

基于能量最小化扩展景深的实现方法

姜晓红 戴芬 姜翰青

(浙江大学CAD&CG国家重点实验室, 杭州 310027)

摘要 提出并实现了一种基于能量最小化的扩展景深方法,该方法使用对比度作为评价标准来拼接源图像,它的创新之处是利用了基于图分割的能量最小化方法提供的空间一致性来解决拼接图像问题,并用最大流算法来解决图分割问题。最后给出了实验结果,将实验结果与缺乏空间规范化的结果进行了比较。实验结果证明,本文方法能够在用户可以接受的时间范围内得到较为优越的扩展景深效果。

关键词 扩展景深 能量最小化 图分割 最大流/最小割

中图法分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2006)12-1854-05

An Energy Minimization Based Extended Depth-of-field Method

JIANG Xiao-hong, DAI Fen, JIANG Han-qing

(State Key Laboratory of CAD&CG, Zhejiang University, Hangzhou 310027)

Abstract This paper proposes a kind of energy minimization based extended depth-of-field method, which uses a measure of contrast to compose source images. The innovation point is that it makes use of the spatial consistency afforded by the energy minimization via graph-cut to solve the problem of composite artifacts, and uses the max-flow algorithm to solve the graph-cut problem. In the end of the paper, we present the result, and then compare it with the result of the algorithm without spatial consistency. The result shows that the method can produce more perfect effect within an acceptable execution time.

Keywords extended depth-of-field, energy minimization, graph cut, max-flow/min-cut

1 引言

使用显微镜看物体标本时,只能看到较小范围的景深,不能看清物体的全貌。同样,人们想要以美丽的风景作为背景为朋友拍摄写真照片时,都感觉到成像中的“景深”问题。可以把人物面部放置在焦平面上,也可以把风景放在焦平面上,但却不能同时对两者进行聚焦。以往的扩展景深方法多数是通过硬件方法解决光学成像的局限,即通过添加立体相位器件来人为地扭曲透镜的二次相位。这样立体扭曲的结果是整个相机深度上的散焦,也就是在焦平面和不足焦平面的物体都模糊了。然后通过数字图像处理来消除模糊。这种方法不仅需要特殊的硬件

结构,而且经相机得到的也只是深度散焦的模糊图像,仍需要经过图像处理方法消除模糊。这对于普通用户来说过于复杂。

Laplacian 金字塔法也曾被用来处理扩展景深问题^[1,2]。1994年,Haerberli利用对比度选择源图像区域的逐像素启发式方法对Laplacian金字塔法进行了简化^[3]。然而,这两种技术由于缺乏空间规范化而产生了噪声。

这里提出一种既不需要硬件支持,又考虑到图像空间一致性,基于图分割最优化和能量最小化的扩展景深方法。用户只需要利用普通的照相机或数码相机拍摄几幅聚焦不同的照片,经此方法处理后能够生成一幅各个层面完全清晰显示的图像。该算法能够得出具有更少接缝图像和更多聚焦区域的结果。

基金项目: 国家自然科学基金重点项目(60533080)

收稿日期: 2006-05-30 改回日期: 2006-06-19

第一作者简介: 姜晓红(1966~),女,副教授。2003年于浙江大学获计算机应用博士学位。主要研究方向为虚拟现实和计算机图形图像处理技术。E-mail: jiangxh@cad.zju.edu.cn

2 系统框架和基本原理

与 Haeberli 方法^[3]类似,使用对比度作为评价标准来拼接源图像。与此同时,受数字蒙太奇方法^[4]的启发,利用了图分割最优化提供的空间一致性来优化拼合结果,减少剩余的噪声,即通过基于图分割的能量最小化方法来减少接缝图像。在求取最小能量的过程中,利用了最大流/最小割的算法。

