

# SURF：快速的鲁棒的特征提取

翻译：picone

摘要：本文提出了一种新的尺度不变的、旋转不变的特征点探测器和描述子，这种算法叫 SURF (Speeded Up Robust Features)。它可以近似甚至超过过去提出的重复性、区别性、鲁棒性的方案，但是现在可以计算得更快速。

这是通过图像卷积的积分图像实现的，建立在现有的探测器和描述子的优点上（在摄像机自动调定中，使用 Hessian 矩阵来度量检测器和描述子），并且简化了这些方法为实现本质。这些结合起来出现了一种新的检测、描述和匹配步骤。本文通过标准的评估集合和真实生活中应用到目标识别的场景的图像进行实现，证明 SURF 有更好的性能。

## 1 简介

寻找两个图像相同的场景或目标是很多计算机视觉应用的一部分，如相机校准、3D 重建、图像登记、目标识别等等。查找不同图像里面一致的地方，可以分为以下 3 个步骤：首先，在不同的地方选择“感兴趣点”，如边角、点和 T 岔口。那些最有价值属性的感兴趣点具有可重复性，即即使在不同的图像展示环境下也能找到相同感兴趣点。其次，在每个感兴趣点的领域里使用准确的特征向量来表示描述子，描述子是独特的，并能够应对噪声有鲁棒性、检测错误、光照和几何形变。最后，使用描述子特征来匹配不同的图像，匹配通常计算描述子特征的距离，如马氏距离 (Mahalanobis distance) 和欧氏距离 (Euclidean distance)。描述子的尺寸直接影响到匹配所花费的时间，所以使用更低维度的描述子。

我们已经开发出检测器和描述子，在不牺牲性能的前提下，比现在最先进的方法有更快的速度。为了实现这个算法，需要平衡以上的要求，如降低描述子的维度和复杂性的同时保持独特性。

目前已经有很多种类的检测器和描述子相关的文章，例如[1-6]。同时，也有一些细节对比和分数评估的数据集，例如[7-9]。在构建这个快速的检测器和描述子时，我们从以前已经完成的工作中获取灵感，并找到对效果有利的方面。在我们对评估图像集和实际目标识别应用的实验中，我们的算法不仅仅更快，而且更有区别性，并有同样的可重复性。

在我们查找特征时，首要问题是不变性。明显地，当我们改变拍摄条件时，

可能会发生几何形变和灯光变化。本文我们致力于尺度不变的和旋转不变的检测器和描述子。这里提供一个很好的当发生形变时在特征复杂度和鲁棒性之间的折中方案。并假设目标被二次倾斜、各向异性缩放和透视变换，这一定程度上提高描述子的鲁棒性。在 Lowe 发表的文章[2]中，加上复杂的仿射不变性特征对会降低鲁棒性，并且效果并不是很好，除非图像角度有很大的转变。在某些情况下，旋转不变性可以被忽略，仅使用尺度不变性的描述子，可以使用“upright SURF”（U-SURF）。事实上，在一小部分的应用上，如移动引导机器人、可视化导游，它们的摄像头通常只是水平的角度旋转。去掉旋转不变性的特性后，不仅能提升速度，还能提升识别能力。对于光线变化，我们设想了一种用缩放因子和偏移因子的简单线性模型。注意，我们的检测器和描述子没有使用到颜色。

本文组织结构如下：第 2 节介绍相关的工作，以及我们的成果是怎么研究的；第 3 节介绍兴趣点检测；第 4 节介绍新的描述子；最后，第 5 节展示实验结果以及第 6 节的结论。

## 2 相关工作

兴趣点检测。最广泛使用的检测器可追溯到 1988 年的基于二阶矩阵特征值的 Harris 角点检测器[10]。但是，Harris 角点检测不是尺度不变的。Lindeberg 介绍了一种尺度自动选择的方法[1]，使得检测图片里的特征点成为了可能。他实验比较了 Hessian 矩阵判别式和 Laplacian（一种跟 Hessian 矩阵一致的方法）提取点特征，Mikolajczyk 和 Schmid 从中提炼出一种鲁棒的、尺度不变的特征检测器 Harris-Laplace 和 Hessian-Laplace[11]，并具有很高的可重复性。他们使用了尺度适应的 Harris 度量或 Hessian 矩阵判别式来选择兴趣点位置，并使用 Laplacian 来选择尺度。为了提升速度，Lowe[12]使用 Difference of Gaussians（DoG）滤波器来近似 Laplacian of Gaussian（LoG）滤波器。

还有一些其他的特征点检测器，例如 Kadir and Brady 的突出区域检测[13]能够最大限度地增大区域内的熵，和 Jurie 的基于边缘的区域检测器[14]，都能够少量地加速特征检测。同时，有一些仿射不变的特征检测器使得能够支持更大的角度改变。但是这些内容超出本文的讨论范围。

通过学习已有的检测器和比较[8]和[15]文章，我们概括出：（1）基于 Hessian 的检测器比基于 Harris 的更稳定和更有可重复性。使用 Hessian 矩阵判别式而不

