

针对图像区域连续化问题的立体匹配算法

徐姗姗,刘应安,徐 昇

(南京林业大学 信息科学与技术学院,江苏 南京 210037)

[摘要] 传统图割算法解决双目立体匹配问题,在高精度的同时需要消耗大量时间.提出一种新的算法,将最小割求取问题转化为贪心问题,从而降低算法复杂度.由于转化后的图割在处理图像区域连续化问题时效率低下,给出了图割与区域匹配相结合的 GR(Graphic Cut in Region)算法,算法不仅将图割理论运用到立体匹配问题中,且在求取初始视差时提出了用窗口单位化匹配代价算法来提高初始视差的精度.实验证明,该算法在图像区域连续化时有较好的效果,明显提高了匹配的精度,且复杂度也大大降低.

[关键词] 图割,图像区域连续化,能量函数最小化,区域匹配,窗口单位化

[中图分类号] TP301.6 **[文献标志码]** A **[文章编号]** 1672-4292(2011)04-0047-06

Continuous Problem of Image Region Stereo Matching Algorithm

Xu Shanshan, Liu Ying'an, Xu Sheng

(College of Information Science and Technology, Nanjing Forestry University, Nanjing 210037, China)

Abstract: Traditional graph cut algorithm to solve the binocular matching problem is time-consuming while it requires great precision. This paper proposes a new algorithm by which the minimum cut graph cut problem is converted into the greedy algorithm to reduce the complexity of the problem. At the same time as the transformed graph cut has a low effect in dealing with the image continuous problem, we propose the graph cuts combined with the region matching algorithm called GR(Graphic Cut in Region) algorithm. They not only apply the Graph cut algorithm theory to the stereo matching problem, but also use a new method which computes the initial disparity cost through scaling the SAD window and twice compute the window cost to improve the accuracy of the initial disparity. The experiments show that this new algorithm in image continuous region has a better effect, and significantly improves the accuracy of matching at this time, and also that the algorithm complexity is greatly reduced.

Key words: graph cut, continuity of image region, minimization of energy function, region matching, scale the window cost

立体匹配是计算机视觉领域的热点问题,它是对图像中的每个像素,计算其在另一幅图像中的对应像素的过程,可以看作每个像素选择一个使之具有最小匹配价值的视差.通常根据视差数据的分布情况,可以将匹配算法分为以边界或角点等图像特征为匹配单元的稀疏视差匹配^[1]和以图像区域灰度作为匹配基元的密集视差匹配^[2].稀疏视差匹配通常都较鲁棒,一般情形下获得比较稀疏的结果;而密集视差匹配给出几乎所有像素的视差信息,但在边界区域会产生一些误匹配.

图割算法是一种密集视差匹配算法,其优点在于,不仅总体精度高,且在不连续区域和低纹理区域的精度也明显高于其他算法.即便有些算法如模拟退火算法^[3]的精度与图割算法的精度接近,但图割算法在优化过程中收敛效果更好,所以图割思想广泛应用于立体匹配中.但是求取图割中最小割问题时,计算复杂度太大,限制了该算法的应用.为此本文研究了基于贪心算法的 GR(Graphic Cut in Region)算法,基本思想是将简化图割与区域匹配算法相结合.实验发现该算法能较好地解决图像区域连续化的立体匹配问题.由于图割算法中初始视差的精度对最后的结果将产生重大的影响,所以在初始视差获得时采取了单位化窗口算法,以提高初始视差的精度.

收稿日期: 2011-08-10.

通讯联系人: 刘应安, 博士后, 教授, 研究方向: 数据挖掘、概率统计. E-mail: lyastat@yahoo.com.cn

1 预备知识

1.1 最小割

设 $G = (V; E)$ 为一个带有非负边权值的有向图,其中 V 为顶点集, E 为边集. V 中包含两个特殊的顶点(称为终端),通常一个称为源,一个称为汇.有向图 G 的一个割 C 就是边集 E 的一个子集,由它所导出的子图即可以表示为 $G(C) = (V, E)$, 其将图 G 的两个终端分离(即两个终端之间没有道路连接),而对 C 其任一子集的诱导图均不能将两个终端分离.割 C 的容量(记作 $|C|$) 定义为它的所有边的权值之和.最小割(minimum cut) 就是图 G 所有割中容量最小的割.

