

# 一种基于最小割的稠密视差图恢复算法\*

彭启民<sup>+</sup>, 贾云得

(北京理工大学 计算机科学与技术系,北京 100081)

## An Algorithm for Dense Disparity Mapping by Using Graph Cuts

PENG Qi-Min<sup>+</sup>, JIA Yun-De

(Department of Computer Science and Technology, Beijing Institute of Technology, Beijing 100081, China)

+ Corresponding author: Phn: +86-10-62614140, E-mail: qimin@ios.cn

Received 2003-11-21; Accepted 2004-02-03

Peng QM, Jia YD. An algorithm for dense disparity mapping by using graph cuts. *Journal of Software*, 2005,16(6):1090-1095. DOI: 10.1360/jos161090

**Abstract:** Vast computation is a great disadvantage of the existing graph cuts based vision algorithms. Lack of adaptability is another issue. An improved global optimal algorithm for dense disparity mapping using graph cuts is presented in this paper. First, adapted occlusion penalty and smoothness penalty are defined based on the intrinsic relation between the disparity changes and the discontinuities in an image. The graph cuts based algorithm is employed to get an optimal dense disparity mapping with occlusions. Secondly, according to the complexity analysis of graph cut algorithms, an operation named restricted  $\alpha$ -expansion operation is defined to control the vertexes generation during graph constructing based on the result of normalized correlation algorithm. It is a great help to reduce the vertexes and edges in the constructed graph, thus the computing is speeded up. The experimental results show performance of the proposed algorithm is improved and it will take a shorter time to compute an accuracy dense disparity mapping.

**Key words:** dense depth map; optimum; graph cut; connected region; restricted  $\alpha$ -expansion

**摘要:** 针对经典最小割算法计算量大和适应性不足的问题,提出一种改进的基于网络最小割计算稠密深度图的全局优化方法。首先,根据视差变化与不连续区域之间的关系,定义了具有一定适应性的平滑约束和遮挡约束,然后使用网络最小割算法,求解遮挡情况下的稠密视差。其次,在分析最小割算法复杂性的基础上,给出了一种受限 $\alpha$ -扩展( $\alpha$ -expansion)操作,该操作根据灰度连通性和特征点匹配的结果对每次网络构造的顶点进行控制,减少网络中顶点和边的数目,可有效提高计算效率。实验结果显示,该算法在保证视差恢复准确性的前提下,能以较快的速度计算出较理想的稠密视差图。

**关键词:** 稠密深度图;优化;最小割;连通区域;受限 $\alpha$ -扩展

中图法分类号: TP391 文献标识码: A

\* Supported by the National Natural Science Foundation of China under Grant No.60473049 (国家自然科学基金)

作者简介: 彭启民(1969—),男,山东滕州人,博士,主要研究领域为计算机视觉,图像处理,图像理解,小波分析;贾云得(1962—),男,博士,教授,博士生导师,主要研究领域为计算机视觉,媒体计算,智能系统。

立体视觉的基本问题是求解两幅或多幅图像之间的对应关系.传统的方法一般需要一致性、唯一性、连续性/平滑性等约束条件,以消除对应点的歧义性.但是这种约束条件在许多应用场合不成立.比如,物体棱角和边缘处的不连续性,随处可见的遮挡等.此外,传统的基于特征匹配的立体视觉算法也无法很好地求解缺乏纹理的图像对应性问题.近几年,人们将图论引入到计算机视觉领域,在求解立体匹配问题方面取得了良好的效果<sup>[1-4]</sup>.

一种直观的用于立体匹配的图构造方式如图 1 所示,这种立体网络的顶点表示可能的匹配像素对,网络的边表示顶点所代表像素间的邻接关系,以及保证最小割与一个匹配关系相对应的有关约束.这种网络假定了每个像素只能对应一个匹配像素,因而无法处理立体视觉中的遮挡问题.Boykov 等人引入外极线约束<sup>[2]</sup>,Ishikawa 引入了相邻外极线间的约束和顺序性约束来解决立体视觉的遮挡问题<sup>[3]</sup>.但是,这些图构造方法仍然具有局限性,比如,对平滑项的凸性要求会导致算法不鲁棒且可能导致 NP 问题.Kolmogorov 等人对上述方法进行了改进,但是该算法采用固定的遮挡约束和平滑约束,适应性差,计算效率也较低<sup>[4]</sup>.