2.1 算法总体思想和框架

扩展景深系统的算法框架如图 1 所示。

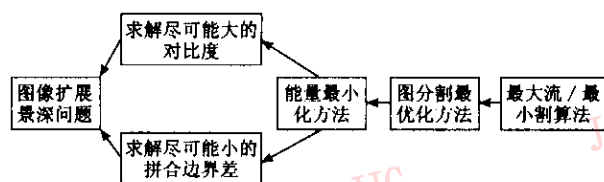


图 1 算法框架图

Fig. 1 Algorithm chart

在扩展景深系统中,需要解决的是两个问题:一方面,要选择源图像序列中和评价标准最一致的像素点,在系统中就是选择对比度尽可能大的像素点;另一方面,要保证图像的拼接质量最少,也就是拼合边界两边像素点之间的差异尽可能地小。为了同时解决这两个问题,可以根据两方面的要求列一个能量方程:

$$\alpha(L) = \sum_p C_d(p, I(p)) + \sum_{p,q} C_i(p, q, I(p), I(q)) \quad (1)$$

其中 C_d 表示像素 p 和评价标准之间的不一致性, C_i 表示的是相邻像素之间的差异程度。因此只需使得这个方程达到最小解,便可以解决上述两个问题。

此方法就是早期和现今计算机视觉问题中常用的能量最小化方法。

在计算最小能量值时,使用了 Boykov 提出的图分割方法^[5,6],并利用最大流/最小割算法^[7]来解决图分割最优化问题。

2.2 利用图分割方法的快速近似能量最小化

利用图分割方法来快速近似地计算方程式(1)的最小能量^[5]。

在扩展景深系统中,有 n 幅预先载入的源景深图像 $S_1, \dots, S_n$ 。为了形成一幅合成图像,必须为每

个像素 p 选择一幅源图像映射 S_i ,代表此像素点的颜色取自源图像 S_i 对应位置的像素点。把像素点和源图像之间的映射叫做“标记法”,并且将每个像素 p 的标记指示为 $I(p)$ 。约定:如果 $I(p) \neq I(q)$,则在合成图像两相邻像素 p, q 之间存在一条接缝。

Boykov 提出了两种计算局部最小量的图分割算法。一种是 α 扩张:对于一种标记 α ,此变动将任意一部分像素集的标记赋为 α ;另一种是 α - β 交换,对于一对标记 $\{\alpha, \beta\}$,此变动将任意一部分标记为 α 的像素集与任意一部分标记为 β 的像素集之间相互交换标记。这两种算法都将最终生成能量最小的标记法。

系统利用了 α - β 交换的方法。算法伪码如下:

- 1 Start with an arbitrary labeling f
- 2 Set $success := 0$
- 3 For each pair of labels $\{\alpha, \beta\} \subset L$
 - 3.1 Find $\hat{f} = \arg \min E(f')$ among f' within one α - β swap of f
 - 3.2 If $E(\hat{f}) < E(f)$ set $f := \hat{f}$ and $success := 1$
- 4 If $success = 1$ goto 2
- 5 Return f

算法第 t 次内部循环迭代将特定的一对标记 $\{\alpha, \beta\}$ 和当前标记 L_t 作为输入,然后计算最佳标记 L_{t+1} 使得 $L_{t+1}(p) = L_t(p)$ 或者 $L_{t+1}(p) = \{\alpha, \beta\}$ 。外部循环在可能的标记对上反复迭代。当对于所有标记的一次循环不再减小惩罚函数值的时候,算法结束。

正如式(1)所定义的,将一个像素标记 L 的惩罚函数 C 定义为两项之和:一项是对于所有像素 p 的数据惩罚因子 C_d ,另一项是对于所有拼合边界上相邻像素对 p, q 的交互惩罚因子 C_i 。

