

# 分层结构直方图及其应用

余旺盛,李卫华,侯志强

(空军工程大学信息与导航学院,陕西西安 710077)

**摘要:** 针对传统直方图特征存在的特征分辨能力不强的问题,提出了一种图像分层结构直方图. 该特征首先将图像按照亮度幅值的大小进行分层,然后根据预先设计的结构图元对所有的分层图进行基于结构图元的直方图统计,最后将所得的直方图进行集成得到最终的分层结构直方图. 以图像匹配与视觉跟踪为例,对分层结构直方图的应用进行了大量的仿真实验,结果表明该特征较参考特征具有更强的特征分辨性能和局部描述能力. 利用分层结构直方图进行图像匹配得到的相似性度量图单峰特性更加明显,能够显著降低视觉跟踪算法的跟踪误差.

**关键词:** 图像匹配; 视觉跟踪; 直方图特征; 分层结构直方图

**中图分类号:** TP751      **文献标识码:** A      **文章编号:** 0372-2112 (2017)11 -2617-08

**电子学报 URL:** <http://www.ejournal.org.cn>      **DOI:** 10.3969/j.issn.0372-2112.2017.11.007

## Hierarchical Structure Histogram and Its Application

YU Wang-sheng, LI Wei-hua, HOU Zhi-qiang

(Information and Navigation College, Air Force Engineering University, Xi'an, Shaanxi 710077, China)

**Abstract:** Aimed to the poor discrimination of the traditional histogram feature, a Hierarchical Structure Histogram (HSH) is proposed in this paper. Firstly, it transforms the target image into hierarchical sub images according to the intensity value. Then, it calculates the histograms based on the structure image element for each hierarchical sub image according to the predefined structure image element. Finally, it integrates the histograms to obtain the final hierarchical structure feature. Take image matching and visual tracking as two representative examples, a lot of experiments were carried out based on the proposed feature. The experimental results indicate that the proposed feature owns better discriminative performance and local description performance than the referenced features. The similarity map obtained from the HSH-based image matching has a much more discriminative single-humped characteristic. The HSH feature can also dramatically decrease the tracking errors for visual tracking.

**Key words:** image matching; visual tracking; histogram feature; hierarchical structure histogram

## 1 引言

图像特征是图像信息在特征空间中的描述形式,是图像处理领域的重要研究内容<sup>[1]</sup>. 图像特征的选择在很大程度上决定了图像处理算法性能的上限. 例如在图像匹配算法中,好的图像特征能够提高匹配精度,减小误匹配的概率. 常用的图像特征主要有颜色特征、纹理特征、轮廓特征和结构特征等. 颜色直方图是一种典型的颜色特征,其利用图像的全局统计颜色信息对图像进行描述,对图像的尺度、旋转变化具有一定的鲁棒性. 然而,颜色直方图仅利用了图像分布的概率统计信息,对图像的纹理、轮廓以及结构特征的分辨性不强,

直接导致匹配精度降低. 为提高特征对纹理图像的分辨能力,Ojala 等人提出了局部二值化模式(Local Binary Pattern, LBP)<sup>[2]</sup>. 该特征通过一种局部二值化的方式来描述图像的纹理特性,在人脸检测与识别问题上取得了非常好的效果. 为充分利用边缘及轮廓信息, Dalal 等人提出了梯度方向直方图(Histogram of Orientated Gradient, HoG)<sup>[3]</sup>. 该特征通过分块统计局部梯度方向信息并进行综合,提高了特征的描述能力,在行人检测问题上取得了较好的应用. 针对传统直方图结构信息缺失的问题, Birchfield 等人提出了空间直方图(Spatial histogram, Spatiogram)<sup>[4]</sup>. 该特征将统计直方图与矩特征结合起来,增强了特征的分辨性能,在视觉跟踪中得到了