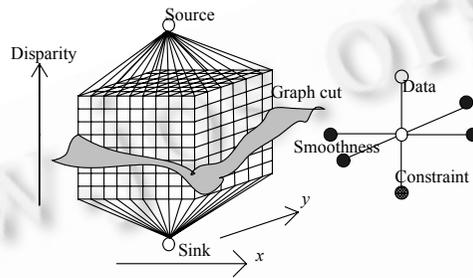


Fig.1 A network for stereo matching

图1 一种用于立体匹配的3D网络

本文在 Kolmogorov 等人工作的基础上提出了一种改进的稠密深度图恢复算法,根据视差的平滑性约束定义了一种受限 $\alpha$ -扩展( $\alpha$ -expansion)操作,对网络构造过程中生成的顶点数进行控制,降低了计算复杂度.同时,根据对数据项和惩罚项的分析,利用遮挡时视差变化的特征,给出了具有一定适应性的遮挡约束和平滑约束,对于真实场景的图像对,可以保证得到更精确的匹配.本文算法在解决立体匹配的遮挡以及纹理缺乏导致的歧义性问题的同时,显著改善了计算效率.

## 1 能量函数和最小割

对于一个有限的场景深度,所有可能的视差是一个有限集合,以所有可能匹配的像素对作为顶点集合,通过邻接及转移关系在适当的约束下建立一个立体网络,则最小割对应一个优化的匹配关系.为了求解该匹配关系 $f$ ,人们普遍使用如下的能量函数形式<sup>[5-7]</sup>:

$$E(f) = E_{data}(f) + E_{occ}(f) + E_{smooth}(f) \quad (1)$$

式中  $E_{data}$  为数据项,表示对应像素匹配一致性程度; $E_{occ}$  为惩罚项,用于保证每个像素尽可能地找到其对应关系; $E_{smooth}$  为平滑项,用于约束邻接像素间具有一致视差.

式(1)中各项的计算是建立在像素间对应关系集合基础上的,本质上属于组合优化问题,常规的求解方法如动态规划法、神经网络、模拟退火、遗传算法等要么难以用于高维数据,要么难以控制收敛,且效率普遍较低.

1956年,Ford和Fulkerson证明了著名的最大流最小割定理,即网络的最大流与最小割的等价性,并给出了求解的具体算法<sup>[8]</sup>.经过多年的发展,已经出现了可以在多项式时间内求解最大流/最小割的算法,如 Boykov 等人的改进算法在最坏的情况下其算法复杂度变为  $O(mn^2|C|)$ ,其中 $|C|$ 为所得到的割集的势, $m$ 和 $n$ 分别为边和顶点的数量,已经可以在一定程度上满足实时应用的需要<sup>[9]</sup>.近年来的工作证明,最小割算法非常适合于求解组合优化问题,本文将用最小割来求解立体视觉匹配全局优化问题.

## 2 最小割的近似算法

传统的最小割算法对于内存的需求随问题的规模呈快速增长的趋势,基于 $\alpha$ -扩展操作的近似算法用局部优化逐步逼近全局最优,通过缩短源到汇的路径长度,可以在有限的内存中实现网络最小割的快速计算<sup>[10]</sup>.