是痕迹 (Laplacian) 似乎更有利, 因为它更少地涉及细长的、局部不正确的结构。

(2) 使用 DoG 来近似能损失很少的精度同时提升速度。

特征描述子。有很多种类的特征描述子, 如高斯导数[16]、不变矩[17]、复杂特征[18][19]、可控滤波[20]、基于相位的局部特征[21], 描述子能表现出不同的兴趣点领域不同的特征, 其中 Loew 的文章[2]有最突出的效果[7], 这是因为他的特征点描述子抓取了大量空间强度模板, 能够对变形和位置错误有很好的鲁棒性。[2]的描述子, 简称 SIFT, 计算感兴趣点周围的局部梯度方向直方图并保存到一个 128 维向量中 (8 个方向, 并且每个方向用  $4 \times 4$  个维度表示)。

已经出现各种各样改良的方法, Ke 和 Sukthankar[4]使用 PCA 梯度图像。这种 PCA-SIFT 使用 36 维的描述子, 因此匹配更快, 但是 Mikolajczyk 的研究[8]证明这种方法比 SIFT 的描述子区分度更小, 并且更慢的特征计算降低快速匹配的影响。在同一份论文[8]里, 做着提出了一种变种的 SIFT, 叫 GLOH, 它的描述子在相同维度的情况下有更大的辨识度, 但是 GLOH 计算量较大。

SIFT 描述子看上去似乎更有吸引力, 因此目前还是使用最广泛的描述子, 它辨识度较大, 且相对较快, 这点对在线的应用至关重要。最近, Se 等人在 Field Programmable Gate Array (FPGA) 上实现了 SIFT, 并速度提高了一个数量级。但无论如何, 高纬度的描述子在匹配时成为了 SIFT 的缺点。对于普通 PC 上的在线应用, 这三步 (检测、描述、匹配) 里的每一步都应该做的更快。Lowe 提出使用 best-bin-first 来替代, 使得速度匹配时候速度更快但降低了准确性。

我们的方法。本文提出了一种新的检测器和描述器, 创造了 SURF (Speeded-Up Robust Features)。检测器基于 Hessian 矩阵[1][11], 但是使用了一种很基础的接近方法: DoG, 一种非常基础的基础 Laplacian 检测器。它使用积分图像来减少计算时间因此叫它“Fast-Hessian”检测器。另一方面, 描述子 Haar 小波响应感兴趣点领域范围内分布的特征。再次强调, 我们使用积分图像来加速。除此之外, 仅使用 64 维的描述子, 减少特征计算和匹配的时间, 与此同时增加鲁棒性。我们同时使用基于 Laplacian 的标志作为索引, 这不仅提高了匹配速度, 还增强了鲁棒性。

为了使论文更具有原创性, 我们简单地介绍了积分图像[23]的概念。他们能够使用盒子滤波器更快地实现。积分图像  $I_{\Sigma}(\mathbf{x})$  在某点  $\mathbf{x} = (x, y)$  表示  $I$  图像内零

点和点  $\mathbf{x}$  所形成的矩形区域内所有像素总和， $I_{\Sigma}(\mathbf{x}) = \sum_{i=0}^{i \leq x} \sum_{j=0}^{j \leq y} I(i, j)$ 。  $I_{\Sigma}$  累加只需要通过计算规矩的独立的矩形区域内 4 个加法运算来计算卷积结果。

### 3 快速 Hessian 检测器

我们的检测器基于 Hessian 检测器，因此它的计算速度和精确性都很好。无论如何，相对于使用不同的度量来选择位置和尺度（正如 Hessian-Laplace 检测器 [11]），我们仅使用 Hessian 判别式。给定的图像  $I$  和点  $\mathbf{x} = (x, y)$ ，Hessian 矩阵  $\mathcal{H}(\mathbf{x}, \sigma)$ ， $\mathbf{x}$  在  $\sigma$  时定义如下：

$$\mathcal{H}(\mathbf{x}, \sigma) = \begin{bmatrix} L_{xx}(\mathbf{x}, \sigma) & L_{xy}(\mathbf{x}, \sigma) \\ L_{xy}(\mathbf{x}, \sigma) & L_{yy}(\mathbf{x}, \sigma) \end{bmatrix} \quad (1)$$

$L_{xx}(\mathbf{x}, \sigma)$  是高斯二阶导数  $\frac{\partial^2}{\partial x^2} g(\sigma)$  和图像  $I$  在点  $\mathbf{x}$  的卷积。同理  $L_{xy}(\mathbf{x}, \sigma)$  和  $L_{yy}(\mathbf{x}, \sigma)$  也是这样。

论文[24]表示高斯函数是尺度空间最优分析。在实践中，高斯函数需要离散和剪裁（图 1 左），并且图像重新采样时也需要高斯滤波。另外，在降采样的时候，没有新的方法支持 2D，目前仅有支持 1D 的[25]。因此，在这方面我们似乎高估了高斯函数的重要性，这里我们测试使用更简单的来替换。因为使用高斯滤波我们别无选择，考虑到 Loew 使用 LoG 近似的成功，我们更进一步地使用盒子滤波（图 1 右）来近似。这些近似高斯二阶微分的方法，能够通过积分图像非常快地计算，并且对图像大小无关。在结果的小节中将会展示使用高斯离散和剪裁和本方法表现对比。

