差概率统计信息得到视差值可信的像素点.

本文第 1 节概述本文算法.第 2 节采用 Mean Shift 方法进行彩色图像分割.在第 3 节,采用改进的基于窗体立体匹配算法得到初始视差值.第 4 节利用初始视差值的概率分布先验知识,采用概率统计的方法确定可信像素点和可信图像分割块.利用可信视差像素点,采用平面拟合的方法估计出能够表示整个场景结构的视差平面集.第 5 节利用分割块之间和分割块内像素之间的两个层次约束来估计不可信像素的视差值.第 6 节是实验结果,同时对本算法性能进行分析.第 7 节给出结论.

1 算法概述

算法流程如图 2 所示:输入图像为两幅经过矫正的彩色图像,使用 Mean Shift 方法分割图像,利用改进的局部立体匹配方法得到初始视差值,用视差值概率统计信息以及分割块的几何约束确定可信像素点和可信分割块,再根据分割块间和分割块内的约束确定不可信像素的视差值.

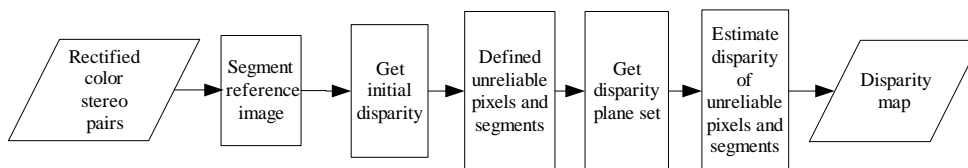


Fig.2 Block diagram of the proposed method

图 2 算法流程图

在本算法中,采用了改进的基于窗体局部立体匹配方法 SAD(sum of the absolute intensity differences)得到像素点的初始视差值.为了得到更准确的初始视差值,在局部的能量方程中增加了梯度信息.实验结果表明,采用修正后的基于窗体局部立体匹配方法提高了立体匹配的精度.

在得到初始视差值之后,经过交叉验证(cross-checking)和视差统计信息,确定可信像素和可信分割块.在推理阶段,主要是从两个层次上进行推理:分割块之间推理和分割块内的像素之间推理.在分割块之间进行推理时,把分割块上下文的几何一致性约束作为全局能量函数的一项进行处理.在每次迭代时,对于选定的分割块,在邻域分割块的几何约束下,通过能量函数最小化,从视差平面集合中选择一个最优视差平面,而不是从邻域图像分割块中选择一个视差平面作为最优解.在分割块内像素之间进行推理时,利用了视差值概率分布先验知识,在得到分割块最优视差平面后,重新确定可信视差点.最后,利用新的可信像素点视差信息得到分割块内不可信像素点的视差值.

2 彩色图像分割

本文采用 Mean Shift^[21]算法对参考图像进行过分割(over-segmentation),把参考图像分割成颜色相同各个均质(homogenous color)的图像块,参数设置为: $h_s=10, h_r=7, M=30$. h_s 是空间带宽(spatial bandwidth), h_r 是颜色带宽(range or color bandwidth), M 是图像分割块的最小像素数(minimum segment size).在实验部分,对参数进行了分析讨论.

3 初始视差

通过改进基于窗体立体匹配算法,得到像素初始视差值.局部能量函数采用 SAD,同时增加了一阶梯度和二阶梯度信息.

$$E_{data} = T_{das} + \omega(T_{dagd1} + T_{dagd2}) \tag{1}$$

其中,

$$T_{das} = \sum_{(i,j) \in N(i,j)} |I_l(i,j) - I_r(i+d,j)| \tag{2}$$

$$T_{dagd1} = \sum_{(i,j) \in N(i,j)} |\nabla_x I_l(i,j) - \nabla_x I_r(i+d,j)| + \sum_{(i,j) \in N(i,j)} |\nabla_y I_l(i,j) - \nabla_y I_r(i+d,j)| \tag{3}$$

$$T_{dagd2} = \sum_{(i,j) \in N(i,j)} |\nabla_x^2 I_l(i,j) - \nabla_x^2 I_r(i+d,j)| + \sum_{(i,j) \in N(i,j)} |\nabla_y^2 I_l(i,j) - \nabla_y^2 I_r(i+d,j)| \tag{4}$$

其中, E_{data} 是总的能量; T_{das} 是测量两个像素的相似程度的数据项; T_{dagd1}, T_{dagd2} 分别是图像一阶梯度和二阶梯度; $N(i,j)$ 是一个以 (i,j) 为中心像素的 5×5 的窗体,在计算 T_{dagd1}, T_{dagd2} 时与文献[21]不同,本文采用了 $N(i,j)$ 是邻域范围内的所有像素; ∇_x, ∇_y 是水平和垂直方向的一阶导数; ∇_x^2, ∇_y^2 是水平和垂直方向的二阶导数; $I(i,j)$ 是图像的颜色

信息,本文把图像RGB通道的相似性进行了简单求和,并且求取其均值作为 $I(i,j)$ 的值;权重 ω 是一个正的恒定值.

4 可信视差判断及视差平面集估计

分割块视差平面是利用每个分割块内可信像素点的视差信息,采用平面拟合的方法得到的.所以,首要任务就是确定可信视差点.

4.1 可信视点的判断

首先采用交叉验证(cross checking)^[22]的方法得到比较可信的视差点,过滤掉遮挡像素点和弱纹理区域像素点. D_l 和 D_r 分别表示左视差图和右视差图.

交叉验证公式如下:

$$|D_l(x,y)+D_r(x+d(x,y),y)|<1 \quad (5)$$

其中, (x,y) 和 $(x+d(x,y),y)$ 分别是左、右图像中对应匹配点.由唯一性约束可以推导出公式(5),唯一性约束是确保左右图像中像素点存在一一对应关系.如果公式(5)不成立,那么该像素的视差值为不可信点;相反,其可信.

然后,利用颜色分割的信息和视差概率分布信息进一步确定可信和不可信像素点.与文献[23,24]类似,再次利用视差概率分布信息判断出可信分割块和不可信分割块.在文献[23]中,把高度可信像素点称为地面控制点GCPs(ground control points),其主要是通过最优视差值的能量和次最优视差值的能量进行比较得到GCPs.在文献[24]中,主要是通过4种先验滤波的方法得到GCPs.而在本文中,是通过视差概率统计信息,确定可信视差值.

在每个图像分割块中,如果不可信像素数和同一图像分割块中所有像素数的比值 λ_{rs} 大于阈值 δ_s ,那么整个分割块中所有像素的视差值都不可信,这个分割块就是不可信分割块;如果 λ_{rs} 小于阈值 δ_s ,那么分割块就是可信分割块,然后按照下面的方法进一步判断其内的可信像素.