1.2 图割能量函数构造

立体匹配等价于一个能量函数最小化问题. 设输入为像素点的集合 P , 视差标注集合为 L , 目标是寻找一个标注 f (从 P 到 L 的一个映射), 使能量函数最小. 通常选择的能量函数具有下列形式^[4]:

$$E(f) = E_{\text{data}}(f) + E_{\text{smooth}}(f). \tag{1}$$

其中:

$$E_{\text{data}}(f) = \sum_{p \in P} D_p(f(p)), \tag{2}$$

$$E_{\text{smooth}}(f) = \sum_{\{p, q\} \in N} V_{p, q}(f(p), f(q)). \tag{3}$$

式中 f 为视差标注; P 为所有像素的集合; D_p 为数据项, 用来度量像素 p 所分配的标号 f_p 与观测数据的不一致程度; N 为邻域像素对的集合; $V_{p, q}(f_p, f_q)$ 度量将 f_p, f_q 分配给相邻的像素 p, q 的代价, 称为平滑项.

本文采用另外一种能量函数. 设 I_1, I_2 为两个灰度图像. 对 I_1 构造标号函数 f , 即 $f(p) \in \{l_1, l_2, \dots, l_n\}$ 为像素 p 的标号. 此处 $f(p)$ 为 I_1 中像素 p 在 I_2 中的可能视差. 本文采用王年^[5] 的构造能量函数方法:

$$E(f) = \sum_{p \in I_1} (| \ln_1(p) - \ln_2(p + f(p)) | + | \mathbf{Gr}_1(p) - \mathbf{Gr}_2(p + f(p)) |), \tag{4}$$

式中 $\ln_k(p), \mathbf{Gr}_k(p) (k = 1, 2)$ 分别表示像素 p 在第 k 幅图像中的灰度以及梯度向量. 注意到式(4)中并没有式(1)中的平滑项, 其实, 式(4)中关于梯度项 $| \mathbf{Gr}_1(p) - \mathbf{Gr}_2(p + f(p)) |$ 在本质上已经体现了平滑约束, 因为若 $p \in I_1$ 位于某区域的内部(或边界), 其在 I_2 中实际的匹配点 $p + f(p)$ 也应在匹配区域的内部(或边界), 此时 $\mathbf{Gr}_1(p)$ 和 $\mathbf{Gr}_2(p + f(p))$ 的距离应当很小. 故 $| \mathbf{Gr}_1(p) - \mathbf{Gr}_2(p + f(p)) |$ 体现了区域内部的连续性和边界的不连续性. 将相邻像素之间的作用定义为像素的梯度向量之间的距离, 主要是为了有利于网络的构造.

1.3 区域连续化问题

区域连续化问题的产生, 如图 1 所示. 图 1 中的左图经过偏移 d 之后得到了图 1 中的右图, 但在计算视差时候, 由于左图大部分区域灰度值完全相同, 所以将会导致其区域偏移值 d 相同, 而仅仅在白色边界区域有不同的偏移值, 这时候算出来的视差图会出现图 2 的结果.



图 1 区域连续化
Fig.1 Continuous regional

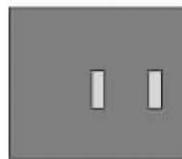


图 2 区域连续化视差图
Fig.2 The disparity map for the continuous regional

可以得出, 若是图割算法仅仅考虑单个像素点必会遇到此类问题, 所以尝试图割算法中添加区域匹配的思想.