对于扩展景深的应用来说,数据惩罚因子是由像素点的未知颜色和评价标准(最大对比度)的 RGB 空间距离定义的,而交互惩罚因子由距接缝目标的 RGB 空间距离定义。

算法的关键是内循环,即寻找最小惩罚函数值。Greig^[8]等人提出了用两端点图的图分割方法来解决这个问题,即通过寻找图中惩罚值最小的分割就能得到每个像素的最佳标记法。但由于二进制图像计算的限制,Greig 的方法在提出后的 10 年内并没有得到广泛的利用。直到 20 世纪 90 年代计算机视觉界中出现了基于图论的最大流/最小割算法,二进制图分割方法才得以推广。

图分割的基本原理:可以将扩展景深问题抽象为拥有两个特殊顶点 α, β (称为端点) 的带权图 $g = \langle V, \varepsilon \rangle$ 。 V 为所有节点的集合, ε 为连接相邻节点的所有边线集, 边的权重值等于与 α, β 端点以及像素节点之间相应的惩罚因子值。节点为景深图像中所有像素点。边代表每个像素点与其周边像素点之间的关系。两端点分别代表图像标记 α 和 β 。一条分割 $C \subset \varepsilon$ 指的是使得 α 和 β 在剩余图 $g(C) = \langle V, \varepsilon - C \rangle$ 中被分离的一个边的集合。另外, 没有任何 C 的真子集将 $g(C)$ 的两端点分离。分割 C 的惩罚值(用 $|C|$ 表示)等于它边的权重值之和。如图 2 所示, 其中源图像序列的像素集为 $P_{\alpha\beta} = P_\alpha \cup P_\beta$, P_α, P_β 为像素集的一个分割, $P_\alpha = \{p, r, s\}$, $P_\beta = \{q, \dots, w\}$ 。

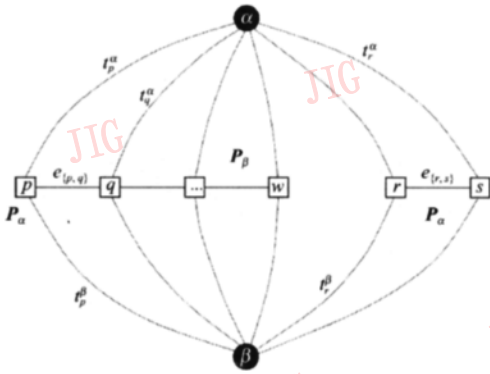


图 2 一个图实例 $g_{\alpha\beta}$ 对于源图像序列的抽象

Fig. 2 An example of the graph $g_{\alpha\beta}$ for a 1D image

我们所要做的就是找到一条最优化的分割 C , 它的总惩罚值最小。它等价于通过最小化吉布斯能量 $E(x)$ 来得到最好的标记法解:

$$E(X) = \sum_{i \in V} E_1(x_i) + \lambda \sum_{(i,j) \in \varepsilon} E_2(x_i, x_j) \quad (2)$$

称 $E_1(x_i)$ 为数据能量, 它解析当节点 i 的标记是 x_i (α 或 β) 时候的惩罚值。在系统中, 它代表节点在 α/β 图像中的对比度的相反值, 将 E_1 定义为关于节点 i 对比度的单调递减函数, 也就是说节点的对比度越大, E_1 就越小, 由此节点所在的图像像素就越可能成为拼接图像中的像素。称 $E_2(x_i, x_j)$ 为交互能量, 它表明当相邻节点 i 和 j 分别是 x_i (α 或 β) 和 x_j (α 或 β) 时的惩罚值。它代表沿接合边缘的梯度而产生的能量值。我们仅仅可以在图像区域拼接边界上获取梯度信息。换句话说, 只有当相邻节点赋予不同标记时 E_2 才能成为一个惩罚因子。将 E_2 定义为关于相邻节点 i 和 j 之间梯度的单调递增