在每个可信的图像分割块 s_i 中,如果具有相同视差值 d 的像素个数和同一分割块 s_i 中的像素个数的比值 λ_{red} 小于阈值 δ_{se} ,那么在这一分割块 s_i 中,具有视差值 d 的所有像素点就是不可信点;相反,就是可信点.

两次判断公式如下:

$$\lambda_{rs} = \frac{NB_{su}}{NB_s} > \delta_s \quad (6)$$

$$\lambda_{red} = \frac{NB_{ed}}{NB_s} < \delta_{se} \quad (7)$$

其中, NB_s 是分割块中所有像素数; NB_{su} 是这一分割块中不可信的像素数; NB_{ed} 是在同一分割块中,具有某一个视差值 d 像素数; λ_{rs} 和 λ_{red} 是两个比例值; δ_s 和 δ_{se} 是大于0的阈值,其取值可以采用固定阈值,也可以采用自适应阈值.本文采用了固定阈值.

4.2 视差平面集的估计

利用可信分割块中的可信像素点来估计视差平面,在每个可信分割块中,利用最小二乘法来估计初始视差平面.

$$d=a_i \cdot x+b_i \cdot y+c_i \quad (8)$$

其中, (a_i,b_i,c_i) 是图像分割块 s_i 视差平面的法向量, (x,y) 是图像分割块 s_i 中可信像素点的坐标, d 是像素 (x,y) 的视差值.在每一个图像分割块的视差平面估计出来之后,把具有相同视差平面并且相邻的图像分割块融合在一起进行重新估计.经过不同次数迭代处理之后,所有这些不同的视差平面就代表了整个场景结构,并构成一个视差平面集合 $PS=\{(a_i,b_i,c_i),i=0,\dots,N\}$.

5 估计不可信像素的视差值

为了确定不可信像素点的正确视差值,利用了图像分割块之间的几何约束和分割块内像素点间的视差概率信息约束两个层次约束,估计不可信像素点的最优视差值.

5.1 分割块间几何约束

对于分割块之间的平滑约束,除了要考虑相邻块之间的颜色和相关性约束以外,还需要考虑它们之间的几何约束,即视差平面的约束.如图 3 所示,这是一个俯视图, F 代表一个曲面, A 和 B 是相邻的图像分割块, $L1$ 和 $L2$ 是分别是图像块 A 和 B 的视差平面, AA' 和 BB' 分别是其法向量.为了让图像块之间保持平滑,必须考虑到它们之间的几何约束.

本文采用的方法是合作算法^[24,25]的核心思想,即采用迭代的方法和非线性处理的思路,把可信的分割块和像素的视差信息传递给不可信的分割块和像素.对每一分割块 s_i ,从视差平面集合 PS 中任意选择其中一个视差平面,计算等式(9)的能量,选择最优的一个视差平面作为分割块 s_i 的视差平面.在计算等式(9)时,其邻域分割块的视差平面保持不变.遍历每一分割块,直到满足迭代次数为止.

针对每个分割块 s_i ,利用邻域可信分割块的视差平面信息进行计算等式(9),而不考虑邻域内不可信的分割块的视差信息.如图 4 所示,图像分割块 $S1,S2,S3,S4,S5,S6,S7$,在每次迭代时,比如对图像分割块 $S4$ 进行计算,已知与其相邻的图像分割块 $S1,S2,S5$ 和 $S6$ 是可信分割块,而 $S3$ 和 $S7$ 是不可信分割块.在计算等式(9)时,考虑可信的分割块 $S1,S2,S5$ 和 $S6$ 的视差信息,而不考虑不可信的分割块 $S3$ 和 $S7$ 的视差信息.在这些可信图像分割块的视差平面保持不变的情况下,从视差平面集合中选择使等式(9)能量最小的视差平面,作为图像分割块 $S4$ 的视差平面.

如果在分割块 s_i 的邻域内没有可信的分割块,那么分割块 s_i 就先不计算.通过多次迭代,可以先确定分割块 s_i 邻域内不可信分割块的视差平面,把不可信分割块变成可信分割块,在下一次迭代中,利用邻域已经确定的分割块视差信息,得到分割块 s_i 的视差平面.如图 5 所示,针对分割块 $S1$,其相邻的分割块 $S2,S3,S4,S7,S10$ 都是不可信的分割块,而分割块 $S5,S6,S8,S9$ 是可信的分割块.我们可以先确定分割块 $S1$ 邻域中部分或者全部不可信分割块 $S2,S3,S4,S7,S10$ 的视差平面,邻域中部分或全部不可信的分割块就变成了可信分割块.然后在下一次的迭代过程中,就可以计算分割块 $S1$ 的视差平面,通过多次的迭代就得到最优视差平面.

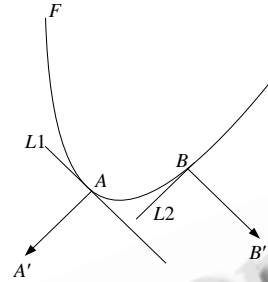


Fig.3 Illustration of the geometric constraint
图 3 几何约束的示意图

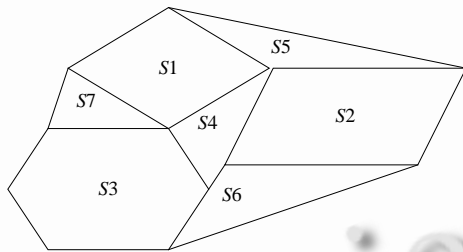


Fig.4 Geometric constraint in segment domain
图 4 分割块之间几何约束示例图

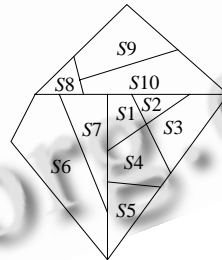


Fig.5 Iterative process in segment domain
图 5 分割块之间的迭代过程

R 是图像分割块的集合, PS 是视差平面的集合.图像分割块 $s_i \in R$ 所对应的视差平面 $f(s_i) \in PS$.能量函数如下:
$$E_{s_i}(f) = E_{s_i,data}(f) + E_{s_i,mismatch}(f) + E_{s_i,smoothness}(f) \tag{9}$$
其中, $E_{s_i}(f)$ 是图像分割块 s_i 的能量, $E_{s_i,data}(f)$ 是 s_i 图像分割块的数据项, $E_{s_i,mismatch}(f)$ 是图像分割块 s_i 没有匹配像素的惩罚项, $E_{s_i,smoothness}(f)$ 是图像分割块 s_i 的平滑惩罚项.