设两幅 $m \times n$ 的图像对 $I_1$ 和 $I_2$ ,允许的视差集合为 $D$ ,则一个可能的匹配构成图中的一个顶点集合:

$$O = \{(x, y, d) | 0 \leq x < m, 0 \leq y < n, d \in D\} \quad (2)$$

给定一个允许的视差值 $\alpha$ ,把 $O$ 中所有视差不为 $\alpha$ 的像素的视差改为 $\alpha$ ,这样的操作成为对 $O$ 的 $\alpha$ -扩展,记这些视差经过更改后得到的顶点集合为 $O^\alpha$ ,令 $\hat{O} = O + O^\alpha$ ,称为 $O$ 的 $\alpha$ -扩展子空间.在 $\hat{O}$ 上定义邻接关系:

$$N = \{(u, v) | u, v \in \hat{O}, v.x = u.x \pm 1, v.y = u.y \pm 1\} \quad (3)$$

按照如下方式构造有向图 $G=(V, E)$ :

$$\left. \begin{aligned} V &= \hat{O} \cup \{s, t\}, E = E_{s,u} \cup E_{v,t} \cup E_{u,v} \\ E_{s,u} &= \{(s, v) | v \in \hat{O}\}, E_{v,t} = \{(v, t) | v \in \hat{O}\} \\ E_{u,v} &= \{(u, v) | u, v \in \hat{O}, (u, v) \in N\} \end{aligned} \right\} \quad (4)$$

取4-邻域,每个点会影响到与其邻接的4个像素.按照上述的网络构造,一个像素的遮挡项应该介于其对邻接像素的最小平滑作用和最大平滑作用之间;取值过小时会出现虚假的遮挡像素,取值过大会使被遮挡像素误匹配的机会增加.此外,平滑项与数据项的取值应该保持适当的比例,若数据项较大,则减弱平滑项作用,导致全局优化性能变差.现定义各边上的权值如下:

从源到 $O$ 中顶点 $a$ 的边以及从 $O^\alpha$ 中顶点 $a$ 到汇的边,权重为 $D_{occ}(a)$ ;

$O$ 中的邻接顶点 $a$ 和 $b$ 之间的边上的权重为 $V_{a,b}$ ;

从源到 $O^\alpha$ 中顶点 $a$ 到边,权重为 $D_{data}(a)$ ;

从 $O$ 中顶点 $a$ 到汇的边,权重为 $D_{data}(a) + D_{smooth}(a)$ ;

从 $O$ 中顶点 $a$ 到 $O^\alpha$ 中顶点 $b$ 的边,权重为极大;

从 $O^\alpha$ 中顶点 $a$ 到 $O$ 中顶点 $b$ 的边,权重为 $V_{occ}$ .

设图中的一条割为 $C$ ,把图中顶点分割为包含源和汇的两个集合,分别记为 $V_s$ 和 $V_t$ ,则代表新的匹配关系的顶点集合为

$$\hat{V} = \{v | v \in O \wedge v \in V_s, v \in O^\alpha \wedge v \in V_t\} \quad (5)$$

若最小割表示的总能量降低,则以新匹配为初始匹配进行下一个计算,直到没有 $\alpha$ -扩展可以进一步降低能量为止.通过对所有允许范围的视差进行上述的计算,用 $\alpha$ -扩展子空间的局部逼近最终获得全局满意解.

## 3 适应性约束条件

实验表明,对于不含纹理或者具有规则纹理的图像,匹配的唯一性单靠灰度信息难以确定,必须引入上下文信息,根据视差的变化和区域的连续性决定平滑项和对遮挡的惩罚量,利用全局约束保证获得正确的匹配.为了降低噪声和离散化等因素的影响,数据项的计算要选取适当的特征.我们选择小波分解的第2层以上的系数组成特征向量<sup>[3]</sup>,得到如下的数据项:

$$D_{data}(a) = (1 + \beta) \|F_1 - F_2\| \quad (6)$$

其中 $-1 < \beta < 1$ 为数据项与平滑项的平衡系数.

在Kolmogorov等人的算法中,当邻域内首次引入一个新的视差时,则增加一次惩罚,当该邻域内再次用到该视差时则不再追加惩罚,显然,如果邻接像素具有相同的视差,则平滑项为0,当对邻域内所有像素赋予各不相同的视差时惩罚最大<sup>[3]</sup>.在比较平坦的区域,出现歧义性的机会将增加,此时应该减少数据项的惩罚量而更多地依靠平滑项,而在边缘处的情况刚好与之相反.因此,需要合理地平衡平滑项与数据项,保证在缺乏纹理的区域也能获得正确的匹配.这种适应性的平滑项表示为

$$D_{smooth}(a) = (1 - \beta) \sum_{\substack{(a,b) \in N \\ b \in \bar{O}, f(b) \neq f(a) \pm 1}} k_{a,b} V_{a,b} \quad (7)$$

其中  $V_{a,b} = \begin{cases} \mu & \text{像素间平滑} \\ 0 & \text{像素间非平滑} \end{cases}$ ,  $\mu > 0$ ,  $k_{a,b}$  为邻接顶点所代表的像素间的相似度.