2 改进能量函数

由于图像区域连续化问题的存在, 可以考虑像素点 p 与周围几个点共同作用, 若点在图像区域连续处则依靠周围的点对其约束, 从而得出该点的真实视差值. 若点处于图像的边界则采用原来的图割算法即

可. 能量函数如下:

$$E(f) = \sum_{p \in I_1} (| \ln_1(w) - \ln_2(w + f(w)) | + | \mathbf{Gr}_1(w) - \mathbf{Gr}_2(w + f(w)) |), \quad (5)$$

$$E(f_o) = \operatorname{argmin}_{p \in P} E(f). \quad (6)$$

式(5)中 w 不是单单的一个像素点,而是关于待求像素点 p 的一个集合.若该点 p 处于图像的边缘处,则集合 w 仅仅含一个元素,即点 p ;若该点处于图像连续处,则需通过计算才能得到 w .本文计算 w 的方法是看点 p 与周围的点灰度差异,分别考察点 p 与左右相邻的点的灰度值,方法如下:

p_i 表示像素点位置为第 i 个点,不妨设 p 此时处于第 k 个点即 p_k ; thresh为预先设定好的阈值.

$i = k$; (保存 p 的位置);

While $| \ln(p_i) - \ln(p_{i+1}) | < \text{thresh}$

$i = i + 1$;

即计算 p 点右边像素点灰度连续区域;

While $| \ln(p_i) - \ln(p_{j-1}) | < \text{thresh}$

$j = j - 1$;

即计算 p 点左边像素点灰度连续区域;

此时 $w = \{p_j, p_{j+1}, p_{j+2}, \dots, p_k, p_{k+1}, p_{k+2}, \dots, p_i\}$.

根据式(6),每个点 p 都能得到标号 f 使得能量函数最小,此时形成集合 f_o 即为所求视差.

3 构造网络

为了最小化上述能量函数(5),构建一个无向网络 $N^{[5]}$,使得标号函数集和网络 N 的(几乎所有的)图割构成双射,并且对于每一个标号 f ,其能量 $E(f)$ 即为其所对应的割的量.根据上述映射关系,能量函数(5)的最小化问题等价于 N 的最小割求解问题,最小割问题可以通过计算最大流来得到解决.

对于标号 L 中元素,不妨假设 L 中元素有下排列: $l_1 < l_2 < l_3 < \dots < l_n$.构建网络如下:

建立坐标系 $(O; X, Y)$,把图像 I_1 置于 OXY 平面,使得 I_1 的原点 X, Y 轴与 OXY 平面的原点以及相应的轴重合.在 Y 的正半轴上从原点开始等距的放置向量 $l_1, l_2, l_3, \dots, l_n, l_1 = 0$,令 $q_0 = l_1 = 0$.对 $i = 1, 2, \dots, n-1$ 在 (l_i, l_{i+1}) 的中点放置点 $q_i, q_n = l_n$.由此,由 OXY 平面中像素点集 w 以及 Y 的正半轴上的点 q_0, q_1, \dots, q_n 构成了一个网格.可以看出,对 $i = 0, 1, 2, \dots, n-1$,每个 Y 的轴上区间 $[q_i, q_{i+1}]$ 恰含一个 l_{i+1} .记 $(w; q_i)$ 为网格上节点, $N(p)$ 为像素点 p 邻域.在网格的上下方添加两个点 s, t ,并在 s 到 I_1 每个像素点之间添加一个边,在 t 到集合 $\{(w; q_n) \mid w \in I_1\}$ 中的每个点,即网格上与 OXY 平面相对立的另一个面上节点,添加一个边.由此获得一个无向图 $G = (V, E)$,如图3所示.其中, $V = \{s, t\} \cup \{(p; q_i) \mid p \in I_1, i = 0, 1, \dots, n\}$, $E = \{(s, (p; q_0)), (t, (p; q_n)) \mid p \in I_1\} \cup E_s \cup E_d$, $E_s = \{((p; q_i), (q; q_i)) \mid q \in N(p), p \in I_1, i = 0, 1, \dots, n-1\}$ 为光滑边构成的集合, $E_d = \{((p; q_i), (p; q_{i+1})) \mid p \in I_1, i = 0, 1, \dots, n-1\}$ 为视差边构成的集合,如图4所示.设 $H = (V, E \setminus E_s)$,即 H 为 $G = (V, E)$ 删除光滑边之后的图.定义 H 中各边的容量如下:

