

# 基于最大窗口的二次立体匹配方法<sup>\*</sup>)

夏永泉 杨静宇

(南京理工大学计算机科学系 南京 210094)

**摘要** 文中提出了一种新颖的基于窗口的立体匹配方法,该方法首先在最大窗口内估计视差,并假设该窗口内视差一致,在此基础上得到两个最大的匹配窗口,然后在两个最大窗口内进行二次匹配,得到基于这两个窗口的逐像素视差,估计视差和像素视差之和就是结果视差。对相对较平滑或平滑均匀的区域,在匹配过程中,会出现匹配最小多值问题,面临如何确定最佳匹配。本文算法采用平滑性测度指标函数来屏蔽平滑或平滑均匀区域,并在匹配完成后,按照最近邻视差均值来估计平滑区域的视差。最后通过立体图像对算法进行了测试。实验结果表明,该方法是可行和有效的。

**关键词** 立体匹配,视差估计,窗口匹配,视差更新,最小多值

## Two Level Stereo Match Approach Based on Maximum-Window

XIA Yong-Quan YANG Jing-Yu

(Department of Computer Science & Technology, NJUST, Nanjing 210094)

**Abstract** A novel window-based stereo matching approach is proposed in this paper. The approach estimates the coarse disparity in maximum window in left and right image according most-min absolute difference method firstly. As result, two maximum windows are established, after then, the second match is done in the two max windows and obtain fine disparity, the result disparity is the sum of coarse disparity and fine disparity. For the smooth or wipe regions, the problem of cost function's multiple most minimums is appeared in the process of stereo matching. In order to overcome it, threshold smooth cost function is presented to avoid the smooth or wipe balance regions. After completing the matching of no-wipe or non-smooth regions, the unmatched wipe and smooth regions are estimated by the average of most adjacent disparity. At last, several stereo images are used to test the algorithm. The results indicate that the approach is valid and possible.

**Keywords** Stereo matching, Disparity estimation, Window match, Disparity update, Multiple most minimums

## 1 引言

立体匹配是机器视觉领域的重要研究方向之一,其中心问题是抽取深度信息,得到视差图。可分为单目、双目、三目以及多目立体匹配,其中双目立体匹配得到了广泛的应用,因为相对来说双目立体匹配更容易实现。

在基于窗口的立体匹配方法中,一些文献采用了不同的策略,例如固定大小窗口<sup>[1]</sup>、自适应可变窗口<sup>[2]</sup>,以及多窗口<sup>[3,4]</sup>等策略。文[1]的方法假设窗口内的视差是一致的,这种方法简单,但误匹配发生的概率增大;文[2]使用了视差估计不确定性作为窗口匹配费用,该方法需要窗口内的视差变化模型和初始视差估计。文[3,4]对每个像素利用几个不同的窗口联合作为该像素的测度。本文的算法提出了一种新的基于窗口的二次匹配方法,首先在最大视差范围内寻找最大匹配窗口,得到整个图像的估计视差,在此基础上将该最大窗口内的数据划分为两个集合:非平滑区域集合和平滑区域集合。原因是平滑区域的匹配很容易陷入匹配函数最小多值问题,在匹配过程中,首先不对平滑区域进行匹配。而在平滑区域匹配完成后,按照连续性约束条件对平滑区域的视差进行估计。实验用的图像是经过校正的灰度图立体图像对,采用从左到右的匹配策略,并以右图像为基准。

## 2 视差获取

用  $f_L(x, y)$  和  $f_R(x, y)$  分别表示大小为  $M \times N$  的左、右

离散图像处的灰度值函数,其中  $(x, y) = \{(u, v) : u \in \{0, M-1\}, v \in \{0, N-1\}\}$ 。  $d(x, y)$  表示右图像在  $(x, y)$  处的视差。因为实验图像是经过校正的立体图像,那么根据双目立体匹配原理有下式成立:

$$C_1(x, y, d) = f_L(x+d(x, y), y) - f_R(x, y) \approx 0 \quad (1)$$

但在实际应用中,由于传感器噪声、平滑等因素的影响,  $C_1(x, y, d)$  作为单独的匹配费用函数不能得到尽量正确的视差图。正是因为立体匹配的复杂性,出现了很多立体匹配方法,但任何方法都离不开(1)式的基本匹配原理。