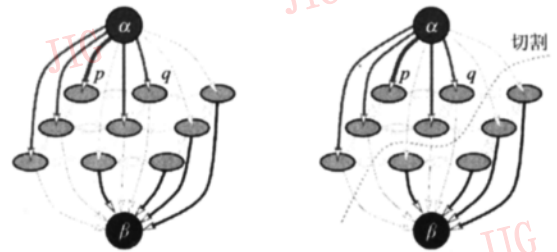
函数, 也就是说两节点之间的梯度越大, E_2 就越大, 由此节点所在的图像位置就越不可能成为 α, β 两图像的拼接区域边缘。

从理论上说, 由此计算出的最小化能量会使得拼接区域像素点的对比度尽可能地大, 同时具有尽可能不明显的拼接边缘, 这样的能量函数设计便达到了用户理想的扩展景深需求。我们设计了合适的数据能量与交互能量函数, 使得扩展景深图像达到最佳的效果。

2.3 用最大流方法解决图分割问题

在寻求最小值切割问题的过程中, 为了避免冗余且缓慢的组合计算方法, 快速地计算最小能量, 我们使用了最大流/最小割的算法^[7]。它是一种基于经典的极大流图论算法的, 在计算机视觉领域有着广泛应用的算法。简单地说, 可以通过寻求源端点 s 至目标端点 t 的最大流的方法来解决最小 α/β 切割问题。

可以将图的边视为最大流量等于边权重值的“有向管道”, 最大流就是指自源端点 s 至目标端点 t 传输的“最大水流量”。如图 3(a) 所示, 其边的粗细代表权重值的大小, 箭头方向指示水流管道方向。图 3(b) 是对此图的一个最小值切割。Ford 和 Fulkerson^[9] 的理论阐明: s 至 t 的最大流使得图中的一些边达到饱和, 将图结点分为对应于一条最小切割的两个不相交的部分 $\{S, T\}$ 。因此, 最大流和最小割问题是等价的。事实上, 最大流值与最小割的总惩罚值是相等的。可以通过求最大流来解决最小割问题。



(a) 最大流

(b) 最小值切割

图 3 用最大流算法解最小分割问题(3×3 图)

Fig. 3 Using max-flow algorithm to solve the graph-cut problem(3×3 graph)

最大流/最小割算法

在扩展景深系统中, 利用 Boykov 提出的基于增广路径的新最大流/最小割算法。和 DINIC 算法^[7]相似, 它通过建立查找树来寻找增广路径, 不同之处

在于它将重用而不是重新构造这棵树。Boykov 的实验结果表明在典型视觉问题中 ,此算法很大程度上优于标准算法。

算法的实现过程如下 :

保存一棵以 α 作为根结点的搜索树 S ,其中自父结点到子结点的边都是非饱和的。不在 S 中的点称为“自由”结点。所有自由结点的集合用 T 表示。搜索树 S 中的结点可分为“主动”和“被动”两类。主动结点可以增长 ,也就是说它们可以从自由结点集中获得新的子结点。被动结点不具有通过非饱和边相连的自由邻结点。因此 ,被动结点不能增长。

算法迭代地重复以下 3 个步骤 :

- (1)生长阶段 搜索树不断增长 ,直到找到 β 。
- (2)增广阶段 已寻找到的路径被增广 ,搜索树被分割为森林。
- (3)回收阶段 森林被重新转换为一棵树。

在生长阶段 ,搜索树在增广 ,主动结点从自由结点集中获得新的子结点。新获得的结点也成为搜索树 S 的主动成员。只要指定主动结点的所有邻结点被遍历过 ,此主动结点就变为被动。当遇到 β 时生长阶段就结束了 ,这样就找到了从 α 到 β 的一条路径。

增广阶段对生长阶段寻找到的路径进行增广。因为我们顺着可能的最大流量推动 ,因而路径中的一些边就饱和了。因此树中的一些结点变成了“孤儿” ,也就是说 ,连接这些结点和它们父结点的边不再有效 (它们饱和了)。事实上 ,增广阶段把搜索树

