

# 基于像元集的置信传递立体匹配

周秀芝<sup>1),2)</sup> 王润生<sup>1)</sup>

<sup>1)</sup>(海军航空工程学院青岛分院训练部虚拟仿真研究系, 青岛 266041)

<sup>2)</sup>(国防科技大学电子科学与工程学院 ATR 国防科技重点实验室, 长沙 410073)

**摘要** 为了提高立体匹配效率和克服处理区域的视差跳跃,提出了一种基于像元集的置信传递立体匹配方法。该方法首先以像素为基元,利用层次置信传递算法得到较为准确的初始视差;然后依次根据颜色和初始视差对参考图像进行分割,再利用分裂合并策略对分割后的像元集进行平面拟合,以消除颜色分割错误对匹配造成的影响;最后在拟合后的像元集空间,利用标准置信传递优化算法得到最终解。采用国际标准图像进行测试的实验结果表明,该方法的匹配效率和精度优于同类方法。

**关键词** 立体匹配 置信传递 图像分割 平面拟合

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2008)01-0506-07

## Pixel-set Based Stereo Matching by Using Belief Propagation

ZHOU Xiu-zhi<sup>1),2)</sup>, WANG Run-sheng<sup>1)</sup>

<sup>1)</sup>(VRSIM Laboratory Department of Training, Qingdao Branch of Naval Aeronautical Engineering Academy, Qingdao 266041)

<sup>2)</sup>(ATR Key Laboratory, School of Electronic Science and Engineering, National University of Defense Technology, Changsha 410073)

**Abstract** In order to improve the efficiency of stereo matching and resolve the problem caused by discontinuity of the disparities in one region, new pixel-set based stereo matching algorithm using belief propagation is proposed in this paper. Firstly, the initial disparity estimate is evaluated by hierarchical belief propagation in pixel domain. Secondly, the reference image is segmented in turn according to color and disparity information. Thirdly, the segmented pixel-sets are fitted to a set of planes based on the strategy of splitting and merging, which can eliminate the mistakes due to color segmentation. Finally, the ultimate disparity estimate is achieved in the pixel-set domain by employing standard belief propagation. Experiments on the international benchmark demonstrate that the performance of our algorithm is comparable to the state-of-the-art stereo algorithms on various data sets.

**Keywords** stereo matching, belief propagation, image segmentation, plane fitting

## 1 引言

立体匹配是近年来计算机视觉领域的研究重点之一,目前已提出了很多方法,Scharstein 和 Brown 对这些方法做了较为详细的分类<sup>[1,2]</sup>。实际上,立体匹配方法可归结为以下两类:基于窗口的局部方法和利用能量函数进行求解的全局方法。

基于窗口的局部方法<sup>[3~6]</sup>的特点是计算复杂度低,执行效率较高,但得到的视差误差较大,匹配精

度较低,难以满足那些对匹配精度要求较高的实际应用;而基于能量的全局方法则由于可以得到较为精确的视差值,因此受到越来越广泛的重视。

基于能量的全局方法中的能量函数一般是以下两部分能量之和:一部分是数据项,用来测量单个基元之间的对应性;另一部分是平滑项,用来惩罚相邻基元间视差的跳跃。构造完能量函数后,即可再利用优化算法来最小化或最大化能量函数和求得视差的最优解。目前在立体匹配中应用最为成功的优化算法是图切割(graph cut, GC)<sup>[7~11]</sup>和置信传递

收稿日期:2006-04-17;改回日期:2006-08-10

第一作者简介:周秀芝(1977~),女。2006年获国防科技大学博士学位,现为海军航空工程学院青岛分院讲师。研究方向为视景仿真、图像理解、图像配准。E-mail: luckyzhouzhou2004@yahoo.com.cn

(belief propagation, BP)算法<sup>[10,12-15]</sup>。匹配中所用到的基元有像元<sup>[8-10,12,13,16]</sup>和像元集两种,像元集是指诸如区域(region)<sup>[11,17,18]</sup>,块(patch)<sup>[7]</sup>等像元集合,由于它可以更好地描述实际场景结构,因此取得了较好的匹配结果。以像元集为基元的方法,一般是先将参考图像根据颜色进行分割先形成像元集,然后利用视差空间的一系列平面来近似像元集;接着求解与不同像元集对应的视差函数;最后在像元集域利用优化算法计算最终视差。由于分割后的基元数目远远小于图像的像素数,因此具有较高的效率。然而,因为此类方法一般是建立在同一区域内的视差满足连续性约束假设的基础上,故Deng等人指出,该类方法存在无法处理区域内部存在视差跳跃的缺点<sup>[7,11,18]</sup>。