统计结果显示,与图像中灰度连续区域相比,灰度不连续处附近像素被遮挡的几率较大,则遮挡项可以表示为

$$D_{occ}(p) = f\left(\frac{1}{k}\right) V_{occ} \quad (8)$$

其中  $k$  为图像的梯度,  $f$  为单调上升的函数,  $V_{occ}$  为预设的常数.

#### 4 受限 $\alpha$ -扩展和基于分割的初始化

Kolmogorov 等人的算法是从一个随机的允许匹配开始进行求解<sup>[4]</sup>,实验结果表明,经过对所有允许视差的一次迭代只能保证 80%的准确率,显示这种随机的初始化方法效率很低.该算法的另外一点不足就是,每一次 $\alpha$ -扩展都对所有视差不为 $\alpha$ 的顶点进行扩展.由图的构造方法和最小割的定义可知,对于远远偏离实际视差的对应,无助于降低优化函数的能量,只会导致生成的网络规模扩大,降低计算速度.根据对最小割算法的复杂性分析,提高效率的关键在于减少顶点和边的数目,因此对 Kolmogorov 算法的改进,显然应该集中在初始化和控制网络的规模上.

由于场景到图像的映射过程导致信息的缺失,传统的相关匹配法虽然简单,但其匹配结果却无法保证其可靠性,特别是在平坦区域或重复纹理区域,由于峰值不唯一,无法确定其对应点,误差显著增加.但边缘和角点等显著特征点,其相关值存在唯一的峰值,因而可以实现准确匹配.只要选择适当的特征,基于特征匹配的算法计算的稀疏深度可以快速计算,辅以连续性和顺序性约束,可以获得比较高的准确度,作为近似计算的初始值无疑是有益的,其中的误匹配将通过基于最小割的全局优化算法逐步得到修正.

Boykov 等人指出,视差与灰度的变化具有内在的关联性<sup>[2]</sup>.通过观察可以发现,灰度的连续性总是伴随着视差的连续性,视差不一致主要出现在边缘或灰度不连续处.因此,在一个基于灰度的连通区域内,所有像素的实际视差总是落在一个有限的连续范围内,这个范围可以由该连通区域内的特征点和区域边缘点的视差共同决定.从视差和遮挡的关系分析,在视差变化区域里,大视差(近距离)对应的区域肯定遮挡小视差(远距离)对应的区域.

只要确定了角点和边缘上的特征点的深度,就可以把其所属连通区域  $i$  内各像素的视差取值限定在一个有限的范围( $d_{min}^i, d_{max}^i$ )内.要实现这个目标,首先要把原图像根据灰度划分为不同的连通区域,我们选择了形态学方法中的分水岭(watershed transform)算法以完成该操作.每次迭代过程中进行 $\alpha$ -扩展操作时,先根据像素的可能视差范围进行判定,若 $\alpha$ 在其允许视差范围内则添加新的顶点,否则不作处理.此外,对于满足唯一性的特征点也不进行扩展处理.我们定义该过程为受限 $\alpha$ -扩展操作.受限 $\alpha$ -扩展操作可以显著减少所构造网络中的顶点数,有效地降低计算强度.本文算法的具体步骤如下:

- (1) 计算图像的梯度图像,用分水岭算法实现对图像灰度连通区域的分割;
- (2) 根据相关匹配算法的结果构造初始匹配,并用特征点视差设置各区域内像素的可能视差范围,初始化最小割代价为极大值;
- (3) 对视差允许范围内的所有取值作未处理标记;
- (4) 选择一个未处理视差,作已处理标记,对当前匹配作受限 $\alpha$ -扩展,完成网络构造,并对邻接边赋权;
- (5) 计算最小割,若代价减小,则构造新的匹配关系;
- (6) 若存在未作已处理标记的视差允许值,转(4);
- (7) 达到预定的循环次数结束,否则转(3).