S 拆分成一个森林。原点仍然是森林中某棵树的根结点 ,孤结点成为其他树的根结点。

回收阶段的目标是恢复为一个以 α 为根结点的搜索树结构。在这个阶段试图为每一个孤结点寻找一个新的有效的父结点。如果没有这样的父结点 ,就把这个孤结点从 S 中移除 ,使它成为自由结点。亦将它以前的子结点声明为孤结点。当没有孤结点留下时 ,这个阶段就结束了 ,于是搜索树 S 就被修复了。因为 S 中的一些孤结点变为自由结点 ,所以回收阶段导致集合 S 的收缩。

回收阶段完成以后 ,算法返回生长阶段。当搜索树不能再生长 ,且 β 没有被找到 (所有的主动结点检查他们的邻结点 ,并且变为被动)时 ,算法结束。

3 实验结果及比较

图版 I 图 1 是采用本文方法和使用对比度选择源图像的逐像素启发式方法的效果比较图 ,缺乏空间规范化的结果具有较多的拼接痕迹 ,并且会有一些噪音。而本文方法的结果具有较好的图像连续性和清晰度 ,拼接效果更佳。

图版 I 图 2 为一组静物的扩展景深效果图。

表 1 对 α - β 交换算法运算效率的两组测试结果进行了比较。其中一次循环的平均时间是指对于整幅源图像的所有像素点进行一次 α - β 交换需要的时间。

表 1 两组图像 α - β 交换算法效率测试比较
Tab.1 Comparisons of the efficiency of α - β swap algorithm of two groups of images

源图像像素	源图片数目(N)	算法的总运行时间(s)	循环次数(C_N^2)	一次循环的平均执行时间(s)
1 344 × 1 021	13	140	78	1.79
1 000 × 667	5	10	10	1.00

从表 1 中可以看出 :无论是对于大面积的还是数目较多源图像 , α - β 交换算法的执行效率相对许多传统计算机视觉问题而言都是较为理想的 ,可接受的。源图片数目影响最终执行的循环次数 ,从而间接影响总的运行时间。因此 ,一次循环的平均执行时间更能准确地反映算法的效率高低。然而 ,可以从中看到 ,源图像像素点个数的多少直接影响着一次循环的运行时间 ,从而影响着总体运行速度的快慢。

表 2 是逐像素启发式方法和 α - β 交换算法的效

表 2 α - β 交换算法和逐像素启发式法的效率测试比较
Tab.2 Comparison of the efficiency of α - β swap algorithm and per-pixel heuristics of salience

	所用的技术方法	
	逐像素启发式法	能量最小化方法 (α - β 交换)
源图像像素	1 344 × 1 021	1 344 × 1 021
源图像数目	13	13
扩展景深算法的总运行时间(s)	14	140

率测试比较结果。逐像素启发式方法虽然在运行速度上具有优势,但这是以产生相比较而言不理想的视觉效果为代价的,而本算法则能够在可以接受的时间范围内得到较为优越的景深效果(如图版 I 图 1(b))。

4 结 论

本文提出了一种有效的基于能量最小化的扩展景深方法。把基于图分割的快速能量最小化方法和最大流/最小割算法很好地应用于扩展景深系统,并得出了较好的结果。本文算法既不需要硬件支持,又考虑到图像的空间一致性。用户只需在系统中输入几幅聚焦不同的照片,经本实验系统处理后能够生成一幅各个层面完全清晰显示的图像。通过试验结果比较可以看出本文方法的结果具有较少的拼接痕迹。

今后,将进一步完善扩展景深系统,在系统中加入图像平滑功能对现有的生成效果进行二次处理,可以利用 Poisson 方程的平滑梯度场引导功能对图像拼接边缘进行融合,除去图像中的噪声点,使得效果图中残余的拼接痕迹得以更大程度地减少,得到更为优化的景深效果。

参考文献(References)

1 Ogden J M , Adelson E H , Bergen J. *et al.* Pyramid-based computer

graphics[J]. RCA Engineer , 1985 , 30(5) : 4 ~ 15.