图 1 的  $9 \times 9$  盒子滤波是近似高斯二阶微分函数， $\sigma = 1.2$ ，并且是最小的尺度。我们用  $D_{xx}$ ， $D_{xy}$ ， $D_{yy}$  来表示我们的近似，使用权重应用于矩形区域使得计算非常简单，但是我们需要更进一步平衡 Hessian 判别式里面的权重， $\frac{|L_{xy}(1.2)|_F |D_{xx}(0.9)|_F}{|L_{xx}(1.2)|_F |D_{xy}(0.9)|_F} = 0.912 \dots \cong 0.9$ ，其中  $|x|_F$  是 Frobenius 范数，因此有：

$$\det(\mathcal{H}_{approx}) = D_{xx}D_{yy} - (0.9D_{xy})^2 \quad (2)$$

此外，滤波相应结果归一化使用滤波核大小，这保证了 Frobenius 范数对于任何滤波核大小权值都是个常数。

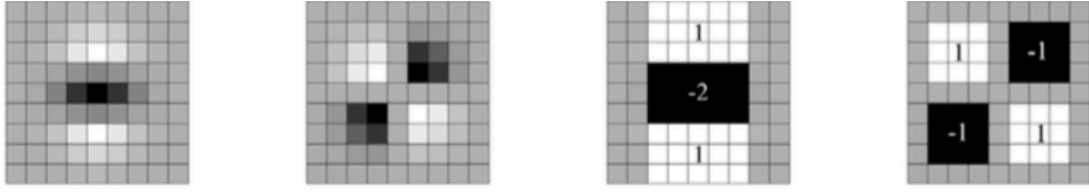


图 1 从左到右为：高斯二阶函数部分导数， $y$  方向、 $xy$  方向上，和我们近似的盒子滤波核，灰色部分是 0。

构建尺度空间通常使用图像金字塔。构建更高层次的金字塔过程中不断重采样并使用高斯平滑。由于我们使用盒子滤波核积分图像，我们不需要迭代使用相同的滤波核并使用上一层输出的结果作为下一层的滤波图像，而是几个不同大小的滤波核应用到完全一样的原始图像中，甚至可以并行执行（不过本文后面没应用）。因此，分析尺度空间通过不断地扩大滤波核大小而不是降采样缩小图像大小。上述 $9 \times 9$ 滤波核被认为是最初始的尺度层，我们把他的的尺度大小 $s = 1.2$ （对应高斯倒数为 $\sigma = 1.2$ ）。考虑到整体图像的离散性质和特定滤波器的结构，接下来的层通过不断扩大滤波核尺寸得到。具体地说，滤波核的尺寸分别是 $9 \times 9$ ， $15 \times 15$ ， $21 \times 21$ ， $27 \times 27$ 等。在更大的尺度空间上，连续的滤波器尺寸之间尺度也相应地改变。因此，对于每个新的组，滤波核尺寸应该翻倍增长（从 6 到 12 到 24）。同时，提取兴趣点的采样间隔也增大了一倍。

当滤波核按比例缩放后，近似的高斯导数比例也随之改变。例如，我们 $27 \times 27$ 滤波器对应 $\sigma = 3 \times 1.2 = 3.6 = s$ 。此外，Frobenius 范数对我们的滤波器来说是常数，他们已经被归一化[26]。为了通过尺度空间找到局部感兴趣点，在感兴趣点附近 $3 \times 3 \times 3$ 领域内应用非极大值抑制。Brown[27]提出使用 Hessian 矩阵行列式的极大值插入到尺度图像空间中。在我们的例子中，尺度空间插值非常重要，因为每一组的第一层的尺度空间差距还是非常大。图 2 左图展示我们使用“Fast-Hessian”检测器检测感兴趣点的结果。

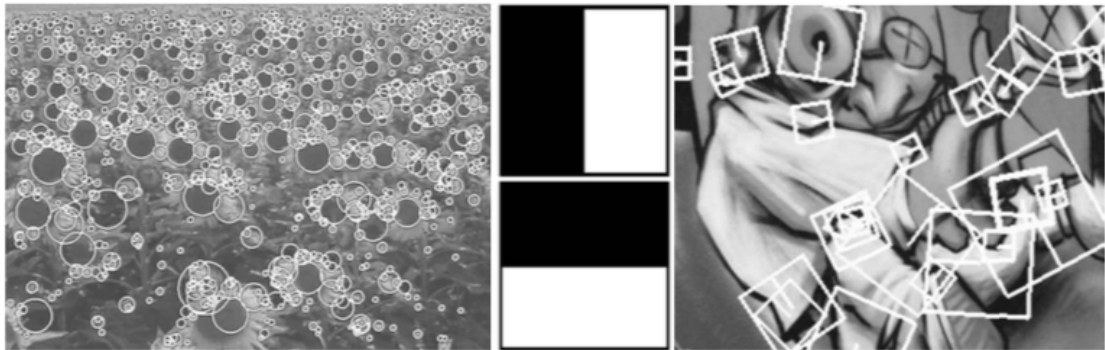


图 2 左：田里的向日葵检测到的感兴趣点，这种场景非常清晰地展示基于 Hessian 检测器的特

征。中：SURF 使用的 Haar 小波类型。右：Graffiti 场景的描述子不同尺度的窗口大小。

## 4 SURF 描述子

与其他描述子相比，SIFT 描述子有良好的表现[8]，它混合了粗糙的局部特征和分布的灰度梯度图像特征，似乎具有很好的独立性，同时避免了尺度空间上的局部错误，利用梯度图像的相对强度减少光线的影响。