$$E_{s_i,data}(f) = \sum_{(x,y) \in N_r} T'_{data}(x,y) \tag{10}$$

其中, N_r 是图像分割块 s_i 内可信视差点的集合.

$$T'_{data} = T_{das} + \omega \cdot T_{dagd} \quad (11)$$

$$E_{s_i, mismatch}(f) = \omega_{mismatch} \cdot NB_{s_i, mismatch} \quad (12)$$

$$\begin{cases} E_{s_i, smoothness}(f) = \sum_{j \in N(i) \setminus i} (\omega_{smoothness} \cdot bolen(s_j, s_i) \cdot geo(s_j, s_i) \cdot colr(s_j, s_i)), f(s_i) \neq f(s_j) \\ E_{s_i, smoothness}(f) = 0, & \text{otherwise} \end{cases} \quad (13)$$

其中, $NB_{s_i, mismatch} = NB_{s_i, u}$ 是图像分割块 s_i 中不可信像素数, 而 $NB_{s_i, u}$ 是经过公式(5)~公式(7)判断之后得到的不可信像素. 其中, 分割块 s_j 是分割块 s_i 邻域中可信分割块, $\omega_{mismatch}$ 和 $\omega_{smoothness}$ 是对没有匹配像素和不连续像素的惩罚系数.

$$bolen(s_j, s_i) = \sum((p_j, p_i) | p_j \in s_j, p_i \in s_i) \quad (14)$$

像素 p_j 与 p_i 是 4 邻域关系, 并且 $s_i \neq s_j$. 其中, s_j 是 s_i 邻域中可信分割块.

$$Colr(s_j, s_i) = abs(mean(s_j) - mean(s_i)) + 0.5 \quad (15)$$

$$mean(s) = \left(\sum_{(x_i, y_j) \in s} ((R(x_i, y_j) + B(x_i, y_j) + G(x_i, y_j)) / 3) \right) / NB_s \quad (16)$$

像素 (x_i, y_j) 是图像分割块 s 中的一个像素点, $R(x_i, y_j), B(x_i, y_j), G(x_i, y_j)$ 是像素 (x_i, y_j) 的 RGB 通道的值, NB_s 是图像分割块中的所有像素数. 其中, 分割块 s_j 是分割块 s_i 邻域中可信分割块.

$$geo(s_j, s_i) = \|N_{s_j} - N_{s_i}\|^2 \quad (17)$$

其中, s_j 是分割块 s_i 邻域分割块中可信的分割块; N_{s_j} 和 N_{s_i} 是图像分割块 s_i, s_j 的视差平面法向量的归一化处理结果, 可以通过下面的公式计算出来:

$$N = \left(\frac{\partial d}{\partial x}, \frac{\partial d}{\partial y}, 1 \right) / \sqrt{\left(\frac{\partial d}{\partial x} \right)^2 + \left(\frac{\partial d}{\partial y} \right)^2 + 1}$$

其中, $\left(\frac{\partial d}{\partial x}, \frac{\partial d}{\partial y}, 1 \right)$ 是图像分割块视差平面的法向量, 主要是通过公式(8)计算得到.

$\|\cdot\|^2$ 是 2 阶范式, 公式(9)的时间复杂度为 $o(M^2)$, M 是 PS 集合中视差平面数. 公式(14)、公式(15)、公式(17)的时间复杂度为 $o(N^2)$, N 是分割块数.

算法.

1. 在每一次迭代过程中, 针对每一个分割块 S , 如果分割块 S 周围没有可信分割块, 转第 5 步; 否则, 执行第 2 步.
2. 如果分割块 S 是可信视差分割块, 则根据公式(9)计算匹配能量 E_1 , 如果分割块 S 是不可信视差分割块, 则其匹配能量 E_1 为无穷大.
3. 从视差平面集合 PS 中选取一个视差平面作为分割块 S 的视差平面, 根据公式(9)计算其能量 E_2 .
4. 比较匹配能量 E_1 和 E_2 : 如果 E_2 小于 E_1 , 更新分割块 S 的视差平面, 如果分割块 S 是不可信视差分割块, 则将其标示为可信的分割块, 把视差平面赋给分割块 S , 并把 E_2 的值赋给 E_1 , 然后返回第 3 步; 如果 E_2 大于 E_1 , 则转第 3 步.
5. 计算下一个分割块, 直到所有分割块都遍历一边, 然后进行下一次迭代.

本算法的时间复杂度是 $o(NM^2T)$, 其中, N 是分割块的数量, M 是 PS 集合中视差平面的多少, T 是迭代次数. 即在每一次的迭代中, 计算传递信息需要花费的时间为 $o(M^2)$, 而总共有 $o(N)$ 信息需要计算, 所以本算法的时间复杂度较高.

5.2 分割块内像素间的约束

利用分割块内像素间约束, 主要目的是进一步剔除奇异点. 在经过以上各步骤处理之后, 每一个图像分割块都得到了最优的视差平面. 然后, 利用分割块视差平面, 得到分割块中每一个像素视差值. 对每一个图像分割块而言, 利用不等式(7)进一步确定新的可信视差点. 再利用这些新的可信点, 采用平面拟合的方法得到分割块新的

视差平面,进而确定不可信视差的视差值.这就是在图像分割块内像素点之间的最优化处理过程.如图 6 所示,横坐标表示像素点,纵坐标表示像素点所对应的视差值.在像素层次上进行最优化处理时,首先经过不等式(7)得到不可信视差点,在图 6(a)中,用“*”表示不可信的 3 个像素点视差值.利用像素间约束处理之后,得到这些不可信像素点的视差值,如图 6(b)所示.

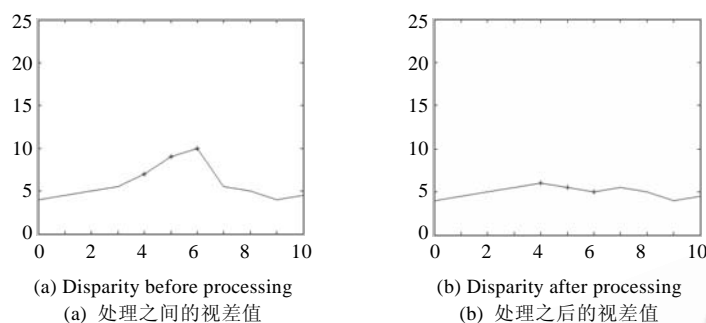


Fig.6 Constraint in pixel domain
图 6 分割块像素间的约束示例图

对于还没有视差值的不可信像素,利用公式(18)寻找合理的视差值作为像素(x,y)的视差值:

$$\arg \min_{(x',y') \in N(x,y) \setminus (x,y)} \text{abs}(I(x,y) - I(x',y')) \tag{18}$$

其中, $N(x,y)$ 是像素(x,y)的 6 邻域范围内所有可信视差像素点的集合.