有效应用. 以尺度不变特征变换 (Scale Invariant Feature Transform, SIFT)<sup>[5]</sup> 为代表的局部特征描述子在图像拼接、三维重构等领域发挥了十分重要的作用. 文献[6]提出一种基于灰度序的局部特征描述子. 该特征采用基于点的灰度序进行特征汇聚, 并通过多支撑区域来提高特征的鲁棒性. 实验结果表明该特征优于 SIFT 特征. 与上述人工构造的图像特征相比, 近年来, 随着深度学习的日益火热, 许多学者开始关注由卷积神经网络 (Correlation neural networks, CNN)<sup>[7]</sup> 自动学习的图像特征. 这些特征在图像查询<sup>[8]</sup>、视觉跟踪<sup>[9]</sup> 等领域取得了非常成功的应用.

本文主要研究如何提高图像特征分辨性的问题. 受梯度方向直方图和空间直方图的启发, 本文提出一

种分层结构化特征, 通过设计结构图元, 分层计算基于结构图元的统计特征, 最后集成为直方图的表现形式. 这里的分层思想与文献[6]中基于灰度序的思想类似. 不同的是, 本文在分层的基础上通过结构图元提取每一个分层图的结构特征, 进一步增强了特征的分辨性.

## 2 分层结构直方图

所谓分层结构直方图, 是指将图像分为不同的图层, 然后统计不同分层对应的空间结构分布, 以充分利用图像的结构信息.

### 2.1 分层结构直方图提取

分层结构直方图提取的原理与过程如图 1 所示. 下面介绍其中的几个关键环节.

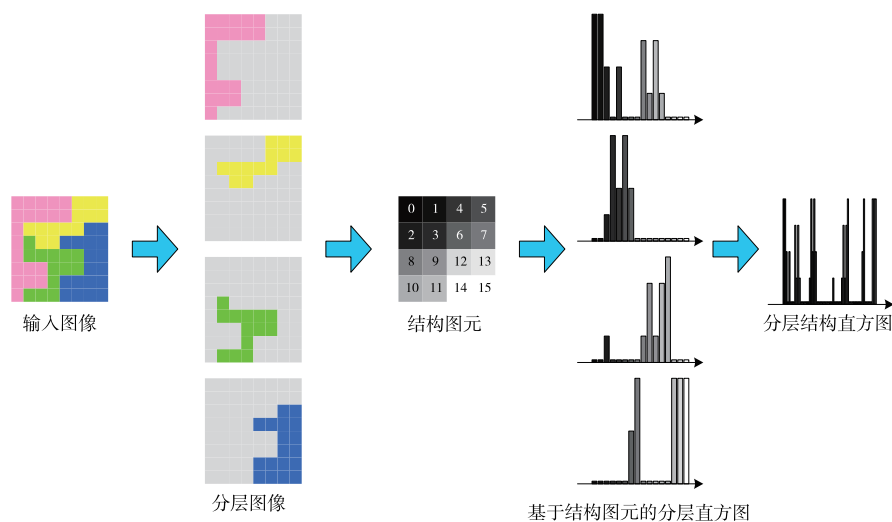


图1 分层结构直方图原理及提取过程示意图

#### (1) 图像分层

图像分层是指将原始输入的图像按照不同的特征属性分解为相对独立的多个图层. 设图像的灰度跨度为  $R (R = I_{\max} - I_{\min} + 1)$ ,  $I_{\max}$  和  $I_{\min}$  分别为图像中灰度的最大值和最小值. 若将图像分解为  $K$  个分层, 则其中第  $k$  个分层为:

$$f_k = \left\{ (x, y) \mid \frac{R}{K} \cdot (k-1) \leq I(x, y) < \frac{R}{K} \cdot k \right\} \quad (1)$$

上式中,  $I(x, y)$  表示点  $(x, y)$  对应的灰度值.

对于彩色图像的分层, 原理与公式(1)类似, 不同的是对不同通道的图像进行分层. 在本文中, 我们将所有彩色图像转化到 HSI 空间, 然后分别将  $H$  通道分解为  $K_1$  层,  $S$  通道分解为  $K_2$  层,  $I$  通道分解为  $K_3$  层, 对应彩色图像的总分层数目为  $K_1 + K_2 + K_3$ .