因此，SURF 描述子也采用相似的属性，其复杂度进一步降低。第一步是根据感兴趣点周围圆形区域的信息确定可重复的方向。然后，我们构造一个与方向所选相一致的正方形区域，并从中提取出 SURF 描述子。后面会解释这两个步骤。此外，我们还提供了一个直接的描述子版本（U-SURF），它并不具备旋转不变性，因此它能更快地计算并应用在一些相机水平的应用上。

### 4.1 方向确定

为了保证旋转不变性，我们为每个兴趣点加上可重复的方向。为此，我们首先计算了，感兴趣点周围半径为  $6s$ （ $s$  是检测到的感兴趣点的尺度）的圆形区域  $x$  和  $y$  方向上的小波响应（如图 2 所示）。同样采样过程也依赖于尺度  $s$  的，小波响应使用当前尺度  $s$ ，因此如果尺度  $s$  很大则小波也很大。因此，我们再次使用积分图像进行快速滤波。在任何尺度上，计算  $x$  和  $y$  方向上的小波响应仅使用 6 次运算。小波的边长是  $4s$ 。

以感兴趣点为中心使用高斯加权（ $\sigma = 2.5s$ ）后小波响应代表的横坐标上的空间响应强度和纵坐标的响应强度。主方向的确定通过角度为  $\frac{\pi}{3}$  的扇形滑动窗口计算窗口内水平和垂直方向的响应强度之和，两个向量相加产生新的向量，选择向量长度最大的向量的方向作为主方向。滑动窗口的大小是个经验参数，这里的选择是经过实验的，这个值过小会导致选择不到主要的小波响应方向，这个值过大会直选择小波响应向量最长的方向，这两种情况都导致了感兴趣点的主方向不稳定。注意 U-SURF 可以跳过这一步。

### 4.2 描述子部分

对于描述子提取，第一步是构造以感兴趣点为中心的矩形区域，并用上一节的方法选择主方向（对于 U-SURF 来说这是不需要的），这矩形区域的边长为  $20s$ ，这些矩形区域如图 2 所示。

区域被均匀地分割成 $4 \times 4$ 个子区域，这些都保存着重要的空间信息。对于每个子区域，我们在 $5 \times 5$ 矩形采样点内计算一些简单的特征。为了简单，我们把水平方向上的 Haar 小波响应叫 $dx$ ，垂直方向方向上的 Haar 小波响应叫 $dy$ （滤波器尺寸为  $2s$ ）。这里的“水平”和“垂直”和所选择的感兴趣点的主方向有关。为了增加鲁棒性来对抗几何形变和局部噪声，对 $dx$ 和 $dy$ 计算之前要用 $\sigma = 3.3s$ 的高斯函数对感兴趣点附近区域进行滤波。

然后，把每个子区域的小波响应 $dx$ 和 $dy$ 之和生成一组元素特征向量。为了获取关于强度变化对立的信息，我们还提取了响应的绝对值 $|dx|$ 和 $|dy|$ 。因此每个子区域有一个 4 维的描述子向量  $\mathbf{v}$  来描述它的潜在的强度结构  $\mathbf{v} = (\sum dx, \sum dy, \sum |dx|, \sum |dy|)$ ，这使得每个描述子向量在 $4 \times 4$ 个子区域一共有 64 维。小波响应在亮度（偏移）中是不变的，把描述子转换成一个单位向量来实现不变性（一个按比例缩放元素）。

图 3 展示了 3 种不同的子区域图像强度模式和它不同的描述子属性。可以把这种局部强度模式组合起来，从而产生一个唯一的描述符。

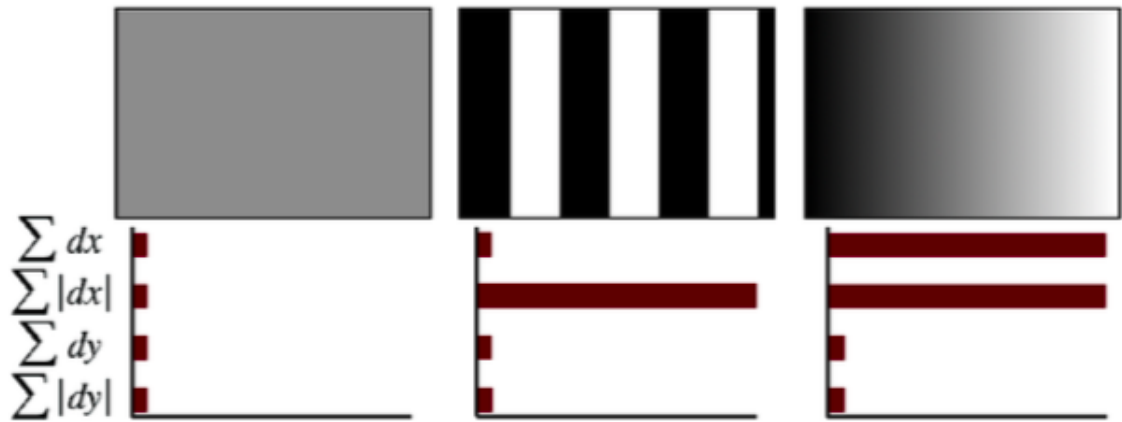


图 3 描述子范围的子区域代表自然情况下不同强度样式。左：对于均匀区域，所有值都相对低。中：对于  $x$  方向安频率强， $\sum |dx|$ 高而其他相对低。如果强度是  $x$  方向渐变的， $\sum dx$ 和 $\sum |dx|$ 都高。