对此,本文提出一种基于像元集的BP匹配算法。该算法首先根据颜色对左图像进行分割,同时利用分层置信传递算法(hierarchical belief propagation, HBP)<sup>[13]</sup>求解到图像的初始视差;然后对颜色分割后得到的区域进行平面拟合,同时根据求解到的视差和能量值的大小确定每个区域是否需进行分裂合并;最后以分割后的区域为基元,利用标准BP算法来得到图像的最终视差。算法的流程如图1所示。该方法具有以下特点:(1)视差初始值的求解不是使用平方差和最小等简单方法,而是用HBP求解到较为准确的视差值,这就为像元集的平面拟合提供了可靠的初值;(2)在像元集拟合阶段,不单单以分割的结果为依据,而是通过将视差初值和颜色分割结果相结合来得到拟合的像元集基元,若某个像元集内视差初值变化超过一定的阈值,则根据视差初值的概率分布来确定是否进行二次分割,若某个像元集与其相邻像元集的视差初值相同,则根据能量最小化原则确定这两个像元集是否合并,这样一种分裂合并策略可以消除由颜色初始分割造成的错误对匹配过程的影响;(3)利用BP算法在像元集域进行优化,不但可以提高效率,而且通过能量函数

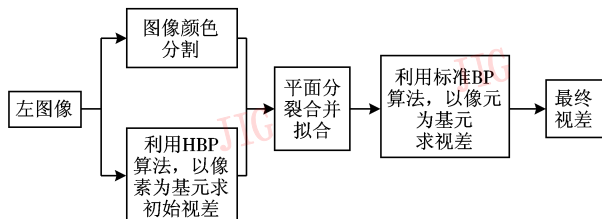


图1 算法实现过程

Fig. 1 The implementation of the proposed algorithm

平滑项中对边缘像素视差跳跃的惩罚来控制不同基元间的相互作用。

## 2 颜色分割和初始视差求解

### 2.1 颜色分割

一般情况下,视差发生变化的地方往往也是图像颜色发生变化的地方。Comanicu等人根据程度不同将分割分为欠分割、过分割和量化分割3种<sup>[19]</sup>。一般基于区域的匹配方法<sup>[11,17,18]</sup>都对图像进行了过分割,在满足效率要求的同时应尽量减少图像细节的丢失,以便保留视差变化的细节。即便这样,仍然不能避免部分视差发生变化的地方被分割到同一区域中,而且由于过分割后基元较多,因此不利于后续步骤的计算。本文采用欠分割的方法,利用较小的计算代价就可以得到视差变化的大致结构。当然这样会有一些视差变化的地方被分割到同一区域中,这个问题将结合初始视差进行二次分割来解决(详细叙述见第3节)。欠分割虽能够保留大部分的细节,但由于欠分割后仍然有很多很小的碎片区域,把它们作为一个区域基元是没有意义的,同时如果将一块小区域嵌在一块大区域中,那么就可以直接用大区域作为基元,因此需对分割后的图像进行去空洞处理和去碎片处理。表1中的第2行和第3行分别给出了网站<http://www.middlebury.edu/stereo>中提供的4幅标准图像过分割和欠分割后的区域数目,由表1可以看出,过分割后的区域数目几乎是欠分割的两倍。表1第4行给出了做过空洞和碎片处理后的区域数目,由此可以看出,处理后的区域数目大为减少。这样就可以减轻后续处理的计算负担。

表1 4幅图像过分割和欠分割后的像元集数目

Tab. 1 The number of pixel-groups for four standard data sets

分割类型	4幅图像分割后的像元集数目			
	Tsukuba	Sawtooth	Venus	Map
过分割	1 167	1 211	1 200	1 257
欠分割	651	656	568	223
欠分割处理后	336	343	260	74
像元集平面拟合后	258	241	195	56

### 2.2 初始视差计算

计算视差初始值的方法有很多,一般利用局部窗口,采用简单的测量,如平方差(squared Intensity difference, SD),绝对差(absolute intensity difference,

AD), 归一化互相关 (normalized cross-correlation, NCC) 等来得到视差初始值。这样尽管所需时间较少, 但得到的视差初始值往往不够精确。Felzenszwalb 提出的 HBP 算法对标准的 BP 算法进行了改进<sup>[13]</sup>, 大大提高了其效率。因此, 本文采用 HBP 算法来计算视差初值。下面先简单介绍一下马尔可夫网络和 BP 算法<sup>[12]</sup>。