(1) $\{(s, (p; q_0)) \mid p \in I_1\}, \{(t, (p; q_n)) \mid p \in I_1\}$ 中边的容量为 $+\infty$ 或者充分大的正数;

(2) 视差边的容量为:对任意 $p \in I_1, i = 0, 1, \dots, n-1$,边 $e_{p, i+1} = ((p; q_i), (p; q_{i+1}))$ 的容量为:
 $c(e_{p, i+1}) = | \ln_1(w) - \ln_2(w + l_{i+1}) | + | \mathbf{Gr}_1(w) - \mathbf{Gr}_2(w + l_{i+1}) |$;

(3) 由此,构建了一个网络 $N = (H, s, t, c)$.

很容易发现最小化问题可转化为有效割集 $EC(N)$ 的最小割问题.而 $EC(N)$ 的最小割即为网络 $N = (H, s, t, c)$ 的最小割.因此,求出 $N = (H, s, t, c)$ 的最小割,此时便获得了式(5)的能量最小化,即求得立体匹配所需要的视差值.对于网络 $N = (H, s, t, c)$ 的每一个割 C ,它必分离 s, t .若 C 不含 s -边或 t -边(即与 s 或 t 关联的边),则对每个像素 p ,它至少经过集合 $\{((p; q_i), (p; q_{i+1})) \mid i = 0, 1, \dots, n-1\}$ (即由像素 p 出发的视差边)中的一条边.由于 s -边或 t -边的容量为 $+\infty$ 或者充分大,显然不可能为最小割.故仅考虑如下类型的割:它仅包含每个像素 p 出发的视差边中的一条边;称这种类型的割为有效割,记有效割的集合为 $EC(N)$.则必定存在割 $C \in EC(N)$, C 的容量为网络 $N = (H, s, t, c)$ 的最小割的容量.确定了 $N = (H,$

s, t, c) 的最小割,也就获得了 $EC(N)$ 的最小割. 则最小割所对应的标号 f' ,便满足 $E = \min E(f')$,从而获得了问题的求解. 由于构造网络的巧妙,没有设置视差边,则图转换为如图 5 所示的简单形式,该形式下最小割简化为贪心算法即可,由于图割求最小割一直很耗时,经过简化后,算法的复杂度将明显降低,精度也有所提高.

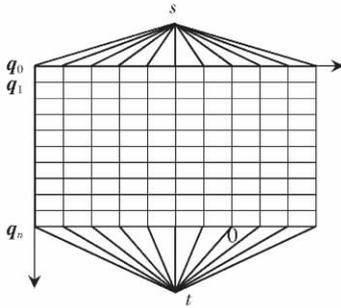


图 3 将图割映射的无向图

Fig.3 The graph cut mapped into an undirected graph

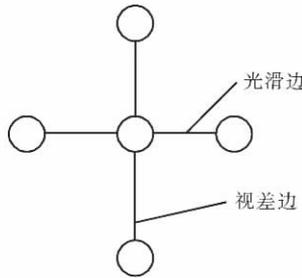


图 4 光滑边和视差边

Fig.4 Smooth edges and disparity edges

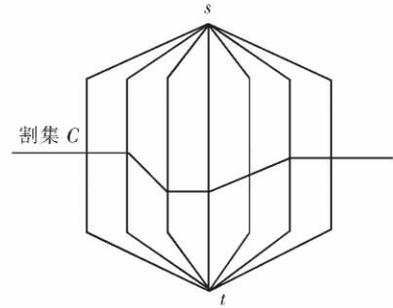


图 5 贪心图割算法运行图

Fig.5 Greedy graph cut algorithm diagram

4 窗口单位化算法

上述求解视差的算法称为 GR(Graphic Cut in Region) 算法,通过算法的分析可知道,转换的有向图的边权值即初始视差的设置很重要. 传统的算法中获得初始视差采用的是 SAD 算法,即:

