

极坐标下基于差分统计的描述器算法

徐小明 杨丹 张小洪 周小龙

(重庆大学软件学院, 重庆 400040)

摘要 特征描述器在许多机器视觉任务中有着重要的应用。为了提高特征描述器的性能,在极坐标下设计出了一种新的特征描述器——小波差分统计描述器(harr wavelet differential histograms, HWDH)。该描述器算法先把选定的特征区域分成9个小区域;接着分别对小区域进行小波变换;然后算出对应的方向差分,同时用改进的尺度适应性二阶矩阵来选择特征点的主方向;最后在极坐标下,对每个小区域的差分,在相应的方向条目内进行统计。与其他基于统计的描述器相比,该算法不需要对特征区域进行旋转,即可减少统计误差。经与尺度不变特征变换(SIFT)的描述器的实验比较表明,该描述器在鲁棒性和速度性能上都很好。

关键词 尺度不变特征变换 主方向 描述器

中图法分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2009)05-961-06

The Algorithm of Descriptor Based on Differential Histograms Under Polar Coordinates

XU Xiao-ming, YANG Dan, ZHANG Xiao-hong, ZHOU Xiao-long

(School of Software Engineering, Chongqing University, Chongqing 400040)

Abstract Descriptor of features point is vital application of many computer vision tasks. Therefore, this paper presents a novel algorithm to design the descriptor of image feature points based on differential histograms under polar coordinates (HWDH). Firstly, the neighbor area of feature point is divided into nine small patches. Then, Gradient differential is produced by harr wavelet on every small patch and a dominant direction of the feature point is computed by Scale Adapted Harris detector which is improved. Lastly, histogram is generated from the differential statistic. In contrast to other descriptor based on differential histograms, the description generated by our algorithm needn't to rotate in the around area of feature point. Therefore, it can deduce the error from statistic. Moreover, the comparative experiments illustrated that the proposed algorithm is more rapid and accurate than SIFT.

Keywords scale-invariant feature transform(SIFT), dominant directions, descriptor

1 引言

基于特征点的特征提取在机器视觉、图像配准等领域有着广泛的应用,其中一个重要的课题就是为特征点设计一个紧凑的、区分度高的描述器,另外,从特征匹配的性能要求上也可以看到,描述器性

能的好坏,很大程度上决定该技术的优劣。

目前,已经提出了许多描述器设计算法, Mikolajczyk 和 Schmid 等人还对各种描述器的性能进行了详细的分析^[1]。比如:形状纹理^[2]、复系数滤波器^[3]、尺度不变特征变换(SIFT)^[4]、PCA-SIFT^[5]描述器等;并给出了以下结论:SIFT是目前最具鲁棒性的描述器算法。为此,实验中,算法实验只与

基金项目:国家自然科学基金项目(60604007);重庆市自然科学基金项目(CSTC2005BA2002)

收稿日期:2007-05-30;改回日期:2007-11-10

第一作者简介:徐小明(1983~),男,重庆大学软件学院硕士研究生。主要研究方向为数字图像处理、计算机视觉等。

E-mail: xxm9136@sohu.com

SIFT 描述器进行比较。SIFT 描述器主要由以下两部分组成:(1)选择主方向,使得该描述器对图像的旋转具有不变性;(2)先将特征区域分成 16 个区域,然后对每个区域在 8 个方向条目上进行梯度统计,最后生成一个 128 维向量,为消除光照的影响,可将向量归一化,并将归一化结果作为该特征点的特征描述器。

从上面的步骤可以看到,由于 SIFT 描述器算法设计的复杂性高,其生成的描述器维数太大,从而导致匹配速度较慢;同时该算法必须先选择一个主方向,然后每个区域还要旋转到这个绝对坐标下才能进行匹配,这对于离散的整数图像来说,必然会带来较大的误差;加上该算法是对图像的梯度进行统计,故对纹理图像的匹配效果不好。对于这些不足,现在出现了一些改进算法,例如, Mortensen, Deng, Shapiro 等人提出了 SIFT + 全局纹理的描述器^[6],虽然该描述器提升了 SIFT 对纹理图像的处理能力,但却增加了算法的复杂性。Ke 和 Sukthankar 等人提出了 PCA-SIFT^[5]以及 Grabner 等人提出了 Fast Approximated SIFT^[7]描述器。这两个描述器虽都提升了 SIFT 描述器的速度,但却牺牲了算法的鲁棒性。针对以上的问题,本文提出了一种新的描述器设计算法——极坐标下基于差分统计的描述器设计算法。该算法不仅提升了 SIFT 描述器的鲁棒性,同时降低了匹配速度。通过实验证明,该算法在鲁棒性和速度上都优于 SIFT 描述器。