马尔可夫网络是一个无方向的图结构。对一个  $N$  个节点的图结构而言, 每一个节点  $p_i$  都与可观察变量  $y_{p_i}$  ( $i = 1, 2, \dots, N$ ) 和隐藏变量  $x_{p_i}$  相关联。令  $\mathbf{X} = \{x_{p_i}\}$ ,  $\mathbf{Y} = \{y_{p_i}\}$ , 这种双马尔可夫网络的后验概率为

$$P(\mathbf{X} | \mathbf{Y}) = \prod_p \Phi_p(x_p, y_p) \prod_p \prod_{q \in N(p)} \Psi_{p,q}(x_p, x_q) \quad (1)$$

其中,  $\Phi_p(x_p, y_p)$  是节点  $p$  的局部证据,  $\Psi_{p,q}(x_p, x_q)$  是节点  $p$  和  $q$  之间的相容函数。BP 算法就是用来最大化式(1)的近似算法。在立体匹配中, 每个像素代表一个节点  $p$ ,  $\Phi_p(x_p, y_p)$  和  $\Psi_{p,q}(x_p, x_q)$  可分别用能量函数的数据项  $E_D(d_p)$  和平滑项  $E_S(d_p, d_q)$  的负指数来代替。实验是利用下式来计算  $E_D(d_p)$ <sup>[12]</sup>:

$$E_D(d_p, I) = -\ln(1 - e) \exp(-|C(d_p, p, I)|/\sigma) + e \quad (2)$$

其中,  $d_p$  表示像素  $p$  处的视差,  $C(d_p, p, I)$  是 Birchfield-Tomasi 匹配代价<sup>[10,20]</sup>, 且该匹配代价对图像采样噪声不敏感。 $e, \sigma$  分别用来控制函数的形状。

$E_S(d_p, d_q)$  可利用下式来计算<sup>[10]</sup>:

$$E_S(d_p, d_q) = \begin{cases} 0 & \text{如果 } d_p = d_q \\ \rho_I(\Delta I) & \text{其他} \end{cases} \quad (3)$$

$\rho_I(\cdot)$  是以两个像素之间的灰度差  $\Delta I$  来定义的, 即

$$\rho_I(\Delta I) = \begin{cases} K \times h & \text{如果 } \Delta I < T \\ h & \text{其他} \end{cases} \quad (4)$$

$T$  是灰度差阈值,  $h$  是两个相邻像素灰度不平滑的补偿量,  $K$  是灰度差小于给定阈值的补偿系数。BP 算法是通过四连通图像所组成的图传递信息。每个信息是一个向量, 该向量的维数与可能出现的视差数量相同。令  $\mathbf{m}_{p,q}^{(t)}$  表示节点  $p$  在  $t$  时刻传递给节点  $q$  的信息。初始时刻,  $\mathbf{m}_{p,q}^0$  为均匀分布, 每次按照下面的式子对信息进行迭代:

$$\mathbf{m}_{p,q}^{(t)}(d_q) = \kappa \max_{d_p} (\exp(-E_S(d_p, d_q)) \cdot \exp(-E_D(d_p)) \cdot \prod_{r \in N(p) \setminus q} \mathbf{m}_{r,p}^{(t-1)}(d_p)) \quad (5)$$

其中,  $N(p) \setminus q$  表示节点  $p$  的邻域像素集 (不包括节点  $q$ )。经过  $n$  次迭代后, 每个节点都可得到一个置

信度向量, 即

$$\mathbf{b}_p(d_p) = \kappa \cdot \exp(-E_D(d_p)) \cdot \prod_{q \in N(p)} \mathbf{m}_{p,q}^{(n)}(d_q) \quad (6)$$

利用 HBP 算法<sup>[13]</sup>, 经过  $n$  次迭代, 即可得到使置信度达到最大的初始视差。

$$d_p^{\text{MAP}} = \arg \max_{d_p} \mathbf{b}_p(d_p) \quad (7)$$

### 3 像元集的平面拟合

由于基于像元集的方法<sup>[7,11,18]</sup>一般是假定在同一像元集内的视差满足连续性约束, 因此该类方法无法处理像元集内部存在视差跳跃的情况。文献[17]则假设同一像元集内只存在两种视差, 然后利用动态分割对存在不同视差的区域进行继续分割, 但却无法对表面是斜平面的物体进行建模。为了解决上述问题, 在平面拟合之前, 本文算法先根据初始视差确定是否对由颜色分割得到的区域进行二次分割, 然后再进行拟合。该阶段的具体实现包括控制点选择、区域分裂、平面拟合和区域合并 4 个步骤, 下面分别予以介绍。

