

# 基于置信度融合的边缘检测性能评价基准图计算

李张帆, 张必银, 张天序

(华中科技大学图像识别与人工智能研究所, 武汉 430074)

**摘要:** 针对边缘检测算法的性能评价, 提出了一种在真实场景下基于统计、自适应强的评价基准图计算方法。定义了算法参数的相关系数, 研究了算法参数相关模型。根据建立的算法参数相关系数得到单算法的预选基准图, 对不同算法预选基准图进行基于置信度的图像融合得到可用于边缘检测算法性能评价的基准图。该计算基准图的方法有助于实现边缘检测算法性能评价的自动化, 实验结果验证了该方法的有效性及其实用性。

**关键词:** 置信度; 相关性; 图像融合; 边缘检测性能评价; 基准图

## Ground-Truth Calculation for Edge Detection Evaluation by Image Fusion Based on Confidence Measures

LI Zhang-fan, ZHANG Bi-yin, ZHANG Tian-xu

(Institute of Pattern Recognition and Artificial Intelligence, Huazhong University of Science and Technology, Wuhan 430074)

**【Abstract】** This paper proposes a new method for calculating the ground truth image by image fusion, and designs the calculation approach for the processing. Using the correlation coefficient between the different detection results, reference ground truth(RGT) images for each detector is obtained. The ground truth (GT) images are calculated by image fusion based on confidence measures, using the RGTs. The method suggests a general tool to assist in practical evaluation of parametric edge detectors where an automatic process is required.

**【Key words】** confidence measures; correlation; image fusion; edge detection evaluation; ground-truth image

边缘检测算法的性能评价方法主要分为主观评价和客观评价, 主观评价复杂且结果多变, 客观评价有利于对边缘检测算法性能的定量和定性分析。根据在评价过程中是否使用基准参考图, 将客观评价划分成两类<sup>[1]</sup>。没有使用基准图的评价方法不易确定检测到的边缘和真实边缘的位移差、检测误判率和漏检率, 而基于基准图的可以定量评价算法性能。

基准图是评价边缘检测算法性能的重要依据, 为保证基准图中含有图像边缘的准确信息(ground truth, GT), 通常采用合成图像<sup>[2]</sup>中预先确定的边缘轮廓图作为基准图或通过手工绘制边缘<sup>[3]</sup>的方法产生基准图, 通过将算法检测结果和基准图的边缘信息比较, 可以定性和定量分析边缘检测的效果, 从而对算法进行评价。虽然使用合成图像作为测试原图能得到精确的基准图, 并能在算法设计时期方便地反馈其效果, 但是考虑真实场景的复杂性和各种噪声对成像的影响, 基于合成图像的评价结果不一定能反映它在真实场景下的性能, 因此, 近几年出现了基于自然场景的通过人工绘制产生基准图的性能评价方法。

### 1 基准图制备方法

基准图被用来和实际的检测结果做匹配, 作为边缘检测算法性能分析的依据。本文提出的制备方法是通过分析已有算法集预定参数组合的检测结果的相关性, 产生单算法预选图像集, 然后对其进行基于置信函数的数据融合得到最终基准图。被检测的原始图像(输入图像)来自真实场景, 选择的5种经典边缘检测算法Canny<sup>[4]</sup>、Nalwa<sup>[5]</sup>、Iverson<sup>[6]</sup>、Bergholm<sup>[7]</sup>和Rothwell<sup>[8]</sup>覆盖了目前边缘检测的不同方法, 每个算法定义

不同的组合参数, 使用参数相关性统计的方法为每个算法选择最优参数。

#### 1.1 算法参数相关性

由于不同图像的边缘强度分布不一致, 噪声的分布也不同, 对所有图像使用固定的算法参数是不合适的。对每个算法, 取有限组参数组合  $S_1, S_2, \dots, S_n$ , 其中, 鲁棒参数  $S_b$  的检测结果和基准图最匹配, 参数  $S_b$  应当根据不同图像自适应地变化。不同的参数组合导致检测算法得到不同结果, 但是这种不同结果间并不是相互独立的, 而是相互关联的。通过统计试验提取出相关信息, 用做边缘点判定依据。参数的取值范围应足够的广泛, 能覆盖检测的各种结果, 参数覆盖范围越广、分布越细, 制备基准图越准确。根据原图像的亮度、密度以及大小应取不同参数组合。原始图像  $I$  中像素  $I(i, j)$  的参数相关系数为  $PC(i, j)$