6 实验结果及性能分析

在实验中,参数 $\omega=5, \delta_s=0.8, \delta_{se}=0.01, \omega_{occ}=5, \omega_{sm}=5$.迭代次数是 2.运行环境是 Pentium(R)D CPU3.4GHZ 主频,1G 内存,在 MATLAB 环境中运行.本文采用从 <http://vision.middlebury.edu/stereo> 下载的标准测试图像 Flowerpot,Wood1,Monopoly 和从文献[19,26,27]得到的具有代表性的测试图 Corridor,Parking meter.因为在这些图像中具有许多倾斜表面,并且非常具有挑战性.但是,Corridor,Parking meter 没有标准的 SCALE 因子,同时也受噪声的影响,所以只能进行定性分析比较,没法进行定量分析.在标准测试图像 Flowerpot,Wood1,Monopoly 上进行了定性和定量分析比较,并对本算性能进行了分析.

图 7 显示了实验中每一步得到的结果图,从左向右是参考图像、分割后的图像、可信视差点的视差图(黑点表示不可信点,非黑点是可信视差点)、本文算法的结果图和文献[16]算法的结果图.第 1 行是图像 Corridor^[19,27],具有 256×256 像素,并且具有 11 个像素的视差范围.第 2 行~第 4 行图像都是具有 80 个像素的视差范围的标准图像,第 2 行是大小为 370×443 的图像 Monopoly.第 3 行图像 Flowerpot 大小为 370×437.第 4 行是图像 Wood1,其大小是 370×457.这几幅测试图像都具有很多倾斜表面和曲面,并且具有很多弱纹理的区域,视察范围也比较大,所以具有较大的挑战性.图 8 是图 7 中测试图像的真实视差图.

利用均方根误差(root-mean-squared error,简称 RMSE)和误匹配像素点的百分比(percentage of bad matching pixels)^[11]对本文的算法和文献[16]的算法进行了定量比较,测试图像是具有多曲面,且具有 80 个视差值的图像 Flowerpot 和具有多倾斜平面且弱纹理比较多的图像 Wood1,Monopoly.表 1 显示了比较结果.

图 9 从迭代次数的角度分析本算法的性能,从中可以看出迭代次数对结果的影响,实验图像是 Flowerpot.图 10 从定量的角度分析图 9 的结果,图 10 绘出了由于迭代次数的不同所引起的误差变化情况(横坐标为迭代次数,纵坐标均方根误差和误匹配率).从结果可以看出,当迭代次数大于某一值时,误差急剧下降.由于采用的优化算法比较简单,将来可以采用较好的优化算法,比如图割(graph cuts)、信念传播(belief propagation)进行优化处理.

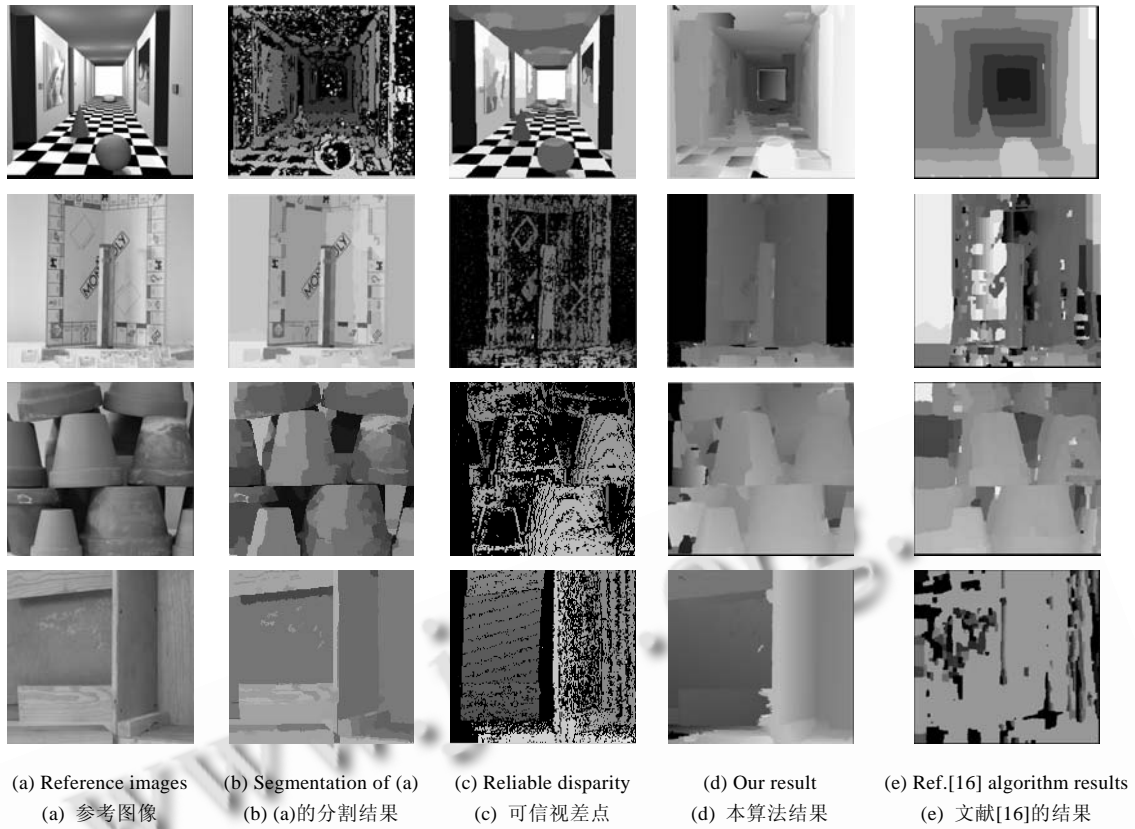


Fig.7 Results of our algorithm each step

图 7 本文算法的每一步结果

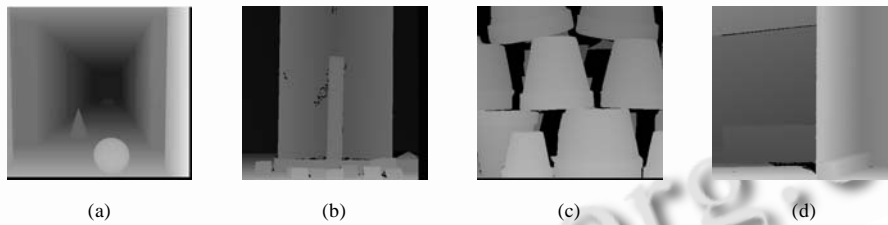


Fig.8 Ground truths of Fig.7

图 8 图 7 图像的真视差图

Table 1 RMSE and the percentage of the bad matching pixels

表 1 均方根误差和误匹配率

Algorithms	Monopoly			Flowerpot			Wood1					
	RMSE	Bad matching >0.5	Bad matching >1	Bad matching >2.5	RMSE	Bad matching >0.5	Bad matching >1	Bad matching >2.5	RMSE	Bad matching >0.5	Bad matching >1	Bad matching >2.5
Our algorithm	0.045 2	0.636 6	0.561	0.514 7	9.431	0.839	0.608 2	0.511 4	5.272 7	0.744 8	0.424 1	0.324 9
Ref.[16] algorithm	0.269	0.806 3	0.731 7	0.686 9	9.999	0.935 7	0.820 8	0.720 5	15.179 3	0.994	0.984 4	0.975 7