为了实现 SURF 描述符，我们实验了一些小波响应特征，如 $d_x^2$ 和 $d_y^2$ 、高阶的小波响应、PCA、中位数、平均数等。通过对上述方法进行评估，使用集合的方法表现最好。然后我们改变子区域采样的数量，考虑到更细分子区域表现不那么鲁棒，并且增加很多匹配时候的耗时， $4 \times 4$ 个子区域的方法表现效果最好。另一方面，使用更短的描述子 $3 \times 3$ 个子区域(SURF-36)表现更差，但匹配速度非常快，且与本文中其他描述子相比还是可以接受的。图 4 显示了一些比较结果 (SURF-

128 将会简短地说明)。

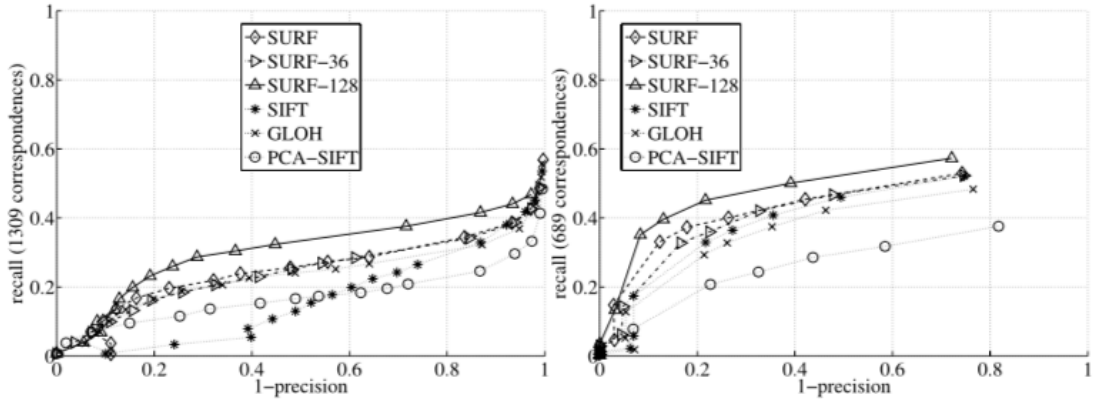


图 4 两种不同的匹配策略和不同的图像描述方法召回对比，测试使用“Graffiti”图像序列（第 1 和 3 帧），它有视觉  $30^\circ$  变化。感兴趣点的计算使用我们的“Fast Hessian”检测器的结果。注意感兴趣点并不是仿射不变的，所以结果不能和[8]的比较。SURF-128 对应扩展的描述子。左：基于相似度阈值的匹配策略。右：基于邻近比率的匹配策略。

我们还测试了另一个版本的 SURF (SURF-128)，它还添加了一些类似的特征。它的使用和之前一样也是求和，但是现在需要更进一步地分开计算这些值。计算  $dx$  和  $|dx|$  时，分开地计算  $dy < 0$  和  $dy > 0$  的情况。相似地，计算  $dy$  和  $|dy|$  依据  $dx$  分开地计算，因此有翻倍的特征。描述子会更有区分性，而且计算描述子不会慢很多，但是匹配的时候会因为高维度速度很慢。

在图 4，参数选择与 Mikolajczyk[8]认为最有挑战性的“Graffiti”场景相比，它包含了外平面旋转、内平面旋转和亮度变化。其中扩展的  $4 \times 4$  子区域 (SURF-128) 描述子表现最好。此外 SURF 性能良好，处理速度快，两种都超越了现有的先进水平。

## 5 实验结果

首先，我们给出一个标准的评估集的结果来测试检测器和描述子。然后，我们讨论实际生活中目标检测的应用。比较的检测器和描述子都是基于作者的原始实现。

基础评估。我们使用 Mikolajczyk<sup>1</sup>提供的图像序列和测试软件来测试检测器和描述子。这些图像有真实的纹理和结构场景。由于篇幅限制，我们不能展示所有的图像序列。对于检测器的比较，我们选择了两个视觉的改变 (Graffiti 和 Wall)，一个缩放、旋转 (Boat) 和亮度变化 (Leuven) 的图像序列 (见下面讨论的图 6)。描述子的评估使用除了 Bark 之外所有的图像序列 (见图 4 和图 7)。