$$PC(i, j) = \sum_{k=1}^n E_k(i, j) \quad (1)$$

对第  $k$  组参数  $S_k$ , 当边缘检测算法判定图像  $I(i, j)$  处为边缘  $E_k(i, j) = 1$ , 否则  $E_k(i, j) = 0$ 。相关系数分布在  $1 \sim n$  之间, 当  $n$  组检测结果同时判断原图  $I(i, j)$  处为边缘, 那么该像素的相关系数  $PC(i, j) = n$ ; 如果只有一组检测结果判断原图中某个像素点为边缘点, 则该像素的相关系数  $PC(i, j) = 1$ ; 没有算

**基金项目:** 国家自然科学基金资助项目(60135020)

**作者简介:** 李张帆(1982-), 男, 硕士研究生, 主研方向: 图像分析, 模式识别; 张必银, 博士研究生; 张天序, 教授、博士生导师

**收稿日期:** 2006-09-26 **E-mail:** lizhangfan@gmail.com

法判断这个像素为边缘,则相关系数  $PC(i,j) = 0$ 。拥有较高相关系数的像素灰度值较高并且清晰,作为图像  $I$  的预选边缘点,而相关系数较低的像素更有可能属于背景或噪声。算法的不同参数的相关系数概率分布可用来判断那些被算法判断为边缘点的像素的可靠性。比如对于 golf car 图像(图 1), Bergholms 和 Rothwell 算法分别取 12 组参数,得到相关系数分布如图 2、图 3 所示。



图 1 golf car 原图

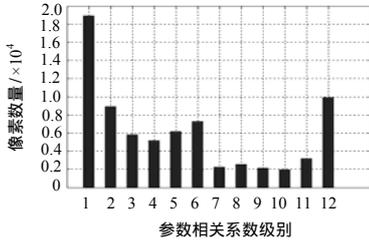


图 2 Bergholms 算法参数相关系数分布

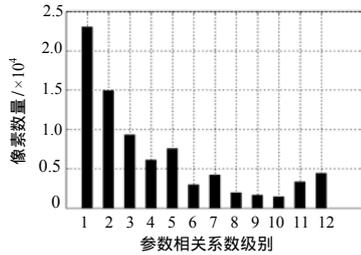


图 3 Rothwell 算法参数相关系数分布

### 1.2 预选基准图的生成

对单个算法,从  $n$  组参数的检测结果中,将被算法判为边缘点的像素(即相关系数  $PC(i,j) > 0$  的像素 ( $1 \leq PC \leq n$ )),按相关系数划分成  $n$  个相关系数等  $CT_1, CT_2, \dots, CT_n$  (correspondence threshold, CT),每级作为像素筛选的阈值,相关系数大于该阈值的像素组成预选基准图(reference ground truth Image, RGT)。将  $CT_i$  对应的  $RGT_i$  ( $0 \leq i \leq n$ ) 分别和所有  $n$  组参数得到的边缘检测结果  $D_j$  ( $0 \leq j \leq n$ ) 进行匹配,如图 4 所示。

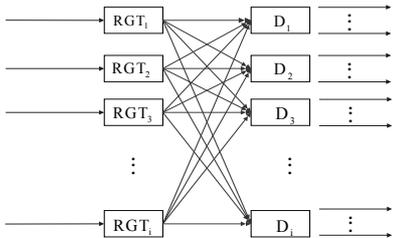


图 4  $RGT_i$  与  $n$  组参数组对应检测结果集的匹配

匹配的结果按照表 1 所示分类(NonE 为非边缘;E 为边缘)。算法判断为边缘的像素可以分成 2 个子集:真边缘集(true positive, TP)和假边缘集(true negative, TN)。真边缘集对应场