### 2.1 最大窗口视差估计

最大窗口视差估计是在(1)式原则下,设图像的整体估计视差为  $E_d$ ,  $d_{\max}$  为允许的最大视差,左图像固定不动,将右图像在左图像上水平移动  $d$ ,  $d \in \{-d_{\max}, d_{\max}\}$ , 计算重叠部分的错误均值:

$$C_2(d) = \sum_{y=0}^{M-1-N} \sum_{x=0}^{d-1} |f_L(x+d, y) - f_R(x, y)| \quad (2)$$

那么整个窗口的估计视差为:

$$E_d = \arg \min_x C_2(d) \quad (3)$$

于是分别得到左右图像的最大匹配窗口为:

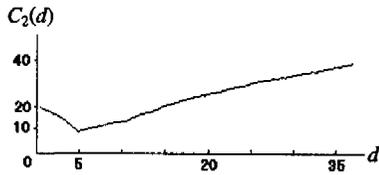
$$W_L(x, y), \text{ 其中 } E_d \leq x \leq M-1, 0 \leq y \leq N-1$$

$$W_R(x, y), \text{ 其中 } 0 \leq x \leq M-1-E_d, 0 \leq y \leq N-1$$

例如,对 Tsukuba 图像进行视差估计,图 1(a)是在窗口范围  $\{-40, 40\}$  内对不同的  $d$  所求得平均错误变化图。由图 1(b)可以看出,随着  $d$  的增大,在  $d > 5$  时,估计视差收敛

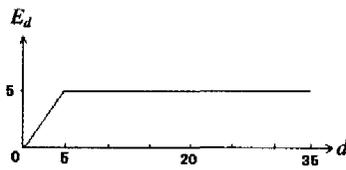
<sup>\*</sup>) 本课题得到国防基础研究项目(编号:J1500C002)基金资助。夏永泉 博士研究生,主要研究方向为计算机视觉、模式识别与人工智能;杨静宇 教授,博士生导师,主要研究领域为计算机视觉、信息融合、模式识别、智能机器人。

到 5, 于是在  $d=5$  处,  $C_2(d)=10$  达到最小值, 于是得到估计



(a) 均值错误

视差为  $E_d=5$ 。



(b) 视差估计

图 1 均值错误和估计视差

$W_L(x, y)$  和  $W_R(x, y)$  就是两个最大的匹配窗口。为了后面描述算法的方便, 用  $f'_L(x, y)$  和  $f'_R(x, y)$  表示  $f_L(x, y)$  和  $f_R(x, y)$  在窗口  $W_L(x, y)$  和  $W_R(x, y)$  内的像素集合:

$$f'_L(x, y) = f_L(x + E_d, y)$$

$$f'_R(x, y) = f_R(x, y)$$

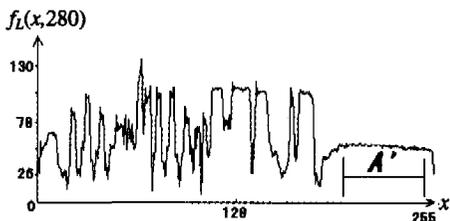
### 2.2 视差更新

经过最大窗口视差估计式(3)得到  $E_d$ , 并假设窗口内的视差一致为  $E_d$ 。显然, 在实际处理中, 这个估计的视差  $E_d$  是不能作为最后的视差结果, 需要进行进一步的视差更新处理。设匹配搜索窗口为  $W_{max}$ , 假设图像经过校正, 那么对  $(x, y)$  处的视差, 需要在另外一副图像中  $\{(x - W_{max}, y), (x + W_{max}, y)\}$  范围内寻找最佳的匹配。