2 Burt P , Kolcztnski R. Enhanced image capture through fusion[A]. In : International Conference on Computer Vision(ICCV 93) [C] , Germany , Berlin , 1993 : 173 ~ 182.

3 Haeberli Paul. A Multifocus Method for Controlling Depth of Field on Grafica Obscura Web Site[EB/OL]. <http://www.graficaobscura.com/depth/index.html> , 1994-08.

4 Agarwala A , Dontcheva M , Agrawala M , *et al.* Interactive digital photomontage[A]. In : Proceedings of ACM SIGGRAPH 2004[C] , USA , California , Los Angeles , 2004 : 294 ~ 302.

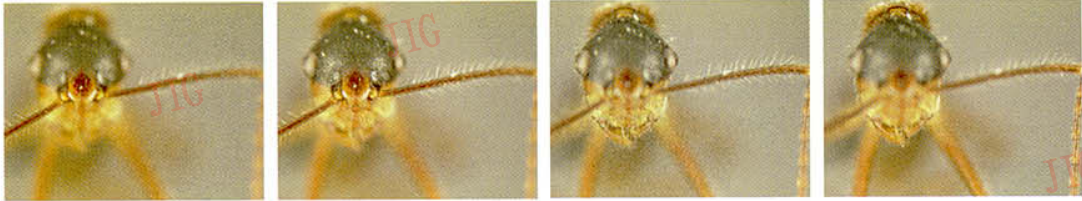
5 Boykov Y , Veksler O , Zabih R. Fast approximate energy minimization via graph cuts[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence , 2001 , 23(11) : 1222 ~ 1239.

6 Kwatra V , Schödl A , Essa I , *et al.* Graphcut-textures : image and video synthesis using graph cuts[A]. In : Proceedings of ACM SIGGRAPH 2003[C] , San Diego , California , USA , 2003 : 277 ~ 286.

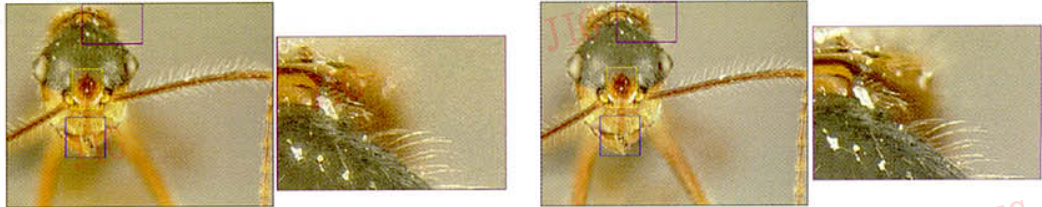
7 Boykov Y , Kolmogorov V. An experimental comparison of mincut/ max-flow algorithms for energy minimization in vision[J]. IEEE Transactions on Pattern Analysis and Machine Interlligence , 2004 , 26(9) : 1124 ~ 1137.

8 Greig D , Porteous B , Seheult A. Exact maximum a posteriori estimation for binary images[J]. Journal of the Royal Statistical Society , Series B , 1989 , 51(2) : 271 ~ 279.

9 Ford L , Fulkerson D. Flows in Networks[M]. New Jersey : Princeton University Press , 1962.



(a) 一组以不同景深拍摄的昆虫显微照片源图像



(b) 对比度选择源图像的逐像素启发式方法得到的效果图

(c) 本文方法的效果图



(d) 显示细节的放大图

图1 本文方法和缺乏空间规范化算法结果比较

Fig.1 A comparison of the result of our algorithm with the result without spatial consistency



(a) 红色相机放置在焦平面上



(b) 聚焦于前方相机上



(c) 扩展景深的结果

图2 静物的扩展景深效果图

Fig.2 Object's extended depth of field result