景中的真实边缘,假边缘集不是场景中的真实边缘,而是边缘检测器误判的边缘。在算法判断为非边缘的像素中,场景中的非边缘点且被算法判断为非边缘的像素称之为假阴性边缘(false negative, FN),而算法漏检的边缘则称之为假阳性(false positive, FP)。每个  $RGT_i$  分别与所有  $n$  组参数的检测结果集  $D_j$  ( $0 \leq j \leq n$ ) 分别匹配后的各个概率平均值用式(2)~式(5)计算。对于大小为  $K \times L$  的图像,使用 1 表示边缘像素;0 表示非边缘像素。

表 1 匹配结果的分类

检测正确的边缘子集		
真边缘集 概率	$TP_{RGT_i, D_j} = P(E_{RGT_i} / E_{D_j})$	像素在 $RGT_i$ 中判为边缘,在 $D_j$ 中也判为边缘
假边缘集 概率 $\times 10^4$	$TN_{RGT_i, D_j} = P(NonE_{RGT_i} / NonE_{D_j})$	像素在 $RGT_i$ 中判为非边缘,在 $D_j$ 中也判为非边缘
检测错误的边缘子集		
假阳性 边缘集 概率	$FP_{RGT_i, D_j} = P(E_{RGT_i} / NonE_{D_j})$	像素在 $RGT_i$ 中判为边缘,但在 $D_j$ 中判为非边缘
假阴性 边缘集 概率	$FN_{RGT_i, D_j} = P(NonE_{RGT_i} / E_{D_j})$	像素在 $RGT_i$ 中判为非边缘,但在 $D_j$ 中判为边缘

$$\overline{TP}_{RGT_i} = \frac{1}{n} \sum_{j=1}^n TP_{RGT_i, D_j} = \frac{1}{n} \sum_{j=1}^n \left( \frac{1}{K \cdot L} \sum_{k=1}^K \sum_{l=1}^L RGT_{i,1} \cap D_{j,1} \right) \quad (2)$$

$$\overline{TN}_{RGT_i} = \frac{1}{n} \sum_{j=1}^n TN_{RGT_i, D_j} = \frac{1}{n} \sum_{j=1}^n \left( \frac{1}{K \cdot L} \sum_{k=1}^K \sum_{l=1}^L RGT_{i,0} \cap D_{j,0} \right) \quad (3)$$

$$\overline{FP}_{RGT_i} = \frac{1}{n} \sum_{j=1}^n FP_{RGT_i, D_j} = \frac{1}{n} \sum_{j=1}^n \left( \frac{1}{K \cdot L} \sum_{k=1}^K \sum_{l=1}^L RGT_{i,1} \cap D_{j,0} \right) \quad (4)$$

$$\overline{FN}_{RGT_i} = \frac{1}{n} \sum_{j=1}^n FN_{RGT_i, D_j} = \frac{1}{n} \sum_{j=1}^n \left( \frac{1}{K \cdot L} \sum_{k=1}^K \sum_{l=1}^L RGT_{i,0} \cap D_{j,1} \right) \quad (5)$$

在式(2)~式(5)中,  $RGT_{i,1}$  和  $RGT_{i,0}$  表示在每个分级基准图( $RGT_i$ )中的边缘像素和非边缘像素;  $D_{j,1}$  和  $D_{j,0}$  ( $0 \leq i \leq n$ ) 表示在单个参数组合的检测结果( $D_j$ )中的边缘像素和非边缘像素。这样的二分类总体上根据  $RGT_i$  中的边缘和非边缘在单算法检测结果中对应的边缘和非边缘集,可以计算出单算法的正确检测比率、灵敏度比率和特异度比率。

$$\text{正确检测比率 } Percent = \overline{TP}_{RGT_i} + \overline{TN}_{RGT_i} \quad (6)$$

$$\text{灵敏度比率 } TPR_{RGT_i} = \frac{\overline{TP}_{RGT_i}}{P} \quad (7)$$

其中,  $P = \overline{TP}_{RGT_i} + \overline{FN}_{RGT_i}$ 。

$$\text{特异度比率 } 1 - FPR_{RGT_i} = 1 - \frac{\overline{FP}_{RGT_i}}{1 - P} \quad (8)$$