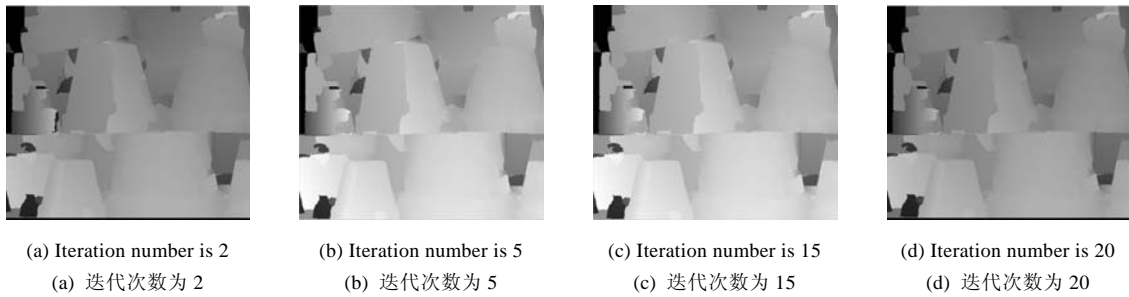


Fig.9 Different iteration number's result in our algorithm
图 9 不同迭代次数时的结果图

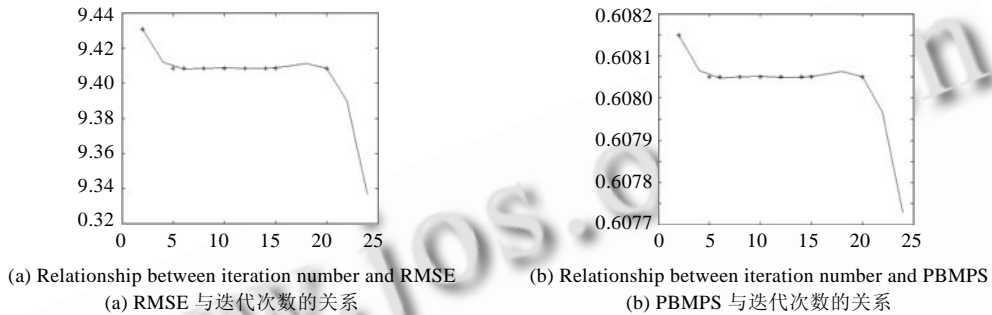


Fig.10 Relationship between iteration number and the RMSE or PBMPs on Flowerpot (error threshold=1)

图 10 当误差阈值为 1 时,在实验图 Flowerpot 上迭代次数与 RMSE,PBMPs 的关系

图 11 显示了添加二阶梯度信息和没有添加二阶梯度信息的结果图,测试图像为 Monopoly.没有添加二阶梯度信息的结果图,其均方根误差为 8.759,阈值大于 1 的误匹配率为 0.636 6;而在添加了二阶导数信息之后,其结果图的均方根误差为 8.707 5,阈值大于 1 的误匹配率为 0.634 7.可以看出,添加的二阶导数信息提高了匹配的准确性.

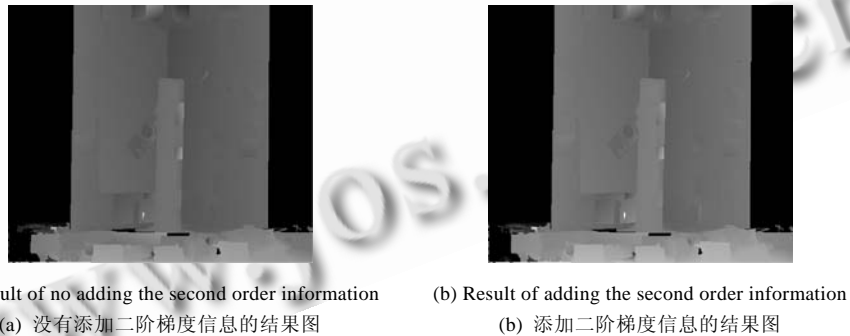


Fig.11 Influence of the second gradient information on the result (error threshold=1)
图 11 当误差阈值为 1 时,二阶梯度信息对匹配结果的影响

表 2 给出了本算法和文献[16]算法的运行时间,由于本算法是在 matlab 环境中运行,通过多次运行,得到一个平均的时间.本算法时间复杂度较高,主要因为大部分时间花费在几何约束信息的计算上.

Table 2 Procedure time on different tested images

表 2 在不同测试图像上的运行时间

	Corridor	Monopoly	Flowerpot	Wood1
Time (s)	3 360.780 400 5	48 226.654 801	32 747.576 682	5 217.029 391 5

图像初始分割的结果对立体匹配具有一定的影响,如果图像的每一个分割块太小,每一个图像分割块图像信息不足,就不容易得到准确的匹配点;如果分割块太大,就容易把两个不同视差的物体误认为是同一视差,也容易产生误匹配.在图 12 和图 13 中,分别从定性和定量的角度分析了在过分割到欠分割各个状态下最终视差图的精度.