2 基于极坐标的差分统计描述器

由于特征点的选择会对匹配结果产生影响,所以为保证实验结果的正确性,本文算法采用 DOG (difference of Gaussians)^[8]来提取特征点,以便使得输入上与 SIFT 保持一致。如图 1 所示,该方法首先建立 DOG 尺度空间,其生成方法为,对于初始的图

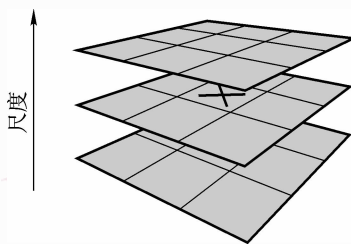


图 1 DOG 点的检测示意图

Fig. 1 The sketch of DOG detection

片是通过连续的高斯平滑来对相邻的图像进行相减;首先得到 DOG 尺度空间;接着找出尺度极值点;最后通过精确定位来获得稳定的特征点。

2.1 极坐标下基于差分统计描述器的原理

在之前的大多数描述器算法当中,都是对图像的梯度进行分析,根据视觉理论可知,人眼对梯度信息非常敏感,因而基于梯度信息设计的描述器模拟了视觉功能,该理论在实验中取得了很好的效果,但有时,梯度信息不能够完全表现物体的信息,例如,一条垂直的黑线与一条水平的黑线,从梯度上来说,两者是没有区别的。另外,对于一个好的描述器,它应具有可识别性与可区分性两个特性,即在图像发生变换时,其相同的点能够很好地被识别,同时不同点也能区分开。这就要求能充分利用特征区域的信息。

基于以上原因,本文设计了一种基于差分信息的描述器。该方法对梯度信息进行了以下两点改进:首先,通过用差分代替梯度来提高区域信息的区分度,例如前面提到的一条水平黑线与一条垂直黑线,在差分描述下,两者显然具有很大的区分性;其次,通过极坐标的形式来减少描述器设计时的误差,由于极坐标可以考虑成方向上的 1 维统计,因而不需要旋转角度来获得方向不变性。

2.2 HWDH 算法的设计步骤

HWDH (harr wavelet differential histograms) 描述器的设计主要分为以下 4 步:

(1)以特征点为中心来选择特征描述区域,该步中,不同于其他的描述器对区域的选择,HWDH 描述器的特征区域与尺度有关,因为不同尺度下的特征点,其邻域内的信息总量是与该点的尺度有关的,根据实验可知,对 DOG 点的输入,该点最好的尺度大小为 8 倍于特征尺度。通过把特征区域与特征点的尺度联系上,就可以保证同一特征点的特征区域所包含的信息量是一致的。选择特征区域后,再将该区域按半径方向分成两部分(如图 2 所示)。区域里面的部分,其大小为 3 倍于特征点尺度,外面的部分 5 倍于特征点尺度。

(2)选择特征点的主方向,本文对于主方向的选择采用了新的设计方法。主要是利用了改进的尺度适应性二阶矩矩阵^[8]。Mikolajczyk 等人提出用尺度适应性二阶矩矩阵来检测尺度不变角点(如式(1)所示):

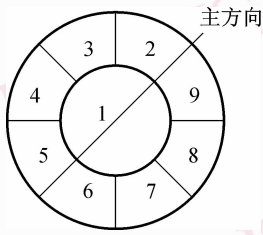


图2 对特征区域的划分示意图

Fig. 2 The sketch of dividing neighbor region of feature point

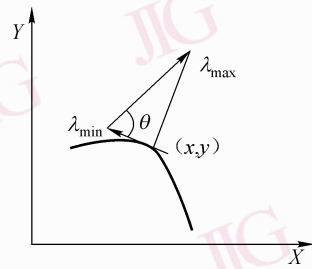


图3 特征点主方向的计算示意图

Fig. 3 The sketch of computing dominator direction of feature points

$$\mu(x, \sigma_1, \sigma_D) = \begin{bmatrix} \mu_{1,1} & \mu_{1,2} \\ \mu_{2,1} & \mu_{2,2} \end{bmatrix} = \sigma_D^2 g(\sigma_1) * \begin{bmatrix} L_x^2(x, \sigma_D) & L_x L_y(x, \sigma_D) \\ L_x L_y(x, \sigma_D) & L_y^2(x, \sigma_D) \end{bmatrix} \quad (1)$$