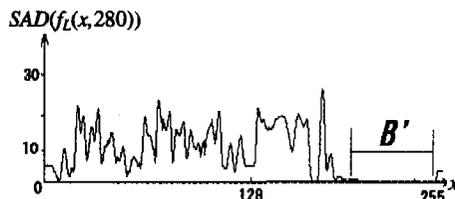
设匹配窗口大小为  $m \times n$ , 则  $(x, y)$  处视差为  $d$  的费用函数为:

$$C_3(x, y, d) = \frac{1}{mn} \sum_{w_x=-m/2}^{m/2} \sum_{w_y=-n/2}^{n/2} |f'_L(x + w_x + d, y + w_y) - f'_R(x + w_x, y + w_y)| \quad (4)$$

在视差范围  $\{0, d_{max} - E_d\}$  范围内寻找满足  $C_3(x, y, d)$  最小的  $d$  即为  $f'_R(x, y)$  在  $(x, y)$  的视差,  $d + E_d$  即为  $f_L(x, y)$  在  $(x, y)$  处的视差, 表示为:



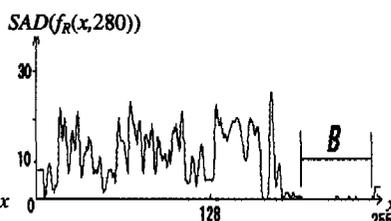
(a) 左图原始数据



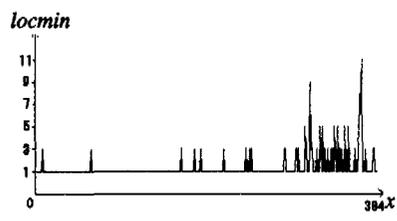
(b) SAD



(c) 右图原始数据



(d) SAD



(e) 匹配最小多值

图 2 灰度图和 SAD

图 2 中(a)、(c)分别为左、右图像第 280 行的原始数据, (b)、(d)分别为对应的 SAD 图, (e)为匹配最小多值图。可以看出, 右图的 A 区域在对应的左图 A' 区域比较平滑, 分别对应的 SAD 图中 B 和 B' 接近于 0。这样, 在匹配时非常容易陷入匹配最小多值问题。图(e)是用式(5)对左、右图像第 280 行进行匹配的最小多值图, 可以看出在寻找 A 区域的对应匹配时, 匹配最小多值问题显得尤为突出。对类似区域 A 部分

$$d_1(x, y) = \arg \min_d (C_3(x, y, d) + E_d) \quad (5)$$

### 2.3 最小多值问题

在立体匹配中, 平滑区域或较平滑区匹配是个困难的问题。因为像素数据的平滑性, 容易造成匹配函数最小多值问题。就是说, 在候选匹配中, 有多个测度值是相同的, 那么采用什么样的方法来选择最佳匹配? 本文算法采用了这样的方法来解决最小多值问题, 首先判断该区域是否是平滑区域? 如果是, 则将其标记不进行匹配, 最后按照最近邻已经匹配处的视差来估计该点的视差, 为此引入一个简单的平滑性测度判断函数。对大小为  $m \times n$  的匹配窗口来说, 计算该窗口内的像素灰度均值, 然后计算该窗口的绝对差和。

窗口灰度均值表示为:

$$\bar{A}_w(f(x, y)) = \frac{1}{mn} \sum_{w_x=-m/2}^{m/2} \sum_{w_y=-n/2}^{n/2} f(x + w_x, y + w_y) \quad (6)$$

那么, 绝对差和可以表示为:

$$SAD(f(x, y)) = \frac{1}{mn} \sum_{w_x=-m/2}^{m/2} \sum_{w_y=-n/2}^{n/2} |f(x + w_x, y + w_y) - \bar{A}_w(f(x, y))| \quad (7)$$

例如对  $3 \times 3$  的窗口, 对 Tsukuba 图像第 280 行用(6)和(7)式计算得到的结果如图 2 所示。

匹配时, 可以选取适当的 SAD 阈值来屏蔽该区域。如图(d)中, 阈值为 5 时, B 区域将被屏蔽不做处理。最后根据最近邻视差来估计该区域的视差。

### 2.4 结果视差

基于匹配最小多值问题的讨论, 在匹配过程中, 为了解决这个问题, 需要对平滑区域或非平滑区域另做处理。本算法对平滑区域或平滑较均匀的区域和非平滑区域分开进行匹

配,先计算非平滑区域的视差,然后按照最近邻视差估计平滑区域的视差。为此,定义两个集合: $S_1(x,y)$ 和 $S_2(x,y)$ ,分别表示非平滑和平滑区域。