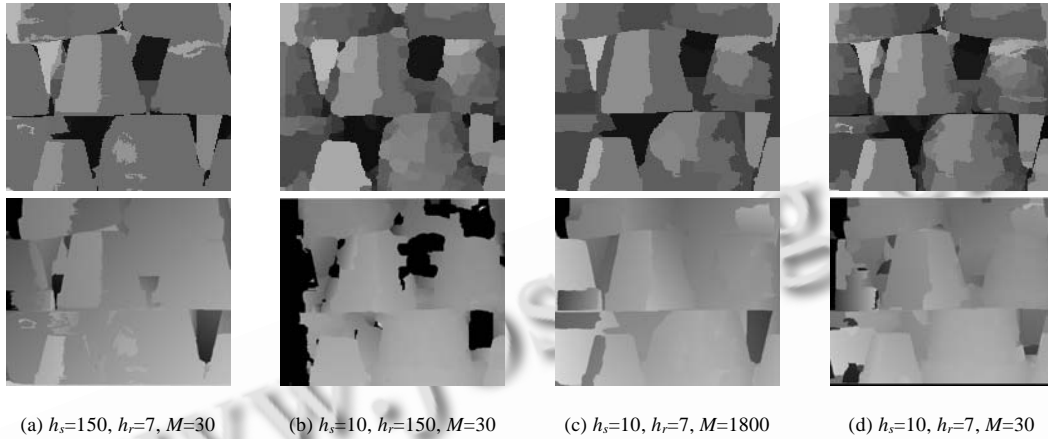


Fig.12 Disparity maps and segment results under the different segmental states

图 12 不同分割状态下的视差图和分割结果图

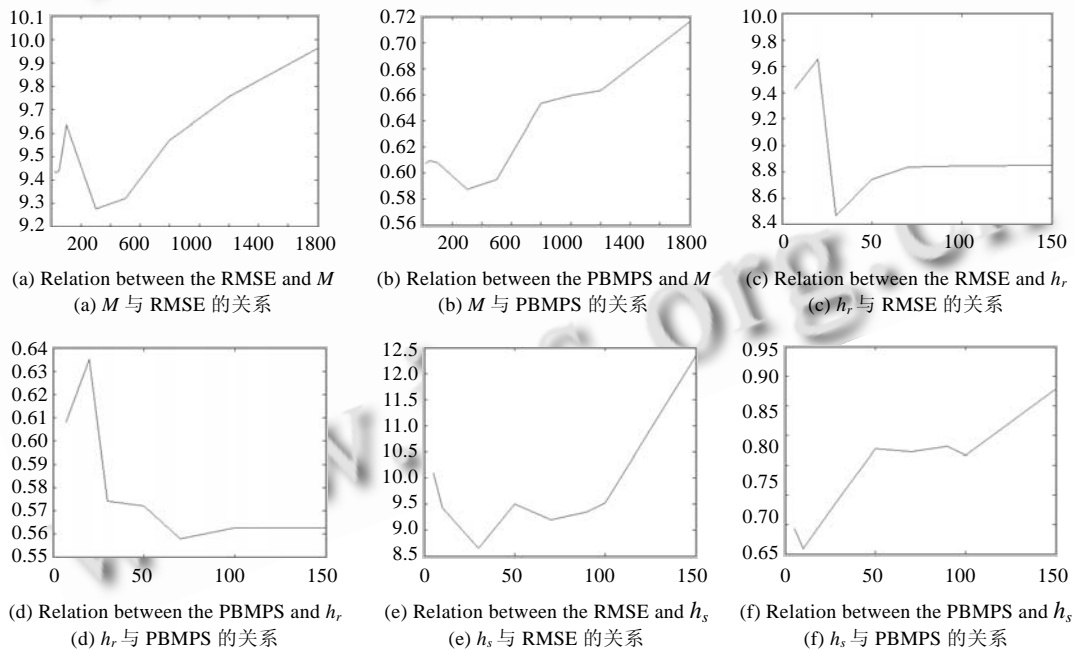


Fig.13 Relationship between RMSE, PBMPs and the different segmental states (error threshold=1)

图 13 当误差阈值为 1 时,不同分割状态与 RMSE,PBMPs 的关系

图 12 显示了不同分割状态下的视差图和分割结果图,不同的颜色代表不同的分割块,第 1 行是分割结果图,第 2 行是最后的视差图,真视差图是图 8(c).在图 13 中,横坐标是分别是分割的空间带宽、颜色带宽和分割块的最小像素数,纵坐标是 RMSE 和 PBMPs(error threshold=1),测试图像是 Flowerpot.从图中可以看出,不同的分割状态下,得到的精度不相同.针对公式(6)和公式(7)的两个阈值 δ_s, δ_{se} ,图 14 显示了其对均方根误差和误匹配率的影响.其中,图 14(a)和图 14(b)是阈值 δ_s 与 RMSE,PBMPs 的关系,图 14(c)和图 14(d)是 δ_{se} 与 RMSE,PBMPs 的关系.

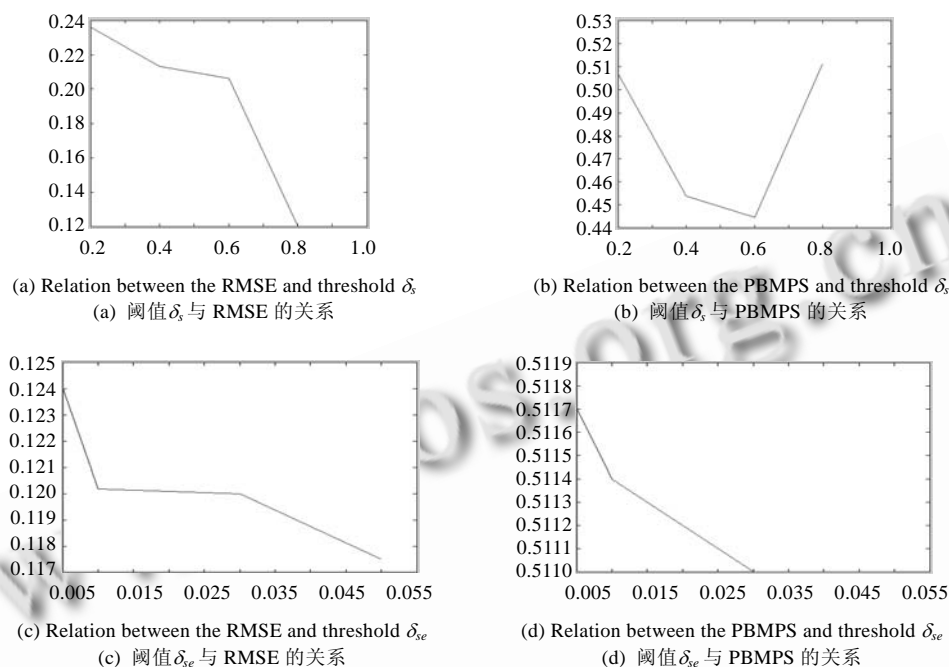


Fig. 14 RMSE, PBMPs (error threshold=1) under the different δ_s, δ_{se} values

图 14 阈值 δ_s, δ_{se} 对 RMSE,PBMPs(误差阈值=1)精度的影响

在图 15、图 16 中,本文的算法与文献[16,19]的算法以及其他算法的结果进行了比较.从结果可以看出:本文的算法比文献[16]的结果要好;与文献[19]的算法比较,在平滑性和边界的保持上比较具有优势,特别是在大块具有倾斜平面的场景中(如墙壁等)具有优势;而与其他算法,比如传统的 sum of squared difference(SSD)、图割(graph cuts)和协作算法(cooperative algorithm)的结果相比较,具有明显好的效果.但对于如图 16 中花的叶子等细小的物体,得到的结果不是很准确;而与其他算法,比如传统的 sum of squared difference(SSD)、图割和协作算法结果相比较,具有明显优势.