式中, σ_1 是积分尺度, σ_D 是差分尺度, 这里 σ_1 由 σ_D 来决定, L_a 是在方向 a 上的微分, a 代表 x, y 方向。该公式描述了尺度为 σ_D 的图像在某一点邻域的梯度分布。其中矩阵 μ 的两个特征值描述了该点邻域的两个主要的方向变化, 基于这个思想, 本文算法对其进行了改进, 设计出了新的主方向选择方法。该方法具体地实现如下: 如图 2 所示, 用 harr 小波对区域 1 进行变换, 首先计算出 x, y 方向的微分 dx, dy, dxy ; 然后求出该区域的 $\sum dx, \sum dy$ 以及 $\sum dxy$, 即可生成以下特征矩阵 μ :

$$\mu(x, \sigma_D) = \begin{bmatrix} \mu_{1,1} & \mu_{1,2} \\ \mu_{2,1} & \mu_{2,2} \end{bmatrix} = \begin{bmatrix} \sum d_x^2(x, \sigma_D) & \sum d_x d_y(x, \sigma_D) \\ \sum d_x d_y(x, \sigma_D) & \sum d_y^2(x, \sigma_D) \end{bmatrix}$$

这里与式(1)不同的是, 该方法是直接对区域进行累加, 而不是用高斯卷积。累加区域的大小与特征点的尺度有关, 本文累加区域选择 3 倍的尺度大小。

矩阵 μ 求出以后, 即可解出 $\lambda_{\max}, \lambda_{\min}$, 其分别为该矩阵的最大、最小特征值 (如图 3 所示), 角度计算如下: $\theta = \text{ant}(\lambda_{\max}/(\lambda_{\min} + 0.5))$, 将其归一化到 $[-\pi, \pi]$, 则结果就是该特征点的主方向。此处分母加上微小数 (取 0.5) 是为了防止 λ_{\min} 为零。综上所述, 该方法不仅计算简单、计算量小, 同时从后面步骤也可以看到, 此处计算的方向差分在后面会用到, 也就省去了多余的计算量。

(3) 首先将外围的区域按方向角度分成 8 份, 其中, 开始点是主方向的位置 (如图 2 中的细实线

所示); 接着对每一个区域用 harr 小波来计算对应的 dx, dy ; 最后, 通过对整个特征区域按划分好的区域条目进行统计来计算出每个条目内的 dx, dy 总和。其中, 对于图 2 中的区域 1, 本文算法把它单独作为一个条目, 同时, 该步不需要对它计算 dx, dy , 因为在第 2 步中, 已经对其进行了计算。该步主要是生成一个紧凑, 区分能力高的描述器, 从这步可以看到, 为了使描述器的鲁棒性更强, 该算法在极坐标下进行统计, 这样做的主要原因是该描述器不需要旋转图像的物理位置, 只需要以主方向为起点划分区域就行, 由于其避免了旋转带来的误差, 从而可达到提高描述器鲁棒性的目的。

(4) 利用前面计算的结果, 就可以生成 HWDH 描述器。由此可很直观地看到, 该描述器只有 $9 \times 4 = 36$ 维, 这意味着该描述器的匹配速度较快, 通过后面的实验也可得出, HWDH 描述器的匹配速度是 SIFT 描述器的 4 倍以上。同时为了尽量减低光照的影响, 该描述器需要进行归一化, 其最终结果可作为该特征点的描述器算子。

该描述器有以下 4 个优点:

(1) 它不需要额外计算主方向, 而是在设计描述器的过程中, 即可对其进行设计;

(2) 该描述器设计不需要对特征区域进行旋转。由于其可避免旋转带来的像素误差, 因而可使得运算的精度得到提高;

(3) 该描述器只是一个 36 维的向量, 从而可使得匹配的速度很快;

(4) 由于该算法是对图像的方向差分进行统计, 从而对于纹理图像也有很好的匹配结果; 通过实验可见, 该描述器对图像的旋转、噪声、光照、尺度等变化具有很强的鲁棒性。同时在匹配速度上提高了 4 倍左右。