$$\text{oldEdata} = \sum_{p \in I_1} | \ln_1(w) - \ln_2(w + f(w)) |. \quad (7)$$

只需求解窗口中的灰度差之和即可. 然而在统计各个不同像素点所对应的匹配代价 Edata 时候,若单单靠一个窗口,来算一个像素点的匹配代价,很难正确得出其真正匹配代价. 在通过传统 SAD 算法得出各个像素点不同视差所消耗的匹配代价后,需进一步考虑像素点周围其他像素点经过传统 SAD 算法后的匹配代价,即设定窗口 D_{win} 来得出该像素点与周围像素点共同的匹配代价. 由于其在对应不同视差值时候代价相差很大,所以需要不同的视差空间匹配代价进行单位化:

$$\begin{aligned} \text{cost}^1 &= \{ \text{Edata}(1, 1), \text{Edata}(1, 2), \dots, \text{Edata}(1, d_{\max}) \}, \\ \text{cost}^2 &= \{ \text{Edata}(2, 1), \text{Edata}(2, 2), \dots, \text{Edata}(2, d_{\max}) \}, \\ &\dots \\ \text{cost}^p &= \{ \text{Edata}(p, 1), \text{Edata}(p, 2), \dots, \text{Edata}(p, d_{\max}) \}, \end{aligned} \quad (8)$$

即 Cost^p 为像素点 p 在取不同视差值时的匹配代价集合, $\text{Edata}(p, d)$ 表示像素点位置 p 在视差为 d 上的匹配代价. 假设最大视差值为 d_{\max} , 则单位化不同视差空间的匹配代价如式 (9) 所示:

$$\text{Edata}(p, d) = \left\{ \text{Edata}(p, d) * \frac{1}{\max(\text{cost}^p) - \min(\text{cost}^p)} \right\}. \quad (9)$$

在单位化后,可同时考虑像素点 p 和周围像素点的匹配代价来得出新的 newEdata,从而减少误匹配:

$$\text{newEdata}(p) = \sum_{k=p-D_{win}/2}^{p+D_{win}/2} \text{Edata}(k). \quad (10)$$

$\text{Edata}(k)$ 即为经过传统的 SAD 算法运算后的各个像素点的匹配代价. 对此代价经单位化后考虑 D_{win} 窗口内的像素点的匹配代价,从而确定 D_{win} 窗口中心像素点 p 的匹配代价. 实验证明其精度比起一次窗口明显增加.

5 实验

5.1 标准数据集

首先在标准数据集上对算法进行测试,标准数据集来自 <http://vision.middlebury.edu/> 网站的立体匹配标准数据集. 选用了 Tsukuba 图片数据集 (384 * 288) (如图 6 所示) 和 Venus 图片数据集 (434 * 383)

(如图7所示) 图片均经过校正,去除了干扰噪声,消除了垂直方向的视差.将结合窗口单位化后的GR算法运用在这两幅图片上,所得结果如图8所示,可见算法在大大降低复杂度同时,能得到较好的视差图.图中出现的亮点是由于窗口Dwin的大小设定问题,若窗口太大则边界将会模糊,若窗口较小则出现斑点,所以窗口选择问题是今后需重点考虑的部分.表1分析了这两幅图片的精度,给出了无遮挡区域(non-occluded)的错误率,测评采用文献[2]中的评估方法.可见,在无纹理区域较多的Venus图上可取得较好的效果,而在无纹理区域较少的Tsukuba改进并不明显.图中白色区域为遮挡区域,计算精度的时候不予考虑.