从所有实验结果可以看出,本文提出的算法具有较好的性能,特别是在图像场景中具有很多倾斜表面和曲面时,本算法得到的效果更好,如图 7 中的 Monopoly 和 Wood1.同时,在边界的保持上也具有明显的优势.但是,本算法对于狭小的物体或细小的物体,如图 16 中灌木的叶子等,立体匹配的准确性不是很好.这主要是由于本算法受到分割的影响,对狭小的物体进行分割之后,分割区域内元素较少,信息量不是很充分,所以引起的匹配精度不是很高.除此之外,本算法的复杂度也是比较高的,将来可以优化代码,提高运行时间.

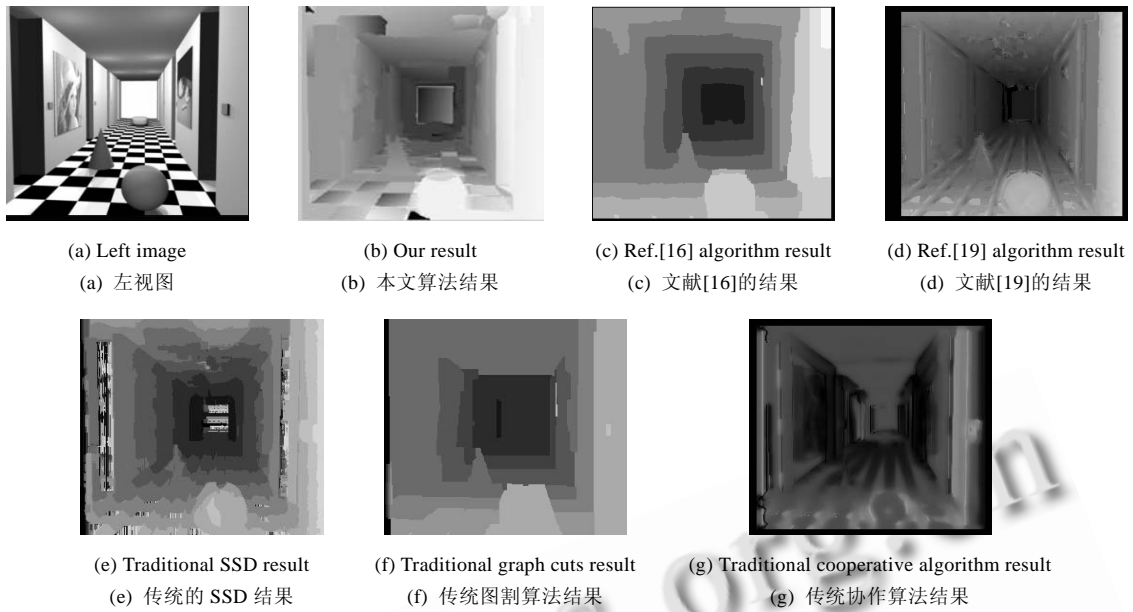


Fig.15 Comparison of different algorithms on Corridor

图 15 Corridor 测试图不同算法结果比较

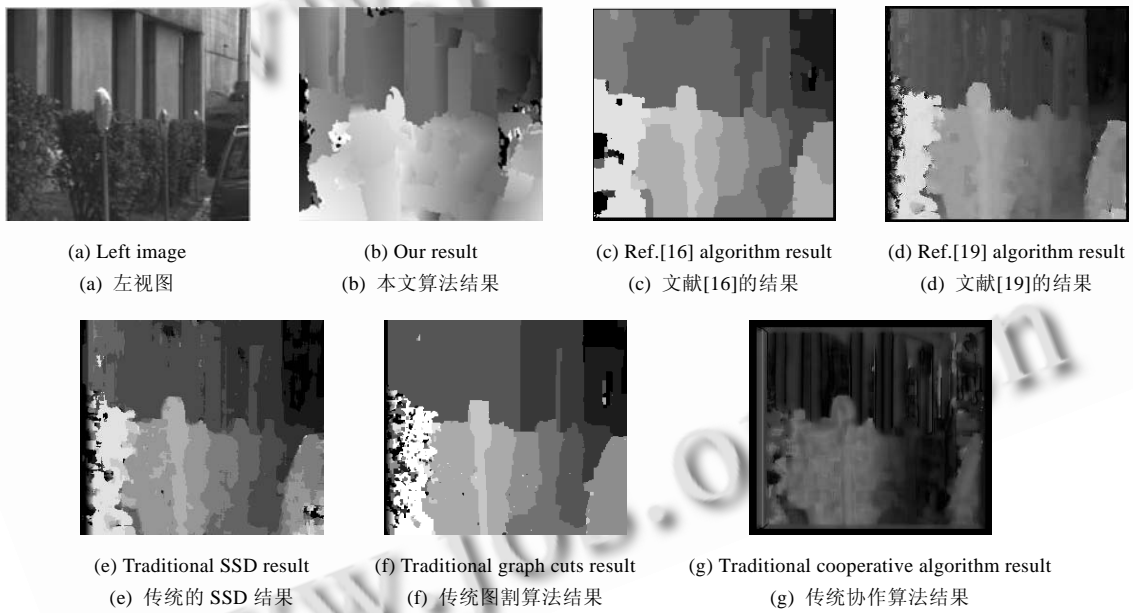


Fig.16 Comparison of different algorithms on Parking meter

图 16 Parking meter 测试图不同算法结果比较

7 结 论

在本文中,为了得到更准确的初始视差值,对局部的能量函数进行改进,添加了梯度信息.针对立体匹配中倾斜平面和曲面的匹配问题,本算法利用分割块的几何约束和视差概率分布信息,采用两级约束推理策略计算

像素的视差值.在分割块之间的推理过程中,采用了上下文分割块的几何约束,通过迭代的方法使每个分割块都得到最优视差平面.在分割块内像素点之间的推理过程中,利用视差概率分布信息准确地估计了不可信像素的视差值.在实验部分,利用具有更富有挑战性的图像数据以及国际上标准的挑战性图像数据进行测试,并且与其他相关算法进行定性和定量比较,并对本算法进行了性能分析.从比较实验结果可以看出,本算法提高了立体匹配的精度,但缺点是时间复杂度较高.由于本文的算法受到图像分割的影响,针对细小物体(如树叶等),效果不是很好.下一步我们将考虑如何提高细小物体的匹配精度,特别是具有曲面的细小物体.另外,在采用更好的优化算法以及提高算法的准确性,并且在判断可信点时,阈值的选择上也有探讨的空间.