#### 3.1 控制点选择

为了消除不可靠匹配的像素对平面拟合过程的影响, 可先从匹配结果中选择可靠匹配的像素, 即控制点 (ground control point, GCP); 然后再进行平面拟合过程。本文参考文献[17]的方法得到 GCP。令  $d_p^{\text{best}} = \arg \max_{d_p} \mathbf{b}_p(d_p)$ , 如果满足式(8)和式(9)的条件, 则节点  $p$  是候选的 GCP。

$$\mathbf{b}_p(d_p^{\text{best}}) \geq \lambda_{\text{amb}} \times \mathbf{b}_p(\hat{d}), \forall \hat{d} \neq d_p^{\text{best}} \quad (8)$$

$$\mathbf{b}_p(d_p^{\text{best}}) \geq \lambda_{\text{amb}} \times \mathbf{b}_{\hat{p}}(\hat{d}), \forall (\hat{p}, \hat{d}), \hat{p} + \hat{d} = p + d_p^{\text{best}} \quad (9)$$

如果  $\lambda_{\text{amb}} = 1$ , 上述的式(8)实质就是选出置信度最高的视差, 而式(9)实质是测试候选点是否满足左右图像的一致性检查, 即若该像素与左图上的其他像素对应于右图上的同一像素, 也就是, 如果该像素的置信度要比其他像素的置信度高, 则该像素才能成为候选 GCP。若候选 GCP 周围没有其他候选 GCP 的支持, 则该点一般是由噪声引起的误差点, 可将其舍弃, 即得到最终 GCP。

#### 3.2 区域分裂

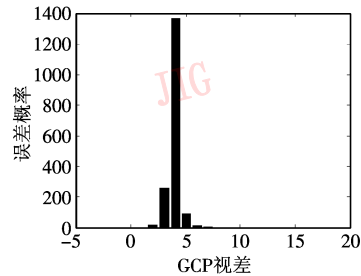
如果单纯根据颜色对图像进行分割, 那么分割一个区域内的像素就可能具有不同的视差, 这一般是由以下两个原因引起的; 一是分割将颜色相似、视差不

同的邻近区域分割到一个区域内;二是物体本身包含斜平面,且在该平面内视差有缓慢的变化,但不明显。因此要根据初始视差,按照不同情况对颜色分割形成的区域进行处理。若某个区域内视差初值变化超过一定的阈值,则将该区域根据视差初值进行二次分割;然后在颜色分割后的区域内统计每个视差出现的概率密度  $f(d = d_q) = \sum_{i=1}^N \delta(d_p = d_q) / N$ , 其中  $N$  为该区域内的 GCP 数量,当  $\delta(d_p = d_q)$  中的  $d_p = d_q$  时,  $\delta$  为 1, 否则为 0。若  $f$  是单峰分布, 则不再进行分割, 如图 2(b) 所示的单峰分布就是图 2(a) 中不规则区域

的 GCP 的视差概率密度, 该区域不必再进行分割; 若  $f$  的峰值大于 2 个, 且与峰值对应的最小视差和最大视差之差大于 3, 则根据视差将该区域进行二次分割, 图 2(d) 就是图 2(c) 中不规则区域的 GCP 视差概率密度分布, 图 2(e) 为根据视差对该区域重新分割后的结果, 新分割后的两个区域的边缘已分别在图中标出; 若  $f$  服从类均匀分布, 且  $a \geq 4$  ( $a$  表示区域内 GCP 所代表的主要视差数量), 则对该区域进行重新分割, 图 2(g) 就是图 2(f) 的视差概率密度, 图 2(h) 为对图 2(f) 进行重新分割后的结果, 新分割后的区域已分别在图中标出。