<sup>1</sup> <http://www.robots.ox.ac.uk/%20CC%83vgg/research/affine/>



对于检测器，我们文章[9]可重复性、可描述性的分数。这代表相对于检测到少量感兴趣点的方法有多少感兴趣点被检测到（只有图像中可见的部分考虑在内）。

检测器与 Lowe[2]提出的使用 difference of Gaussian（DoG）的检测器和 Mikolajczyk[15]提出的使用 Harris-Laplace 和 Hessian-Laplace 的检测器作比较。所有检测器发现的感兴趣点数量平均值很接近，这适用于所有图像，包括使用目标检测实验的数据库，详见表格 1。我们可以看到，我们的“Fast-Hessian”检测器比 DoG 快了 3 倍，比 Hessian-Laplace 快了 5 倍。同时，我们使用 Graffiti, Leuven, Boats 甚至 Wall 来比较可重复性。Graffiti 和 Wall 图像序列中有平面外的旋转，导致放射变换，但是我们比较的检测器只有旋转不变性和尺度不变性，因此，这种形变需要鲁棒性的特征来解决。

表格 1 我们比较的检测器的阈值、感兴趣点数量及消耗时间（Graffiti 的第一帧，800×640）

检测器	阈值	感兴趣点数量	消耗时间（ms）
Fast-Hessian	600	1418	120
Hessian-Laplace	1000	1979	650
Harris-Laplace	2500	1664	1800
DoG	默认	1520	400

描述子的评估使用像[4]和[8]的有召回 1 个准确的图片。对于每个评估集合，我们使用图像序列中 1 和 4 图像，除了 Graffiti 使用 1 和 3，Wall 使用 1 和 5，分别对应视觉的 30° 和 50° 变化。在图 4 和图 7 中，我们用“Fast-Hessian”检测器检测到的感兴趣点使用 SURF 描述子比较 GLOH、SIFT 和 PAC-SIFT 描述子，几乎所有测试中 SURF 的表现最优秀。在图 4 中，我们分别使用了两种不同的比较方法，一种使用阈值相似度另一种使用最邻近比例（详见[8]对这些方法的讨论），这对描述符的分数排名有影响，但是 SURF 依然表现最好。由于篇幅限制，只有阈值相似度的方法在图 7 中展示，因为该方法更适合用来区分描述子的特征空间[8]并更广泛应用。

SURF 描述子在规律性和特征性上都表现突出，在相同的精度上有时候会超过 10%的的召回率（见表格 2）。在第 4 节中给出的精确版本（SURF-128）比常规的 SURF 效果稍好，但是对于速度要求高的应用来说它的速度就比较慢了。

在整个论文中，包括目标识别实验，我们通常使用相同的参数和阈值（见表

格 1)，时间评估基于 Linux PC（Pentium IV，3GHz）。

表格 2 测试了 Graffiti 第 1 帧图像，计算了检测和描述的总时间，阈值适应对应的图像，这些速度也代表了其他图像。

	U-SURF	SURF	SURF-128	SIFT
时间（ms）	255	354	391	1036

目标识别。我们还在实际应用中测试了新的特征，目的是识别博物馆里的艺术品。测试数据库里有 216 个图像 22 个对象。测试集的图像（116 张照片）是在各种条件，包括极端的亮度变化、通过玻璃反射的目标、角度变化、放大、不同的相机质量等等。此外，图像很小 320×240 很多图像的细节被忽略，因此更具有挑战性。

为了从数据库中识别对象，我们进行了如下操作：将测试集的图像与参考集中所有的图像进行比较，以匹配它们各自的感兴趣点。匹配到的目标是匹配感兴趣点最多的参考图像作为匹配目标。

匹配的执行方式如下：将测试图像中的感兴趣点和参考的感兴趣点进行比较，通过计算它们描述子之间的欧氏距离进行比较。最邻近比对策略[2][7][18]为，如果匹配点的距离比之前匹配的近 0.7 倍，那么这是一个新的匹配对。显然，额外进行几何匹配能减少匹配错误点的影响，这个可以在匹配器上完成。但是现在并没有这样做，因为这些可能会掩盖方法的缺点。平均识别率反映了评估结果的性能：SURF-128 平均识别率是 85.7%，U-SURF 是 83.8%，SURF 是 82.6%。其它描述子，GLOH 是 78.3%，SIFT 是 78.1%，PCA-SIFT 是