分层数目的设定, 需要同时考虑特征的分辨性和计算效率. 适当增大数值, 可以提高特征的分辨性, 但会降低计算效率; 数值取得太小则会降低特征的分辨性. 综合上述因素, 本文对灰度图像设定为  $K = 8$ , 对彩

色图像设定为  $K_1 = 8, K_2 = K_3 = 4$ .

得到分层图像以后, 我们用“1”表示属于当前分层图像的像素 (图 1 中第 2 列彩色方格所示), “0”表示不属于当前分层图像的像素 (图 1 中第 2 列灰色方格所示).

#### (2) 结构图元设计

结构图元是预先设计的用于表征像素空间分布情况的模板图像. 简单起见, 这里以边长为 4 的结构图元来介绍其设计过程.

首先将图元分为上下左右对称的四个区域, 按照“0”“1”编码的规则, 上半区为“0”, 下半区为“1”, 左半区为“0”, 右半区为“1”. 编码后得到“00”, “01”, “10”, “11”四个区域. 对于每一个区域, 按照上述区间划分和编码办法, 即可得到从“0000”到“1111”共 16 个二进制编码数字. 将这些二进制编码数字转化为十进制, 即得到最终结构图元中每一个元素对应的数值. 结构图元的具体设计过程如图 2 所示.

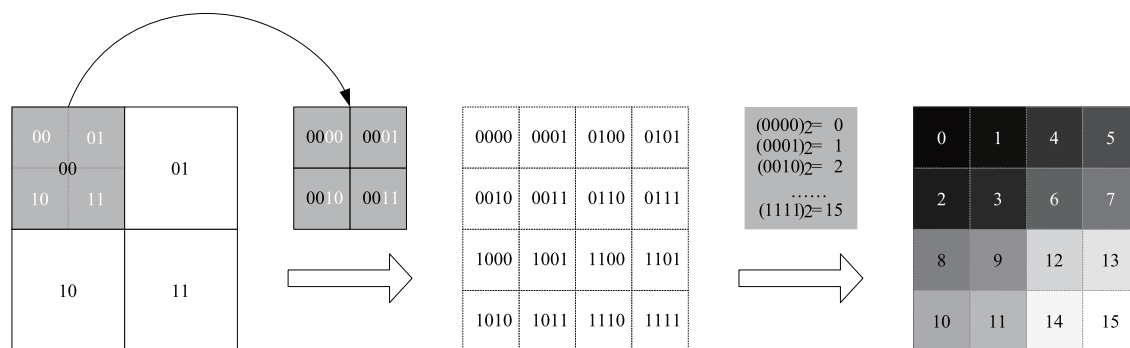


图2 结构图元的设计过程示意图

在实际应用中,结构图元的大小需要与待计算特征的图像大小相匹配,可将图像划分为边长为4的正方形网格,然后采用上述规则进行编码。

### (3) 特征提取

获取输入图像的分层图像之后,对应于每一幅分层图像,将其与结构图元重叠,统计分层图像中元素“1”对应的结构图元中的所有结构元素的直方图。如图1中所示,第一个分层图像中元素“1”为粉色所示区域。统计这些区域所对应的结构图元中结构元素的直方图,得到该分层图像的直方图数据为“4420200031310000”。当所有分层图像均完成基于结构图元的直方图统计处理后,将所得直方图数据组成一个新的直方图数据,即为最终的分层结构直方图。

## 2.2 分层结构直方图分析

与传统的颜色直方图特征相比,分层结构直方图是基于结构图元的像素空间结构分布统计,因此分层结构直方图具有更强的分辨性能。尤其对于像素频率分布相同,但空间分布不同的图像,分层结构直方图能够很容易地将其分辨开来。