### 5 实验结果

所有算法用 VC++6.0 在 Windows2000 下实现,计算机 CPU 主频为 P4 1.6G,128M 内存.我们分别用 Tsukuba 测试图像和我们自己采集的真实图像进行算法验证.图 2(a)是 Tsukuba 测试图像对,大小为 384×288 像素,最大视差为 16 个像素.实验中对图像进行了 3 层的小波分解,采用 3×3 邻域进行相关匹配.表 1 为有关算法的计算结果与真实数据误差大于 1 的有关统计数据,只有未被遮挡的像素参与统计.从表中数据可以看出,本文算法由于选用了小波特征,较 Kolmogorov 单纯基于灰度的匹配总误差率有所下降,适应性的约束条件也改善了深度不连续处的正确率,而在平坦区域的指标则相同.

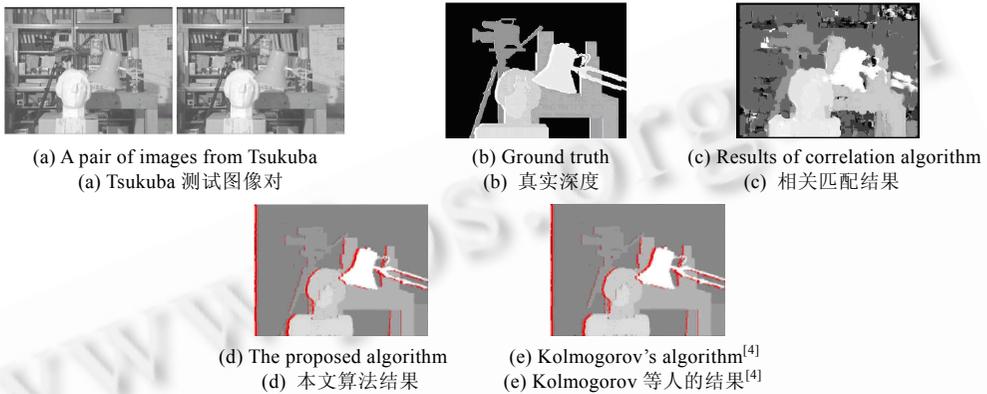


Fig.2

图 2

Table 1 Error comparison of disparity mapping results

表 1 视差计算结果误差对比

Algorithm	Total (%)	Flat region (%)	Discontinuous region (%)
Correlation	5.23	3.80	24.66
Boykov <sup>[2]</sup>	1.86	1.00	9.35
Kolmogorov <sup>[4]</sup>	1.27	0.43	6.90
Proposed	1.24	0.43	6.71

图 3(a)为我们自己采集的测试图像对,大小为 384×288,水平方向最大视差为 47 个像素,图 3(b)是经过均衡化处理及校正后进行了 5×5 LoG 滤波的结果,目的是为了消除噪声和光线变化的影响.图 4(a)是用分水岭算法获得的连通区域,图 4(b)是相关匹配结果,其中黑色区域代表不确定深度部分,图 4(c)是本文算法得到的结果,白色代表被遮挡区域.具体的运行时间见表 2.可以看出,在没有降低匹配准确度的前提下,由于减少了每次迭代处理的顶点数和边数,基于受限 $\alpha$ -扩展的优化算法运算效率有很大提高,当视差搜索范围较大时效果更明显.由于每个连通区域的视差取值范围是由多个特征点共同决定的,从统计的角度分析,该范围以较大概率覆盖了区域内各像素真实的视差范围,即各像素可以获得对应真实视差的扩展节点.即使选取的个别特征点获得的初始视差存在误差,也是以较大概率导致真实的视差范围被扩大,因此只会影响到受限 $\alpha$ -扩展操作的效果而不会对最终结果产生根本性的影响.