其中,  $P = \overline{TP}_{RGT_i} + \overline{FN}_{RGT_i}$ 。

由此,得到鲁棒参数  $S_b$ , 该参数对应的  $RGT_i$  得到的检测结果最好。

$$\forall i, 1 \leq i \leq n, S_b = \text{Max}(TPR_{RGT_i} \cdot FPR_{RGT_i}) \quad (9)$$

将由上述算法得到的最佳结果作为预选基准图。

### 1.3 基准图的生成

边缘同图像的一阶导数局部峰值有关,梯度是函数变化的一种度量,但图像梯度值较大的地方,还不足以确定为边缘。为统计出整图中确切的边缘,需要细化幅值图像中的峰值。在大规模实验的基础上,对于初始阶段得到的  $N$  幅预选基准图,在融合决策时采用按权值关联。预选基准图  $k$  在位置  $(i,j)$  值为  $f_k(i,j)$ , 权重  $\omega_k$  计算如式(10)所示。

$$\omega_k = \frac{\sum_{i=0}^m \sum_{j=0}^n f_k(i, j)}{\sum_{k=1}^N \sum_{i=0}^m \sum_{j=0}^n f_k(i, j)} \quad (10)$$

则预选基准图加权关联结果

$$F(i, j) = \sum_{i=1}^N (\omega_i \cdot f_i(i, j)) \quad (11)$$

用初始算法计算出来的预选基准图边缘线较粗，对边缘的定位不确切。为此，需使用边缘置信度技术来实现对预选基准图的进一步融合，利用处理对象之间的相互关系与相容条件，逐步对边缘点做出准确的定位。具体步骤是：

(1)由预选基准图加权关联后的数据计算可能边缘点的初始置信度，迭代系数  $k=0$ ，初始置信度  $T^{k=0}(e)$  取加权关联后的值除以 255；

(2)在置信度  $T$  的计算中，如果像素点  $e$  的邻域点中至少有 2 个点的置信度较高(如大于 0.6)且与  $e$  成直线，迭代后置置信度将增加，取  $T^k(e) = \text{Min}(1, T^{k-1}(e) + \delta)$ ，执行步骤 4，否则执行步骤 3；

(3)如果像素点  $e$  的邻域点中至少有 7 个置信度均较低(如小于 0.3)，则迭代后置置信度将减小，取  $T^k(e) = \text{Max}(0, T^{k-1}(e) - \delta)$ ，否则置信度不变；

(4)如果  $T^k(e) = 0$  或 1，执行步骤 5，否则  $k=k+1$ ，执行步骤(2)；

(5)取置信度为 1 的作为最终边缘点，组成评价基准图(其中， $\delta$  是常数，取 0.1~0.3； $k$  是迭代次数)。图 5 是基准图的制备流程。

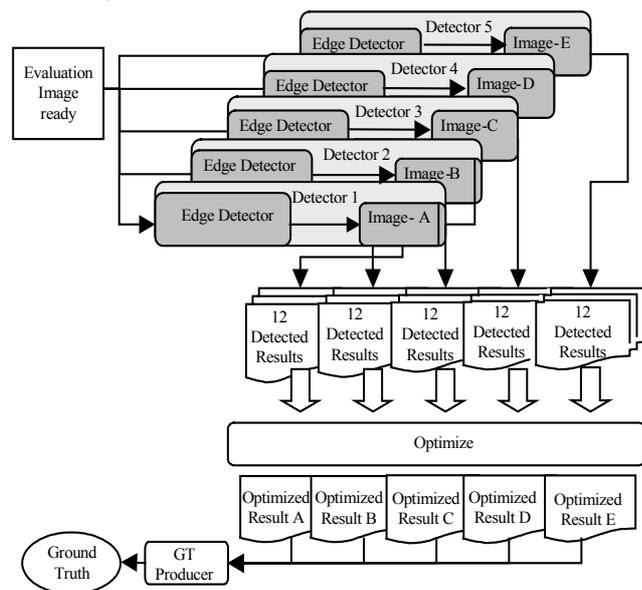
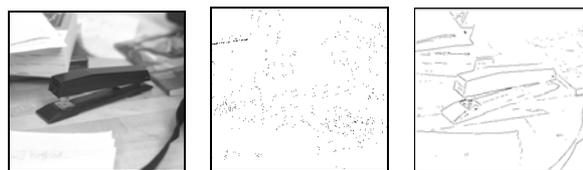