LBP 侧重于图像的局部纹理描述,虽然可以通过调整圆周的半径来扩大纹理描述的尺度,但对图像的结构特征描述能力很弱。HoG 通过图像分块之后统计局部范围内的梯度方向信息,然后综合集成最终的直方图特征。该特征对图像的显著边缘以及轮廓特征具有很强的描述能力,但对于结构变化比较复杂的图像,其分辨性能会受到影响。SpatioGram 在传统直方图中加入了像素的空间分布信息。其利用矩特征对像素的空间分布进行描述,但矩特征的空间描述能力较弱。文献[6]中采用灰度序对支撑区域进行划分,利用了像素的空间结构信息。通过样本点与特征点之间的坐标位置关系来构建旋转不变特征,对旋转变化的鲁棒性。

分层结构直方图通过将图像分层和基于结构图元的直方图统计,不仅考虑了像素的频率分布与空间分布信息,增强了特征的分辨性,而且沿袭了传统直

方图特征的表现形式,使得计算与相似性度量更加简单。

## 3 分层结构直方图的应用

作为一种基于区域的特描述符,分层结构直方图具有非常广泛的适用领域。下面以图像匹配和视觉跟踪为例来验证其有效性与优越性。

### 3.1 图像匹配

#### (1) 实验设计

实验思路为已知模板图像,通过穷举搜索的方式确定模板图像在原始图像中的匹配位置。所采用的对照特征为 Histogram 特征, LBP 特征, HoG 特征以及 SpatioGram 特征。本文特征记为 HSH。

#### (2) 结果分析

图3所示为图像匹配所得的相似性分布图。比较可知, Histogram 匹配所得的峰值不明显,致使真实结果容易受到干扰。LBP 在 Histogram 的基础上有明显改进,所得的匹配峰值比较明显。除第3幅图像中的行人目标外, HoG 特征的图像效果并不十分理想。SpatioGram 特征的效果改善比较明显,所得的相似性分布图中峰值更加明显。与上述特征比较, HSH 匹配所得结果的峰值最为明显。

图4所示为测试图像1在峰值点沿  $x$  轴和  $y$  轴的剖面图(考虑论文篇幅, 仅以图像1为例进行说明)。由图可明显看出本文所提出的分层结构直方图具有分辨性最强的峰值。为进一步量化比较各种特征匹配结果的峰值分辨性强度, 采用 Liu 等人<sup>[10]</sup>提出的度量指标 PSR (peak-to-sidelobe ratio) 来进行度量。其计算公式如下:

$$\text{PSR} = \frac{\max(\mathbf{S}) - \mu(\mathbf{S})}{\sigma(\mathbf{S})} \quad (2)$$

式中,  $\mathbf{S}$  表示匹配所得的相似性分布矩阵,  $\max(\cdot)$ 、 $\mu(\cdot)$ 、 $\sigma(\cdot)$  分别表示计算矩阵的最大值、均值和均方差。