图 5 其中一个例子，左面是参考集，右面是测试集。注意两幅图的视觉和颜色不同。

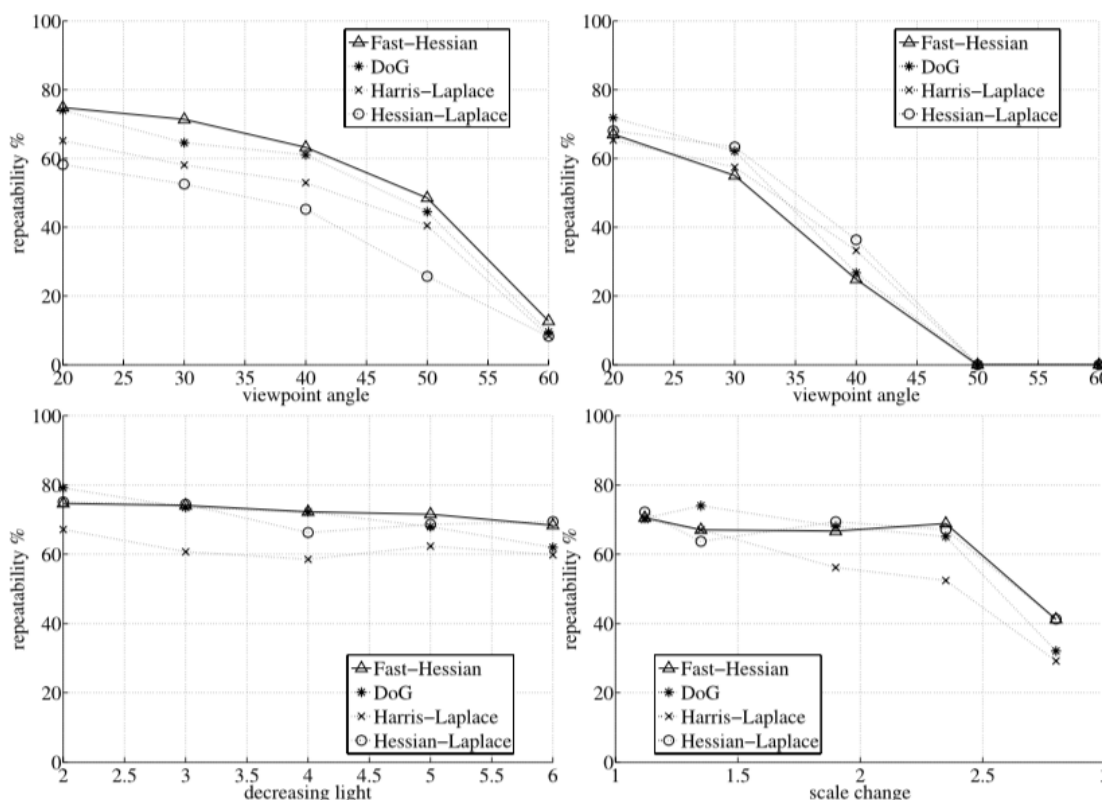


图 6 图像序列的可重复性分数，从左到右，从上到下分别是 Wall 和 Graffiti（角度变化），Leuven（亮度变化）和 Boat（缩放和旋转）。

## 6 结论

我们提出了一个快速的和表现不错的感兴趣点检测器和描述子的方案，在速度和准确度上超越了目前最优秀的方法。描述子很容易扩展成支持放射变换不变性。未来的工作将致力于提高优化代码来提高速度，一个最新的版本的二进制代码已经放在互联网上<sup>2</sup>。

致谢：作者感谢 Swiss SNF NCCR project IM2，Toyota-TME 和 Flemish Fund for Scientific Research 的支持。