图 5 基准图制备流程

## 2 基准图生成实验结果

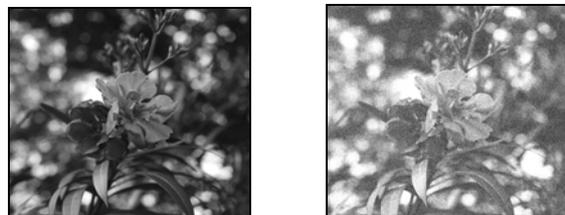
本文通过 VC++ 编程实现性能评价基准图像的生成的流程。测试图像选用一幅人造物体景物图 stapler (图 6(a))和一幅植物图 flower(图 7(a))所示。

图 6(b)是对无噪声图像采用 Canny 算子进行边缘检测的结果，图 4(c)是本文生成基准图结果，根据置信度得到的最终基准图清晰地描述场景中主要物体，图 7 是对含噪声图像的处理。对原图像加入随机噪声之后(如图 7(b))，Canny 算子检测的边缘严重地受噪声干扰(如图 7(c))很难辨认，而新的基准图能够抑制大量噪声，提取出清晰的边缘(如图 7(d))，对基准图作用的发挥很有利。

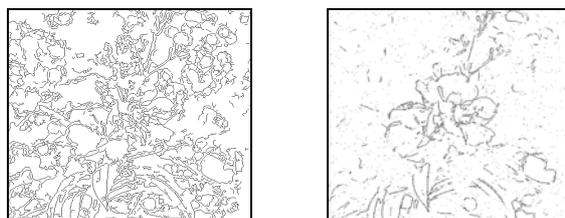


(a)原图 stapler (b)Canny 检测结果 (c)本文基准图结果

图 6 无噪声图像及其检测



(a)原图 flower (b)原图添加噪声



(c)Canny 检测结果 (d)本文基准图结果

图 7 有噪声图像及其检测

## 3 结论

基于图像融合技术和对边缘处理性能分析的结合，本文提出了新的基准图融合方法，该方法通过分析真实场景下建立的单算法预选基准图，进行基于置信度的图像融合得到可用于边缘检测算法性能评价的基准图。该基准图可信度高、抗噪鲁棒性好，能有效地解决抑制噪声和保留精细边缘之间的矛盾，形成较完善的基准图生成过程。此外，图像融合技术也是边缘检测算法发展的一个新的思路。

### 参考文献

- 1 Heath M, Sarkar S, Sanocki T. et al. A Robust Visual Method for Assessing the Relative Performance of Edge Detection Algorithms[J]. IEEE Trans. on Pattern Analysis and Machine Intelligence, 1977, 19(12): 1338-1359.
- 2 Abdou I E, Pratt W K. Quantitative Design and Evaluation of Enhancement/Thresholding Edge Detectors[C]//Proceedings of the IEEE, 1979, 67(5): 753-763.
- 3 Bryant D J, Bouldin D W. Evaluation of Edge Operators Using Relative and Absolute Grading[C]//Proc. of IEEE Computer Society Conference on Pattern Recognition and Image Processing, Chicago, 1979: 138-145.
- 4 Canny J. A Computational Approach to Edge Detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986, 8(6): 679-698.
- 5 Nalwa V S, Binford T O. On Detecting Edges[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986, 8(6): 699-714.
- 6 Iverson L A, Zucker S W. Logical/linear Operators for Intelligence, 1995, 17(10): 982-996.
- 7 Fredrik B. Edge Focusing[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1987, 9(6): 726-741.
- 8 Rothwell C A, Mundy J L, Ho\_man W, et al. Driving Vision by Topology[C]//Proc. of International Symposium on Computer Vision. Coral Gables, Florida. 1995: 395-400.