$$S_1(x,y) = \{f'_R(x,y) : |SAD(f'_R(x,y)) \geq \delta\}$$

$$S_2(x,y) = \{f'_R(x,y) : |SAD(f'_R(x,y)) < \delta\},$$

$\delta$ 为阈值。如图3所示,黑色区域为 $S_2(x,y)$ ,灰色区域为 $S_1(x,y)$ 。其中(a)为实验1用图像, $\delta=5.35$ ; (b)为实验2用图像, $\delta=5.5$ 。本算法首先对 $S_1(x,y)$ 集合内像素进行匹配,然后根据 $S_1(x,y)$ 的视差来估计 $S_2(x,y)$ 区域内像素的视差。

对集合 $S_1(x,y)$ 内像素视差,按照式(5)计算;对 $S_2(x,y)$ 集合内像素视差的求取,按照下式进行:

$$d_2(x,y) = \frac{1}{k} \sum_{i=1}^k d_i \quad (8)$$

其中 $(x,y) \in S_2(x,y)$ ,  $d_i \in d_1(x,y)$ ,  $d_i$ 为与 $(x,y)$ 最近邻视差,  $k$ 为最近邻视差的数量,文中算法中采用了最近邻视差简单均值来估计平滑区域的视差,也可以按照其它的方法进行插值处理。于是,结合式(5)和式(8)得到最后的视差函数为:

$$d(x,y) = \begin{cases} d_1(x,y), & (x,y) \in S_1(x,y) \\ d_2(x,y), & (x,y) \in S_2(x,y) \end{cases}$$

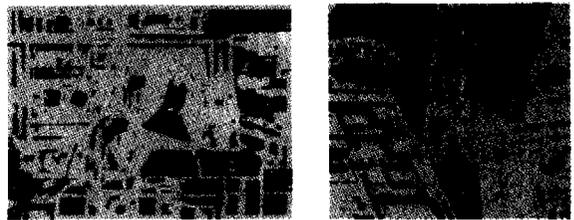


图3 平滑区域和非平滑区域

### 3 实验结果

为了验证算法的有效性,采用了 Sawtooth 和 Tsukuba 立体图像进行实验。实验1采用了 Tsukuba 立体图像对,大小为  $380 \times 330$ ,采用  $7 \times 7$  的匹配窗口,得到的视差结果如图4(c)所示。实验2采用了 Venus 立体图像对,大小为  $384 \times 288$ ,采用  $11 \times 11$  的匹配窗口,得到的结果如图5(c)所示。