3 实验与讨论

3.1 特征点对的匹配准则和评估指标

判断某两个特征点是否相似,主要有以下 3 个指标:

- (1) 基于欧几里德距离的阈值;
- (2) 邻接点距离;
- (3) 第 1 邻接点和第 2 邻接点的比值。本文实验采用的匹配准则是第 3 个,因为用第 1 邻接点和第 2 邻接点的比值更能反映描述器的区分能力。

查全率 (Recall) 和查准率 (Precision) 在信息检索中是一个非常重要的评估指标,它们的定义如下:

$$R_{\text{recall}} = N_1 / N$$

$$R_{\text{precision}} = n_1 / n$$

其中, R_{recall} 为查全率, $R_{\text{precision}}$ 为查准率, N_1 为匹配点对数, N 为数据库中对应的匹配点对数, n_1 为正确命中数, n 为总命中数

不过查全率和查准率很难同时达到最佳,一般查全率高,查准率就低,反之,查准率高,查全率就低,实际应用中常把查全率,查错率 ($1 - R_{\text{precision}}$) 作为衡量指标。在实验中,要做的是验证基于 HWDH 的描述器的有效性,也就是在保证匹配点对多的同

时还要尽可能地保证正确率。

3.2 实验结果

本文选取了 4 组不同的图像,用来对该算法和 SIFT 描述器进行评估。其中 HWDH 描述器是 36 维, SIFT 描述器采取的还是通用的 128 维。图 4 显示了选择的图像集。它们分别代表了不同的图像变化:图 4(a) 为光照变化的图像,图 4(b) 为受噪声影响的图像,图 4(c) 为视角变化的图像,图 4(d) 为尺度和方向变化的图像。下面对这 4 组图像分别进行实验和讨论。

图 5 是在不同阈值下,图 4 图像集两种算法的匹配结果:图 4(a) 右图是图 4(a) 左图的比较结果,图像间发生了光照变化;图 4(b) 右图是图 4(b) 左图的比较结果,图 4(b) 右图是在图 4(b) 左图上增加了噪声;图 4(c) 右图是图 4(c) 左图的比较结果,图像发生了视角变化;图 4(d) 右图是图 4(d) 左图的比较结果,图像间发生了尺度和方向变化。从图 5(a)、图 5(b) ~ 图 5(d) 中可以看出, HWDH 描述器对于旋转、噪声、光照、尺度变化图像的匹配结果优于 SIFT 描述器。这主要归功于 HWDH 描述器紧凑的设计:在对特征区域进行统计的时候, HWDH 描述器在极坐标下,由于使用了更巧妙的方法使得不需对图像进行物理上的旋转,从而提高了统计的



图 4 不同变换下的图像对集

Fig. 4 Imageset including every transformation

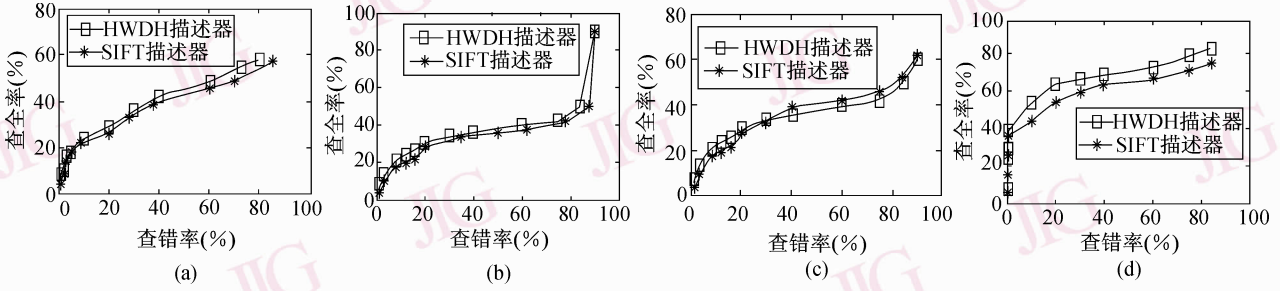
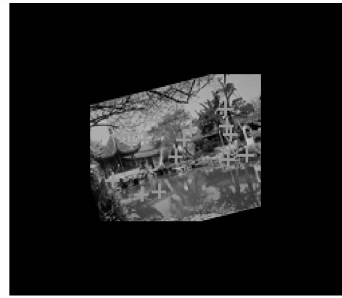
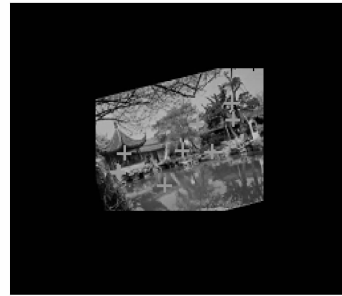