<sup>2</sup> <http://www.vision.ee.ethz.ch/~surf/>

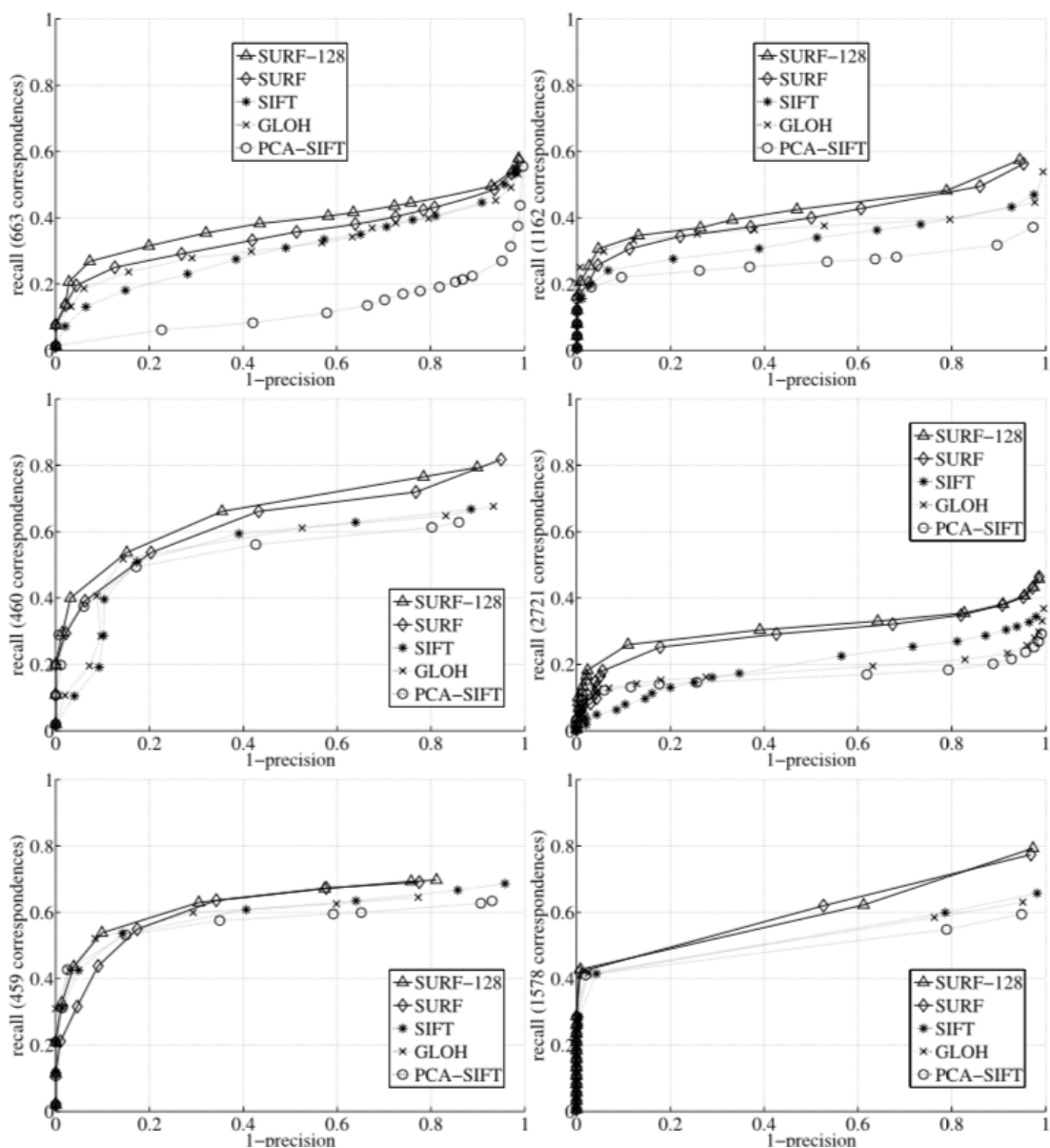


图 7 从左到右，从上到下依次是视觉改变 50° (Wall)、2 个尺度因子 (Boat)、图像模糊 (Bikes 和 Trees)、亮度变化 (Leuven)、和 JPEG 压缩 (Ubc)

## 参考文献

- [1] Lindeberg, T.: Feature detection with automatic scale selection. IJCV 30(2) (1998) 79 – 116
- [2] Lowe, D.: Distinctive image features from scale-invariant keypoints, cascade filtering approach. IJCV 60 (2004) 91 – 110
- [3] Mikolajczyk, K., Schmid, C.: An affine invariant interest point detector. In: ECCV. (2002) 128 – 142
- [4] Ke, Y., Sukthankar, R.: PCA-SIFT: A more distinctive representation for local image descriptors. In: CVPR (2). (2004) 506 – 513

- [5] Tuytelaars, T., Van Gool, L.: Wide baseline stereo based on local, affinely invariant regions. In: BMVC. (2000) 412 – 422
- [6] Matas, J., Chum, O., M., U., Pajdla, T.: Robust wide baseline stereo from maximally stable extremal regions. In: BMVC. (2002) 384 – 393
- [7] Mikolajczyk, K., Schmid, C.: A performance evaluation of local descriptors. In: CVPR. Volume 2. (2003) 257 – 263
- [8] Mikolajczyk, K., Schmid, C.: A performance evaluation of local descriptors. PAMI 27 (2005) 1615–1630
- [9] Mikolajczyk, K., Tuytelaars, T., Schmid, C., Zisserman, A., Matas, J., Schaffalitzky, F., Kadir, T., Van Gool, L.: A comparison of affine region detectors. IJCV 65 (2005) 43–72
- [10] Harris, C., Stephens, M.: A combined corner and edge detector. In: Proceedings of the Alvey Vision Conference. (1988) 147 – 151
- [11] Mikolajczyk, K., Schmid, C.: Indexing based on scale invariant interest points. In: ICCV. Volume 1. (2001) 525 – 531
- [12] Lowe, D.: Object recognition from local scale-invariant features. In: ICCV. (1999)
- [13] Kadir, T., Brady, M.: Scale, saliency and image description. IJCV 45(2) (2001) 83 – 105
- [14] Jurie, F., Schmid, C.: Scale-invariant shape features for recognition of object categories. In: CVPR. Volume II. (2004) 90 – 96
- [15] Mikolajczyk, K., Schmid, C.: Scale and affine invariant interest point detectors. IJCV 60 (2004) 63 – 86
- [16] Florack, L.M.J., Haar Romeny, B.M.t., Koenderink, J.J., Viergever, M.A.: General intensity transformations and differential invariants. JMIV 4 (1994) 171–187
- [17] Mindru, F., Tuytelaars, T., Van Gool, L., Moons, T.: Moment invariants for recognition under changing viewpoint and illumination. CVIU 94 (2004) 3–27
- [18] Baumberg, A.: Reliable feature matching across widely separated views. In: CVPR. (2000) 774 – 781
- [19] Schaffalitzky, F., Zisserman, A.: Multi-view matching for unordered image sets, or “How do I organize my holiday snaps?”. In: ECCV. Volume 1. (2002) 414 – 431
- [20] Freeman, W.T., Adelson, E.H.: The design and use of steerable filters. PAMI 13 (1991) 891 – 906

- [21] Carneiro, G., Jepson, A.: Multi-scale phase-based local features. In: CVPR (1). (2003) 736 – 743
- [22] Se, S., Ng, H., Jasiobedzki, P., Moyung, T.: Vision based modeling and localization for planetary exploration rovers. Proceedings of International Astronautical Congress (2004)
- [23] Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. In: CVPR (1). (2001) 511 – 518
- [24] Koenderink, J.: The structure of images. Biological Cybernetics 50 (1984) 363 – 370
- [25] Lindeberg, T.: Discrete Scale-Space Theory and the Scale-Space Primal Sketch, PhD, KTH Stockholm,. KTH (1991)
- [26] Lindeberg, T., Bretzner, L.: Real-time scale selection in hybrid multi-scale representations. In: Scale-Space. (2003) 148–163
- [27] Brown, M., Lowe, D.: Invariant features from interest point groups. In: BMVC. (2002)