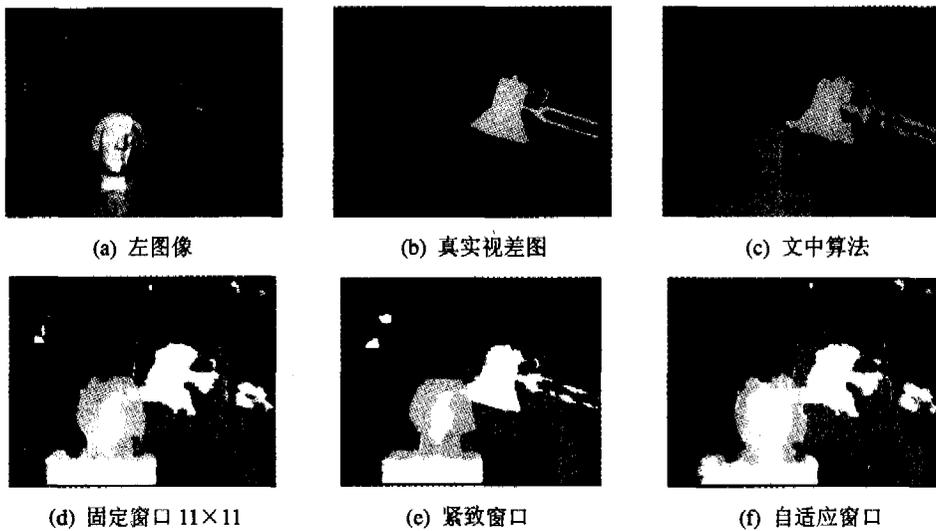


图4 实验1



图5 实验2

图4中(d)为采用固定窗口得到的视差图,(e)为文[5]方法得到的视差图,(f)为文[2]方法得到的视差图,(c)为本文算法得到的视差图。本文算法采用了二次匹配的方法,在二次匹配时,采用了固定窗口的方法,也可以采用文[2]的自适应窗口方法和文[3]的多窗口等方法,最后得到的视差图根据视差连续性约束,通过简单的局部窗口对结果视差图进行了平滑。另外,算法对平滑区域阈值的选取是根据实验来得到的。能否根据图像的灰度特征自动获取阈值,则是以后要讨

论和解决的问题。

**结论** 本文根据基于窗口的立体匹配原理和关于基于窗口匹配文献的基础上,提出了一种新的基于最大窗口的二次立体匹配方法,该方法利用最大窗口视差估计的方法,以左右两副图像相关窗口最小均值错误来对整个图像进行粗匹配,然后以该窗口为基础进行二次匹配,得到细节的视差,并对粗的估计视差进行更新,得到最后的结果视差。通过几个实验来验证算法的有效性,实验结果表明,该方法是可行和有效

的。

### 参考文献

- Okutomi M, Kanade T. A multiple-baseline stereo. In: Proceeding IEEE Conference on Computer Vision and Pattern Recognition, 1991. 63~69
- Kanade T, Okutomi M. A Stereo Matching Algorithm with an Adaptive Window: Theory and Experiment. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1994, 16(9): 920~932

- Fusiello A, Roberto V, Trucco E. Efficient stereo with multiple windowing. In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition, 1997. 858~863
- Geiger D, Ladendorf B, Yuille A. Occlusions and binocular stereo. In: Proc. European Conf. on computer Vision, 1992. 425~433
- Veksler O. Stereo matching by compact windows via minimum ratio cycle. In: Proceeding International Conference on Computer Vision (ICCV), 2001, 1: 540~547

(上接第 201 页)

(800 篇)、教育(800 篇)、法律(800 篇)、艺术(800 篇)、体育(800 篇)和军事(800 篇)的六个类别的中文网页,本文在实验中分别用模糊 C-均值聚类算法、ISODATA 算法和 FPCM 算法进行聚类,迭代次数分别为 5、8 和 3,然后用 Naive Bayes 增量学习方法分类,其分类的准确率如图 1 所示。

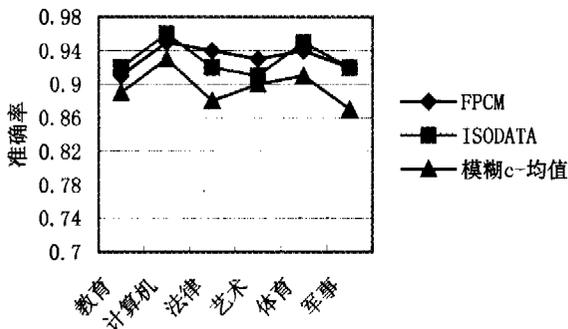


图 1 分别用 FPCM、ISODATA 和模糊 C-均值三种聚类方法结合 Naive Bayes 增量学习进行无标记 Web 文本分类

从实验中我们得出,FPCM 和 ISODATA 聚类结合 Naive Bayes 增量分类的结果准确率比较相近,模糊 C-均值聚类结合 Naive Bayes 增量分类准确率比前两个方法要差。在实验中我们还观察到在整个分类的过程中,采用 FPCM 先聚类的分类速度最快,其次是模糊 C-均值,ISODATA 的速度相对较慢。在本文对聚类的应用精度要求不是很高的情况下,FPCM 无疑是最好的选择,因为 FPCM 比模糊 C 均值方法的精度高,比 ISODATA 的速度快。因此,在无标记文本分类中,用 FPCM 聚类作为 Naive Bayes 分类的前期标记工作,能达到比较满意的分类结果。