致谢 感谢 Felzenszwalb 和 Huttenlocher 公开文献[16]的代码,从他们的网站上可以下载.同时感谢 Scharstein 和 Szeliski 提供了标准的国际测试图像数据.感谢匿名审稿专家教授和编辑部老师提出的宝贵修改意见.

References:

- [1] Hong L, Chen G. Segment-Based stereo matching using graph cuts. In: Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, Vol.1. IEEE Computer Society, 2004. 74–81.
- [2] Yang QX, Wang L, Yang RG, Stewenius H, Nister D. Stereo matching with color-weighted correlation, hierarchical belief propagation, and occlusion handling. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2009,31(3):492–504. [doi: 10.1109/TPAMI.2008.99]
- [3] Taguchi Y, Wilburn B, Zitnick CL. Stereo reconstruction with mixed pixels using adaptive over-segmentation. In: Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition. IEEE Computer Society, 2008. 1–8.
- [4] Bleyer M, Gelautz M. Graph-Based surface reconstruction from stereo pairs using image segmentation. In: Proc. of the SPIE—The Int'l Society for Optical Engineering, Vol.5665. 2005. 288–299.
- [5] Sun J, Zheng NN, Shum HY. Stereo matching using belief propagation. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2003,25(7):787–800. [doi: 10.1109/TPAMI.2003.1206509]
- [6] Xu L, Jia JY. Stereo matching: an outlier confidence approach. Proc. of the European Conf. on Computer Vision, 2008,4:775–787.
- [7] Chen W, Zhang MJ, Xiong ZH. Dense matching algorithm based on region boundary restriction via graph cuts optimization. Journal of Software, 2005,16(6):960–969 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/16/960.htm> [doi: 10.1360/jos160960]
- [8] Wen GJ. A global algorithm for straight line stereo matching based on feature grouping. Journal of Software, 2006,17(12): 2471–2484 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/17/2471.htm> [doi: 10.1360/jos172471]
- [9] Agrawal M, Davis L. Window-Based, discontinuity preserving stereo. In: Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, Vol.1. 2004. 166–173.
- [10] Li MX, Jia YD. Adaptive aggregation based cooperative stereo vision. Journal of Software, 2008,19(7):1674–1682 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/19/1674.htm> [doi: 10.3724/SP.J.1001.2008.01674]
- [11] Scharstein D, Szeliski R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. Int'l Journal of Computer Vision, 2002,47:7–42.
- [12] Li G, Zucker SW. Differential geometric consistency extends stereo to curved surfaces. In: Proc. of the European Conf. on Computer Vision. 2006. 44–57.
- [13] Sun J, Li Y, Kang S, Shum HY. Symmetric stereo matching for occlusion handling. In: Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, Vol.II. 2005. 399–406.
- [14] Wen GJ, Zhou XZ. Stereo matching and large occlusion detection based on disparity points. Journal of Software, 2005,16(5): 708–717 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/16/708.htm> [doi: 10.1360/jos160708]
- [15] Lin XY, Liu Y, Dai WZ. Study of occlusions problem in stereo vision. In: Proc. of the 7th IEEE World Congress on Intelligent Control and Automation. 2008. 5062–5067.
- [16] Felzenszwalb PF, Huttenlocher DP. Efficient belief propagation for early vision. Int'l Journal of Computer Vision, 2006,61(1): 55–79.

- [17] Oh JD, Ma SW, Kuo CCJ. Stereo matching via disparity estimation and surface modeling. In: Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition. 2007. 1–8.
- [18] Ogale AS, Aloimonos Y. Stereo correspondence with slanted surfaces: Critical implications of horizontal slant. In: Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition. 2004. 1568–1573.
- [19] Li G, Zucker SW. Stereo for slanted surfaces: First order disparities and normal consistency. In: Proc. of the Energy Minimization Methods in Computer Vision and Pattern Recognition. 2005. 617–632.
- [20] Li G, Zucker SW. Surface geometric constraints for stereo in belief propagation. In: Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, Vol.2. 2006. 2355–2362.
- [21] Comaniciu D, Meer P. Mean shift: A robust approach toward feature space analysis. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2002,24(5):603–619. [doi: 10.1109/34.1000236]
- [22] Klaus A, Sormann M, Karner K. Segment-Based stereo matching using belief propagation and a self-adapting dissimilarity measure. In: Proc. of the 18th Int'l Conf. on Pattern Recognition. 2006. 15–18.
- [23] Wei YC, Quan L. Region-Based progressive stereo matching. In: Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, Vol.1. 2004. 1106–1113.
- [24] Bobick AF, Intille SS. Large occlusion stereo. Int'l Journal of Computer, 1999,33(3):181–200.
- [25] Wang ZF, Zheng ZG. A region based stereo matching algorithm using cooperative optimization. In: Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition. 2008. 1–8.
- [26] Huang XF. Cooperative optimization for energy minimization: A case study of stereo matching. 2007. <http://front.math.ucdavis.edu/author/X.Huang.cs.CV/0701057>
- [27] Frohlinghaus T, Buhmann JM. Regularizing phase-based stereo. In: Proc. of the 13th Int'l Conf. on Pattern Recognition. 1996. 451–455.

附中文参考文献:

- [7] 陈旺,张茂军,熊志辉.基于区域边界约束和图割优化的稠密匹配算法.软件学报,2005,16(6):960–969. <http://www.jos.org.cn/1000-9825/16/960.htm> [doi: 10.1360/jos160960]
- [8] 文贡坚.一种基于特征编组的直线立体匹配全局算法.软件学报,2006,17(12):2471–2484. <http://www.jos.org.cn/1000-9825/17/2471.htm> [doi: 10.1360/jos172471]
- [10] 李鸣翔,贾云得.基于自适应聚合的立体视觉合作算法.软件学报,2008,19(7):1674–1682. <http://www.jos.org.cn/1000-9825/19/1674.htm> [doi: 10.3724/SP.J.1001.2008.01674]
- [14] 文贡坚,周秀芝.基于视差点的大遮挡检测和立体匹配方法.软件学报,2005,16(5):708–717. <http://www.jos.org.cn/1000-9825/16/708.htm> [doi: 10.1360/jos160708]



翟振刚(1979—),男,山西河津人,博士生,助教,主要研究领域为计算机视觉,模式识别,三维目标识别.



赵红(1979—),女,博士生,讲师,主要研究领域为图像超分辨率,模式识别.



陆耀(1958—),男,博士,教授,博士生导师,主要研究领域为计算机视觉,模式识别,人工智能.