图 5 在不同阈值下,对于图 4 中的图像集两种算法的匹配结果

Fig. 5 The experiment result under different thresholds was produced by testing imageset which was shown in Fig. 2 using two algorithms which was present



(a) HWDH描述器的匹配结果(邻接比取0.4,获得16个点对,匹配正确率为100%)



(b) SIFT描述器的匹配结果(邻接比取0.4获得8个点对,匹配正确率为100%)



(c) SIFT描述器的匹配结果(邻接比取0.52,获得16个点对,匹配正确率为75%)

图 6 SIFT 描述器与 HWDH 描述器的匹配结果比较

Fig. 6 HWDH descriptor was compared with SIFT descriptor

精度。从图 5(c)中可以看到,在视角变化的时候,HWDH 描述器表现并不总是优于 SIFT 描述器,在查错率较大的时候,HWDH 描述器的查准率要低于 SIFT 描述器,这是因为梯度方向对视角比较敏感所致。总的说来,HWDH 描述器在总体性能上要优于 SIFT 描述器。

图 6 通过综合各种影响来对 HWDH 描述器和 SIFT 描述器的匹配效果进行评估。在这一组图像中,物体发生了方向、尺度、亮度以及小视角上的变化,实验结果表明,在限定同样对应匹配点时(选择 16 对匹配点),HWDH 描述器的匹配准确率优于 SIFT 描述器。同样,在限定同样的邻接点比值时(实验中取 0.4),HWDH 描述器找出的匹配点数明显多于 SIFT 描述器。

表 1 给出了 HWDH 描述器与 SIFT 描述器的生成速度和匹配速度的比较,实验中,选取大约 3 000 个特征点。从表中可以看到,尽管两者的计算时间非常接近,但是,在匹配速度上,由于 HWDH 描述器的维数很小,所以速度上要明显快于 SIFT 描述器。

表 1 HWDH 描述器与 SIFT 描述器的运行速度比较

Tab. 1 Comparison of time spent between HWDH and SIFT descriptor

算法	描述器生成时间(s)	匹配时间(s)
HWDH 描述器	1.6	0.6
SIFT 描述器	1.7	2.9

4 结 论

本文在极坐标下通过对特征点邻域内的方向差

分进行分析,设计了一种新的描述器算法——HWDH。通过与 SIFT 描述器的比较可以得出,本文提出的描述器不仅具有很强的鲁棒性,而且匹配速度也很快。不过,与 SIFT 描述器类似,本算法在比较大的视角变化下,其匹配效果不佳,所以对于它的视角处理能力将是需要研究和完善的。

参考文献 (References)

- 1 Mikolajczyk K, Schmid C. A performance evaluation of local descriptors [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(10): 1615-1630.
- 2 Belongie S, Malik J, Puzicha J. Shape matching and object recognition using shape contexts[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(4): 509-522.
- 3 Schaffalitzky F, Zisserman A. Multi-view matching for unordered image sets[A]. In: Proceedings of the 7th European Conference on Computer Vision[C], Copenhagen, Denmark; 2002:414-431.
- 4 Lowe D G. Distinctive image features from local scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 2(60): 91-110.
- 5 Ke Y, Sukthankar R. PCA-SIFT: a more distinctive representation for local image descriptors[A]. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition[C], Washington, D C, USA, 2004. 511-517.
- 6 Mortensen E N, Deng H, Shapiro L. A SIFT descriptor with global context[A]. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition [C], San Diego, CA, USA, 2005: 184-190.
- 7 Grabber M, Grabber H, Bischof H. Fast approximated SIFT[A]. In: Proceedings of Asian Conference on Computer Vision [C], Hyderabad, India, 2006: 918-927.
- 8 Mikolajczyk K, Schmid C. Scale and Affine Invariant Interest Point Detectors[J]. International Journal of Computer Vision, 2004, 60(1):63-86.