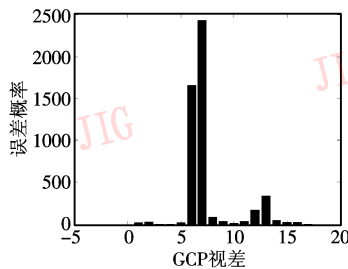
(a) 视差单峰分布的图像



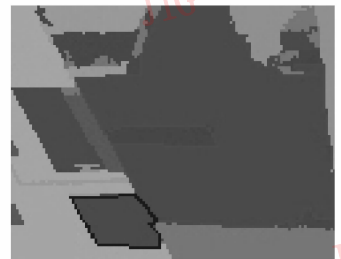
(b) 图 2(a) 中 GCP 的视差概率密度分布



(c) 视差双峰分布的图像



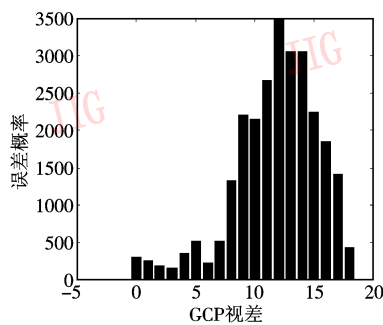
(d) 图 2(c) 中 GCP 的视差概率密度分布



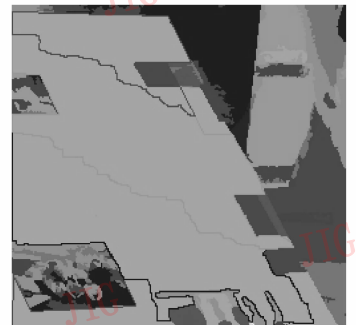
(e) 图 2(c) 重新分割后的结果



(f) 视差变化缓慢的区域



(g) 图 2(f) 中 GCP 的视差概率密度分布



(h) 图 2(f) 重新分割后的结果

图 2 不同图像的 GCP 概率密度分布和重新分割结果

Fig. 2 The probability density distribution of different images and the re-segmented results

### 3.3 平面拟合

Tao 介绍了如何根据初始视差值和分割得到的像元集来拟合一个平面<sup>[18]</sup>。一个平面可表示为  $d = c_0 + c_1x + c_2y$ , 其中,  $c_0, c_1, c_2$  是平面参数,  $d$  是像

素  $(x, y)$  处的视差。 $(c_0, c_1, c_2)$  是下面的线性系统的最小二乘解:

$$P \times [c_1, c_2, c_0]^T = B \quad (10)$$

其中, 矩阵  $P$  的第  $i$  个行向量  $p_i$  是  $[x_i \ y_i \ 1]$ , 列

矢量  $\mathbf{B}$  的第  $i$  个元素是  $d(x_i, y_i)$ 。第 1 次拟合后, 应再不断迭代更新平面参数。每次迭代时, 像素视差将根据上次迭代的结果在一定范围内变化, 然后根据修改后的视差更新平面参数, 直到平面参数趋向于一个稳定的值或超过一定的迭代次数为止。

### 3.4 像元集合并

虽然根据图像分割和初始视差处理后得到的图像像元集是零散和不规则的, 而且视差相同的像元集往往被分割成几个小块, 但可通过像元集的合并来解决这个问题。若两个相邻像元集的平面参数  $(c_0, c_1, c_2)$  相同, 则根据式 (1) 分别求利用初始视差得到的后验概率  $P^{\text{MAP1}}(\mathbf{X}|\mathbf{Y})$  和利用平面参数得到的后验概率  $P^{\text{MAP2}}(\mathbf{X}|\mathbf{Y})$  为

$$P^{\text{MAP1}}(\mathbf{X}|\mathbf{Y}) = \prod_p \Phi_p(d_p^{\text{MAP}}, y_p) \prod_p \prod_{q \in N(p)} \Psi_{p,q}(d_p^{\text{MAP}}, d_q^{\text{MAP}}) \quad (11)$$

$$P^{\text{MAP2}}(\mathbf{X}|\mathbf{Y}) = \prod_i \Phi_p(c_1 x_p + c_2 y_p + c_0, y_p) \times \prod_{p,q,p < q} \Psi_{p,q}(c_1 x_p + c_2 y_p + c_0, c_1 x_q + c_2 y_q + c_0) \quad (12)$$

若  $P^{\text{MAP2}}(\mathbf{X}|\mathbf{Y}) \geq P^{\text{MAP1}}(\mathbf{X}|\mathbf{Y})$ , 则将两个像元集合并为一个新的像元集, 否则保留两个像元集。重复上述过程, 直到所有的像元集都被访问过。由于经过像元集合并, 基元的数目就大大减少, 从而降低了以像元集为基元进行匹配的 TRBP 算法的计算量。表 1 列出了文献 [1] 在网站 <http://www.middlebury.edu/stereo> 提供的 4 幅标准图像在过分割、欠分割、欠分割经过空洞和碎片处理后以及结合初始视差进行像元集平面拟合合并后的像元集数目。