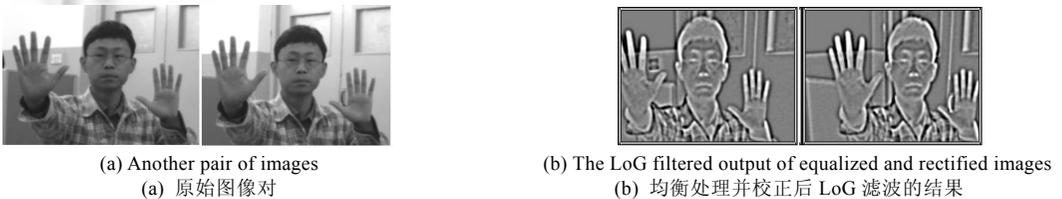


Fig.3

图 3

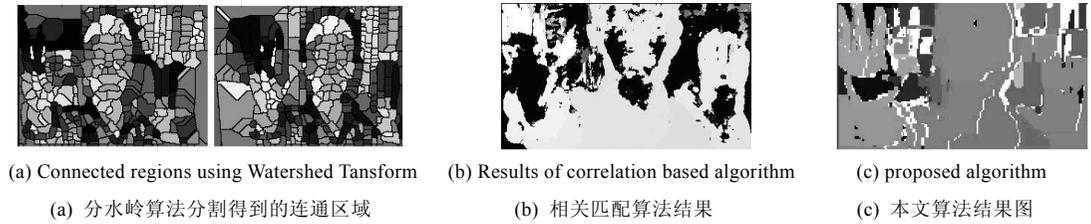


Fig.4  
图 4

Table 2 Time consumption comparison (s)

表 2 运行时间的对比(s)

	$\alpha$ -expansion	Restricted $\alpha$ -expansion
Tsukuba image	47	29
Real image	152	61

## 6 结束语

本文在分析立体匹配中有关约束以及遮挡的基础上,定义了具有一定适应性的约束条件,根据深度/视差与灰度变化内在的关联性,在把图像进行基于灰度连通性分割的基础上,提出了一种受限 $\alpha$ -expansion的操作,可有效地对每次网络构造的顶点进行控制,提高了计算效率.通过对网络中各类边权值关系的分析,合理设定平滑项与数据项的平衡,保证了在缺乏纹理的区域也能获得正确的匹配,直接得到稠密的深度信息.连通区域的合理设置对本算法的最终结果有一定影响.

## References:

- [1] Szeliski R, Zabih R. An experimental comparison of stereo algorithms. In: Proc. of the Int'l Workshop on Vision Algorithms: Theory and Practice table of contents. LNCS 1883, London: Springer-Verlag, 1999. 1-19.
- [2] Boykov Y, Veksler O, Zabih R. Markov random fields with efficient approximations. In: Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition. Washington, DC: IEEE Computer Society, 1998. 648-655.
- [3] Ishikawa H. Global optimization using embedded graphs [Ph.D. Thesis]. New York: University of New York, 2000.
- [4] Kolmogorov V, Zabih R. Computing visual correspondence with occlusions using graph cuts. In: Proc. of the Int'l Conf. on Computer Vision (ICCV 2001), Vol. II, 2001. 508-515.
- [5] Geiger D, Ladendorf B, Yuille A. Occlusions and binocular stereo. Int'l Journal of Computer Vision, 1995,14(3):211-226.
- [6] Bobick AF, Sintille S. Large occlusion stereo. Int'l Journal of Computer Vision, 1999,33(3):1-20.
- [7] Veksler O. Efficient graph-based energy minimization methods in computer vision [Ph.D. Thesis]. Cornell University, 1999.
- [8] Cook WJ, Cunningham WH, Pulleyblank WR. Combinatorial Optimization. John Wiley & Sons, 1998.
- [9] Boykov Y, Jolly MP. Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images. In: Proc. of the Int'l Conf. on Computer Vision (ICCV 2001). 2001. 105-112.
- [10] Boykov Y, Veksler O, Zabih R. Fast approximate energy minimization via graph cuts. IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI), 2001,23(11):1222-1239.