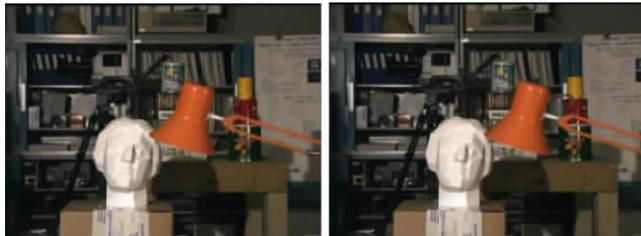


图6 Tsukuba的左图与右图
Fig.6 The Tsukuba images



图7 Venus的左图与右图
Fig.7 The Venus images

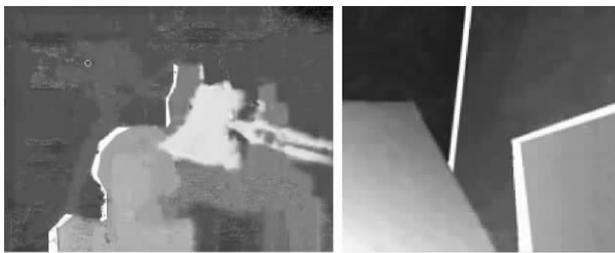


图8 Tsukuba的左图与右图
Fig.8 The Tsukuba images

表1 Tsukuba与Venus的精度

Table 1 The accuracy for the Tsukuba and Venus

图片	SAD	GC	GR	GR + Dwin
Tsukuba	9.03	2.04	7.46	6.16
Venus	12.33	1.90	5.54	3.02

5.2 测试数据集

以门廊合成图像为研究对象.该图像未经过同标准数据集一样的仔细校正,垂直方向存在一定的偏移,而且无纹理区域较多,用GR算法和其他算法进行比较,可以进一步体现出GR算法的优越性.

如图9所示,在面对门廊图时,由于图像的区域连续性,用传统的一些方法将会得到错误的视差.图中(a)与(b)为门廊图片左右不同视角所拍的图片,(c)为门廊图片的真实视差,(d)为门廊图片的真实深度.如(c)所示,由真实的视差图可以反应距离视角近的地方图像亮,而远的地方较暗,且物体的轮廓边缘很明显,能真实地显示出物体,图像较光滑连续.如(d)所示,可进一步地反应图像偏移,将其与真实视差图与深度图进行比较.(e)为基于GR算法所得的视差图,可以看出物体边缘轮廓相当明显,但当物体在图像区域灰度值相同的时候,如地板内部,则算法效果很差.(f)为区域匹配算法,其能很好地处理区域灰度值内部相同的视差,但对于物体的轮廓边界显得模糊甚至消失.与区域匹配算法相似,Alvarez^[6]提出了基于能量匹配的能量方程算法,也能处理灰度值相同的问题,但如(g)所示,存在很多视差计算错误之处,如球体亮度过强,且边界也较为模糊.(h)为本文采用的将区域匹配算法与图割算法相结合的GR方法,经过两次窗口代价的计算,使其既能处理图像区域连续化问题,又可保持轮廓的清晰.

6 结论

本文针对图割算法复杂度过高且难以满足对时间的需求,研究了基于贪心图割的处理方法.由于图像区域连续化效果低下,又提出使其与区域匹配算法相结合的思想.考虑到图割算法初始视差对结果的影响很大,提出了窗口单位化来提高匹配视差的精度,使其成为处理像素集合的新的图割方法.通过实验验证,本文算法在处理图像区域连续化问题的时候能得到较好的结果,且算法复杂度较小,是一种行之有效的图像匹配算法.