## 4 最终视差求解

这一节介绍如何利用 BP 算法在像元集域求解最终视差值的过程, 关于 BP 算法的具体实现已在 2.2 节做了详细描述, 因此这里只叙述如何在像元集域构造能量函数。

令  $\mathbf{R}$  为参考图像分割拟合后的像元集空间,  $\mathbf{D}$  为视差平面的估计值空间。匹配的目的就是找到标记  $\mathbf{L}$ , 使得每一个像元集  $\mathbf{g} \in \mathbf{R}$  都能得到相应的标记  $\mathbf{L}(\mathbf{g}) \in \mathbf{D}$ ,  $\mathbf{L}$  应该满足平滑性约束, 同时应与观察数据相吻合。这个问题可作为空间  $\mathbf{R}$  的一个能量最小化问题来求解:

$$E(\mathbf{L}) = E_D(\mathbf{L}) + E_S(\mathbf{L}) \quad (13)$$

$E_D$  是基元数据的能量项, 它包括将不同视差平面参数赋给像元集的代价

$$E_D(\mathbf{L}) = \sum_{\mathbf{g} \in \mathbf{R}} \text{CS}(\mathbf{g}, \mathbf{L}_{\mathbf{g}}) \quad (14)$$

其中,  $\text{CS}(\mathbf{g}, \mathbf{L}_{\mathbf{g}}) = \sum_{p \in \mathbf{g}-o} \mathbf{C}(d_p, p, l), \mathbf{C}(d_p, p, l)$  是 Birchfield-Tomasi 匹配代价,  $d_p = \mathbf{p}_i * \mathbf{L}_{\mathbf{g}}^T, \mathbf{p}_i = [x_i \ y_i \ 1]$  是由像素  $p$  的坐标组成的行向量,  $\mathbf{L}_{\mathbf{g}} = \mathbf{L}(\mathbf{g}) = [c_{1,g} \ c_{2,g} \ c_{0,g}]$  是视差平面的参数。为了排除遮挡像素对匹配代价的影响,  $p$  应该是平面  $\mathbf{g}$  内未被遮挡的像素。这里所谓的遮挡是检测像素  $p$  是否满足一致性约束, 也就是, 若像素  $p$  和  $q$  都对应于右图上的同一个像素, 则  $p$  和  $q$  都被判为遮挡, 不参与匹配代价的计算。

能量函数的另一项  $E_{\text{smooth}}$  通过惩罚视差平面间的不连续性来加入平滑性约束。

$$E_S(\mathbf{L}) = \sum_{(\mathbf{g}, \mathbf{g}')} \sum_{i \in E_{\mathbf{g}} \cap E_{\mathbf{g}'}} \delta(\mathbf{p}_i \times \mathbf{L}_{\mathbf{g}}^T \neq \mathbf{p}_i \times \mathbf{L}_{\mathbf{g}'}^T) \quad (15)$$

函数  $\delta(x \neq y)$  当  $x \neq y$  时为 1, 其他情况下为零。

## 5 实验及讨论

文献 [1] 总结了各种立体匹配方法, 同时给出了 4 幅测试图像, 并用这些图像对现有各种算法进行了实验比较, 关于它们的详细介绍见网址 <http://www.middlebury.edu/stereo>。笔者将本文方法应用于测试图像, 得到的视差图如图 3 第 3 列所示。从图 3 可以看出, 所得视差图与标准视差图除了细节方面有所不同外, 差别不是很大。尤其对于“Map”图像, 由于该图像中纹理灰度分布的对比度较低, 因此单纯依据颜色分割将会导致许多分割错误, 而本文方法在颜色分割后又根据视差进行了二次分割, 由于消除了颜色分割带来的错误, 因而效果较好。

本文又利用文献 [1] 中提供的定量参数对算法所得匹配结果进行了验证。在参数计算过程中, 一个“坏”匹配像素的定义为该像素的视差与标准视差的绝对差大于 1, 3 个定量参数  $b_0$ 、 $b_T$  和  $b_D$  分别统计了非遮挡区域的坏像素占总像素的百分比、非纹理区域的坏像素占总像素的百分比和视差不连续区域的坏像素占总像素的百分比。笔者将本文方法得到的结果上传到网站 <http://www.middlebury.edu/stereo>, 得到了本文算法的 3 个定量参数。表 2 列出了本文算法与排名在前 15 名的算法的参数比较。表 2 中的