以下我们用 FPCM 结合 Naive Bayes 增量学习分类的结果与模糊 C-均值聚类、ISODATA 聚类的结果进行比较,如图 2 所示。

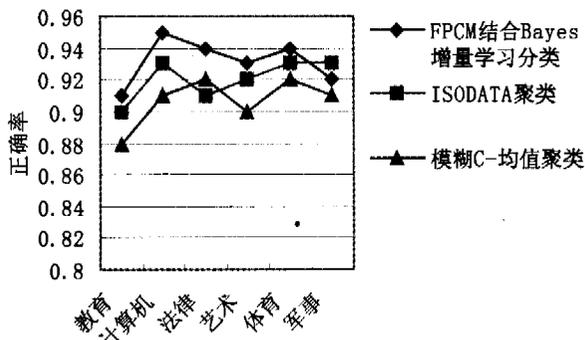


图 2 FPCM 结合 Bayes 增量学习分类结果和 ISODATA 聚类、模糊 C-均值聚类结果的比较

表 1 FPCM 与 Naive Bayes 增量学习分类相结合的结果

|      | 教育    | 计算机   | 法律    | 艺术    | 体育    | 军事    |
|------|-------|-------|-------|-------|-------|-------|
| TP   | 705   | 752   | 758   | 762   | 750   | 758   |
| FN   | 69    | 38    | 43    | 57    | 46    | 61    |
| FP   | 95    | 48    | 42    | 38    | 50    | 42    |
| Pre. | 91%   | 95.3% | 94.6% | 93%   | 94.1% | 92.5% |
| Rec. | 88.1% | 94%   | 94.8% | 95.2% | 93.6% | 94.8% |

从实验中,模糊划分聚类 FPCM 的迭代次数  $t$  为 3,控制模糊聚类参数  $m=2$ ,停止阈值  $0.4 \leq \lambda \leq 0.6$ , $\lambda$  越小,聚类越细,模糊度越小,聚类精度越高,且所需要的时间也就越长,所以在本文中我们没有将  $\lambda$  选得很小,尽量减少计算时间。通过实验我们观察到 FPCM 与 Naive Bayes 增量学习相结合的分类方法,比单纯的两种模糊聚类的分类准确度高,分类性能好。表 1 给出了 FPCM 与 Naive Bayes 增量学习相结合的实验结果。

**结束语** 在 Web 文本分类中,获得已标记训练样本集的成本是很高昂的。本章针对无标记训练样本的文本分类问题进行了研究,提出了模糊划分聚类 FPCM 与 Naive Bayes 增量学习相结合的分类方法。

本文利用无监督聚类不依赖于预先定义的类和带类别标记的训练实例的特点,对模糊聚类进行改进,提出了模糊划分聚类方法 FPCM。使用该方法通过对 Web 无标记文本的模糊聚类,得到少量标记文本,从而为监督学习找到了分类依据,与 Naive Bayes 增量学习方法的结合,提高了文本分类精度,对候选样本的平衡选取,进一步增强了分类器的性能。对于这个模型,我们进一步的工作是要解决经验选取初始点的敏感问题,以及尽量减小由于数据的分散性所带来的对孤立点的敏感程度,减少人工的干预,同时又不能增加太多的计算复杂度,在不同规模的数据集上测试该方法的有效性,以构造更高性能的分类器。

### 参考文献

- Linoff G S, J. a. Berry M. Mining the web, America, 2001, 348
- Mena J. Data Mining your website. America, 2000, 368
- Wang Shi, Gao Wen. Web data mining. Computer Science, 2000, 27(4): 237~240
- Hutter M. Distribution of Mutual Information. In: Proc. of the 14th Intl. Conf. on Neural Information Processing Systems, NIPS-2001
- 边肇祺,张学工,等编著,模式识别(第二版),清华大学出版社,2000
- Keogh E J, et al. Learning Augmented Bayesian Classifiers: A Comparison of Distribution-based and Classification-based Approache, 2002 <http://citeseer.nj.nec.com/context>.
- 官秀军,等. 主动贝叶斯网络分类器. 计算机研究与发展, 2002, 5: 574~579