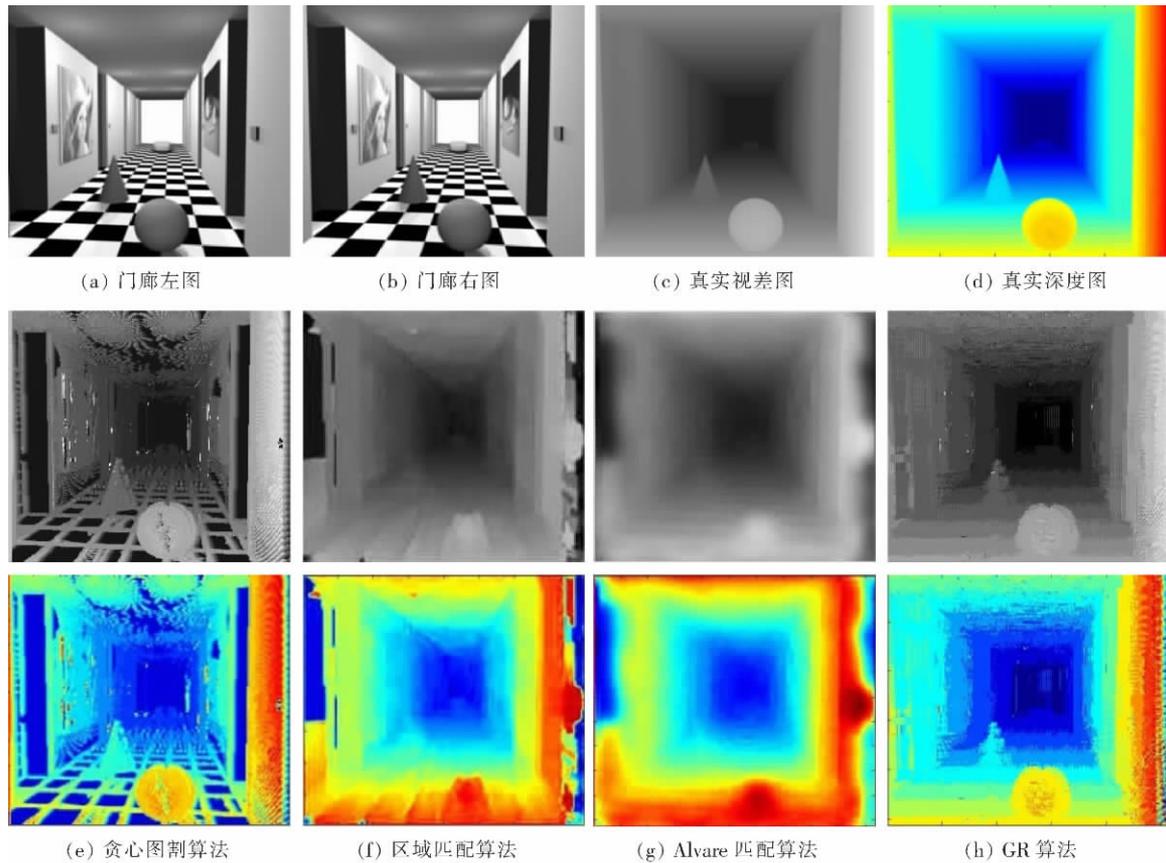


图 9 门廊匹配实验

Fig.9 Matching experiment

[参考文献](References)

[1] Hsieh Y C ,D McKeown ,F P Perlant. Performance evaluation of scene registration and stereomatching for car-tographic feature extraction [J]. IEEE Trans ,1992 ,PAMI-14(2) :214-238.

[2] D Scharstein ,R Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms [J]. International Journal of Computer Vision ,2002 47(1) :7-42.

[3] Sun Dongmei , Qiu Zhengding. A new non-rigid image matching algorithm using thin-plate spline [J]. Acta Electronica Sinica , 2002 ,30(8) : 1 104-1 107.

[4] Boykov Y ,Kolmogorov V. An experimental comparison of mincut/maxflow algorithms for energy minimization in vision [J]. IEEE Transon Pattern Analysis and Machine Intelligence ,2004 26(9) :1 124-1 137.

[5] 王年 ,范益政 ,鲍文霞. 基于图割的图像匹配算法 [J]. 电子学报 2006 34(2) :232-236.
Wang Nian ,Fan Yizheng ,Bao Wenxia. Graph matching algorithm base on graph cut [J]. E-Journal 2006 34(2) :232-236. (in Chinese)

[6] Alvarez L ,Deriche R ,S´anchez J ,et al. Dense disparity map estimation respecting image derivatives: a PDE and scale-space based approach [J]. Journal of Visual Communication and Image Representation ,2002 ,13(1/2) : 3-21.

[责任编辑: 严海琳]