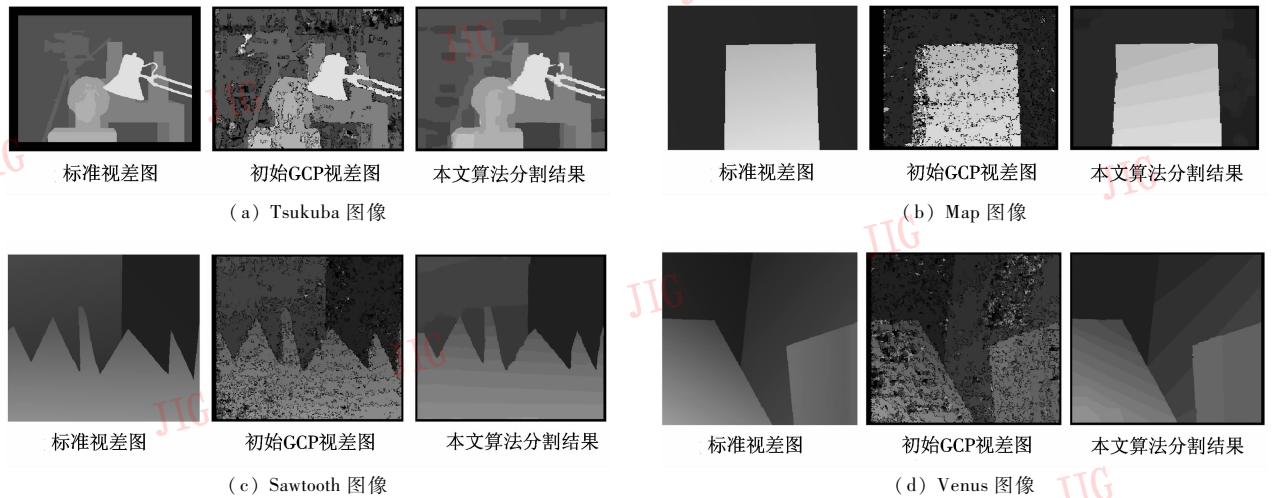


图 3 标准视差图和本文算法的分割结果  
Fig. 3 The ground truth map and our results

表 2 4 幅标准图像的 3 个定量参数  $b_o$ 、 $b_T$ 、 $b_D$  在不同立体匹配方法下的测试结果比较  
Tab. 2 The results of different stereo matching algorithms testing by four standard images

算法	4 幅标准图像的 $b_o$ 、 $b_T$ 、 $b_D$ 测试结果										
	Tsukuba			Sawtooth			Venus			Map	
	$b_o$	$b_T$	$b_D$	$b_o$	$b_T$	$b_D$	$b_o$	$b_T$	$b_D$	$b_o$	$b_D$
本文算法	1.13[3]	0.34[5]	6.49[5]	0.29[4]	0.00[1]	2.81[5]	0.33[5]	0.13[5]	4.86[6]	0.37[5]	3.24[2]
Sym. BP + occl.	0.97	0.28	5.45	0.19	0.00	2.09	0.16	0.02	2.77	0.16	2.20
Patch-based	0.88	0.19	4.95	0.29	0.00	3.23	0.09	0.02	1.50	0.30	4.08
Segm. -based GC	1.23	0.29	6.94	0.29	0.30	3.24	0.08	0.01	1.39	1.49	15.46
Graph + segm	1.39	0.28	7.17	0.25	0.00	2.56	0.11	0.02	2.04	2.35	20.87
GC + mean shift	1.13	0.48	6.38	1.14	0.06	3.34	0.77	0.70	3.61	0.95	12.83
Segm. + glob. vis.	1.30	0.48	7.50	0.20	0.00	2.30	0.79	0.81	6.37	1.63	16.07
Belief. prop.	1.15	0.42	6.31	0.98	0.30	4.83	1.00	0.76	9.13	0.84	5.27
Layered	1.58	1.06	8.82	0.34	0.00	3.35	1.52	2.96	2.62	0.37	5.24
2-pass DP	1.53	0.66	8.25	0.61	0.02	5.25	0.94	0.95	5.72	0.70	9.32
Region-Progress	1.44	0.55	8.13	0.24	0.00	2.64	0.99	1.37	6.40	1.49	17.11
GC + occl.	1.19	0.23	6.71	0.73	0.11	5.71	1.64	2.75	5.41	0.61	6.05
MultiCam GC	1.85	1.94	6.99	0.62	0.00	6.86	1.21	1.96	5.71	0.31	4.34
Improved Coop	1.67	0.77	9.67	1.21	0.17	6.90	1.04	1.67	13.08	0.29	3.65
Adapt. weights	1.51	0.65	7.24	1.14	0.27	5.48	1.14	0.61	4.49	1.47	13.58

本文算法一行在列出所得指标的同时,还给出了本文算法该项指标在所有算法中的排名(即[]中的数字)。由此可以看出,本文算法对图像具备普遍的适应性,不仅所得到的指标参数较好,且排名位置比较平均。这不仅证明了本文算法的有效性,同时说明本文算法在与同类方法的性能指标的比较中位于

前列。

## 6 结 论

本文提出了一种新的立体匹配算法,该算法首先以像素为基元求得初始视差值,然后结合图像颜

色分割结果对场景结构进行平面拟合,最后以拟合后的像素集为基元,利用 BP 优化算法得到最终解。实验结果表明,该算法可以克服单独以像素或以分割结果为基元进行匹配的缺点,并能正确处理颜色分割后的区域内部存在视差跳跃的情况,即可得到较为理想的结果。

### 参考文献 (References)

- 1 Scharstein D, Szeliski R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms [J]. *International Journal of Computer Vision*, 2002, **47**(1): 7 ~ 42.
- 2 Brown Z Myron, Burschka Darius, Hager G D. Advances in computational stereo [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, **25**(8): 993 ~ 1008.
- 3 Kanade T, Okutomi M. A stereo matching algorithm with an adaptive window: theory and experiment [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1994, **16**(9): 920 ~ 932.
- 4 Scharstein D, Szeliski R. Stereo matching with nonlinear diffusion [J]. *International Journal of Computer Vision*, 1998, **28**(2): 155 ~ 174.
- 5 Veksler O. Stereo matching by compact windows via minimum ratio cycle [A]. In: *Proceedings International Conference on Computer Vision [C]*, Vancouver, Canada, 2001: 540 ~ 547.
- 6 Veksler O. Fast variable window for stereo correspondence using integral images [A]. In: *Proceedings of the 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition [C]*, Madison, Wisconsin, USA, 2003: 556 ~ 561.
- 7 Deng Y, Yang Q, Lin X, *et al.* A symmetric patch-based correspondence model for occlusion handling [A]. In: *Proceedings International Conference on Computer Vision [C]*, Beijing, China, 2005, **2**: 1316 ~ 1322.
- 8 Kolmogorov V, Zabih R. Computing visual correspondence with occlusions using graph cuts [A]. In: *Proceedings International Conference on Computer Vision [C]*. Vancouver, Canada, 2001, **II**: 508 ~ 515.
- 9 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.
- 10 Tappen M F, Freeman W T. Comparison of graph cuts with belief propagation for stereo, using identical MRF parameters [A]. In: *Proceedings International Conference on Computer Vision [C]*, Beijing, China, 2003: 900 ~ 907.
- 11 Hong L, Chen G. Segment-based stereo matching using graph cuts [A]. In: *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition [C]*. Washington, DC, USA, 2004, **1**: 74 ~ 81.
- 12 Sun J, Zheng N N, Shum H Y. Stereo matching using belief propagation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, **25**(7): 787 ~ 800.
- 13 Felzenszwalb P F, Huttenlocher D P. Efficient belief propagation for early vision [A]. In: *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition [C]*, Washington, DC, USA, 2004, **1**: 261 ~ 268.
- 14 Sun J, Li Y, Kang S B, *et al.* Symmetric stereo matching for occlusion handling [A]. In: *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition [C]*, San Diego, CA, USA, 2005, **2**: 399 ~ 406.
- 15 Meltzer T, Yanover C, Weiss Y. Globally optimal solutions for energy minimization in stereo vision using reweighted belief propagation [A]. In: *Proceedings International Conference on Computer Vision [C]*, San Diego, CA, USA, 2005, **1**: 428 ~ 435.
- 16 Kolmogorov V, Wainwright M J. On the optimality of tree-reweighted max-product message-passing [A]. In: *21st Conference on Uncertainty in Artificial Intelligence [C]*, Edinburgh, Scotland, 2005: 316 ~ 323.
- 17 Wei Y, Quan L. Region-based progressive stereo matching [A]. In: *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition [C]*, Washington DC, USA, 2004: 106 ~ 113.
- 18 Tao H, Sawhney H S, Kumar R. A global matching framework for stereo computation [A]. In: *Proceedings International Conference on Computer Vision [C]*, Vancouver, Canada, 2001, **1**: 532 ~ 539.
- 19 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): 603 ~ 619.
- 20 Birchfield S, Tomasi C. A pixel dissimilarity measure that is insensitive to image sampling [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1998, **20**(4): 401 ~ 406.