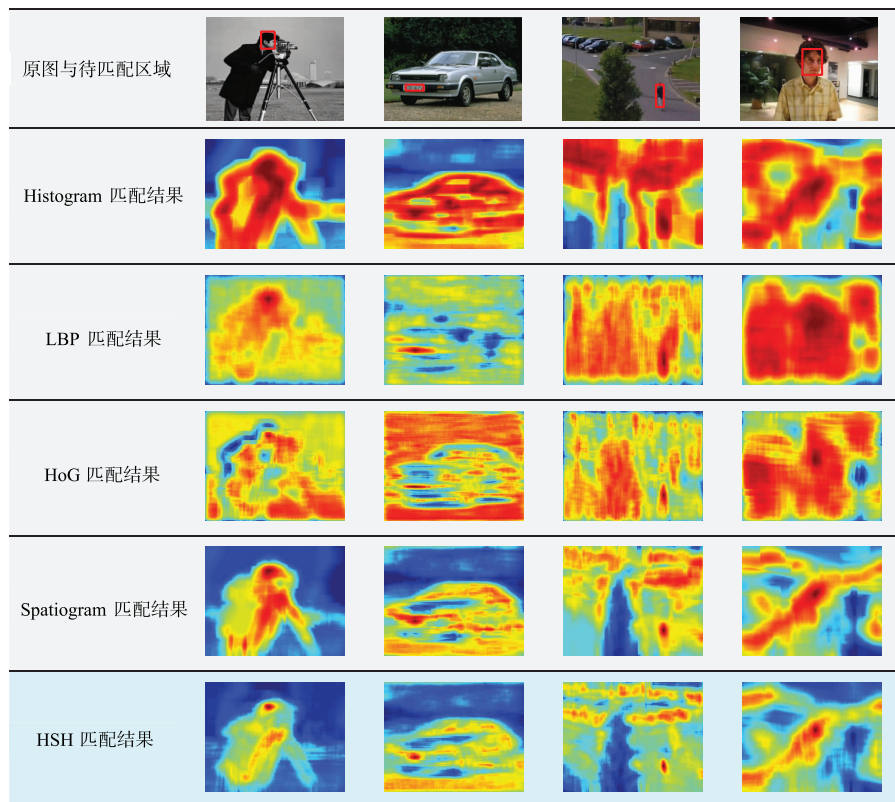


图3 图像匹配所得相似性分布图比较

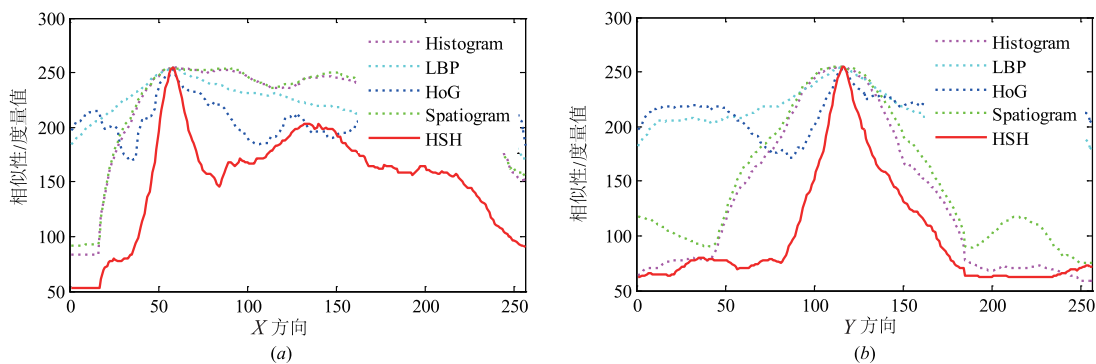


图4 图3中第1幅图像匹配所得相似性分布沿峰值处的切面图

表1 图像匹配所得相似度分布图的PSR值比较

测试图像	Histogram	LBP	HoG	Spatiogram	HSH
图像1	1.6786	2.1926	2.8289	2.6711	3.6739
图像2	1.6909	2.9487	1.6709	2.1883	4.5074
图像3	1.4397	2.4698	2.9913	2.3555	3.4121
图像4	1.6284	1.4437	1.9746	2.5042	3.1085

通过表1中数据可以看出,LBP、HoG以及Spatiogram三种特征均在Histogram特征的基础上有不同程度的改进,所得的PSR值在绝大多数情况下都得到明显的提高.HSH在所有实验结果中均得到了最大的

PSR值,表明利用HSH进行图像匹配能够得到分辨性更强的结果,峰值特性更加明显,这对提高图像匹配的精度以及减小误匹配均具有非常重要的意义.

### 3.2 视觉跟踪

#### (1) 实验设计

本文所提特征的另一个重要应用领域是视觉跟踪.在基于粒子滤波的视觉跟踪中,首先需要将被跟踪目标用图像特征进行描述.为证明本文所提分层结构直方图的有效性 with 优越性,在粒子滤波算法框架下<sup>[11]</sup>进行了大量视觉跟踪实验,采用的参考对照特征为Histogram,LBP,HoG以及Spatiogram.测试视频来自于通用测试平台<sup>[12]</sup>,跟踪时粒子采样数目统一设为300.

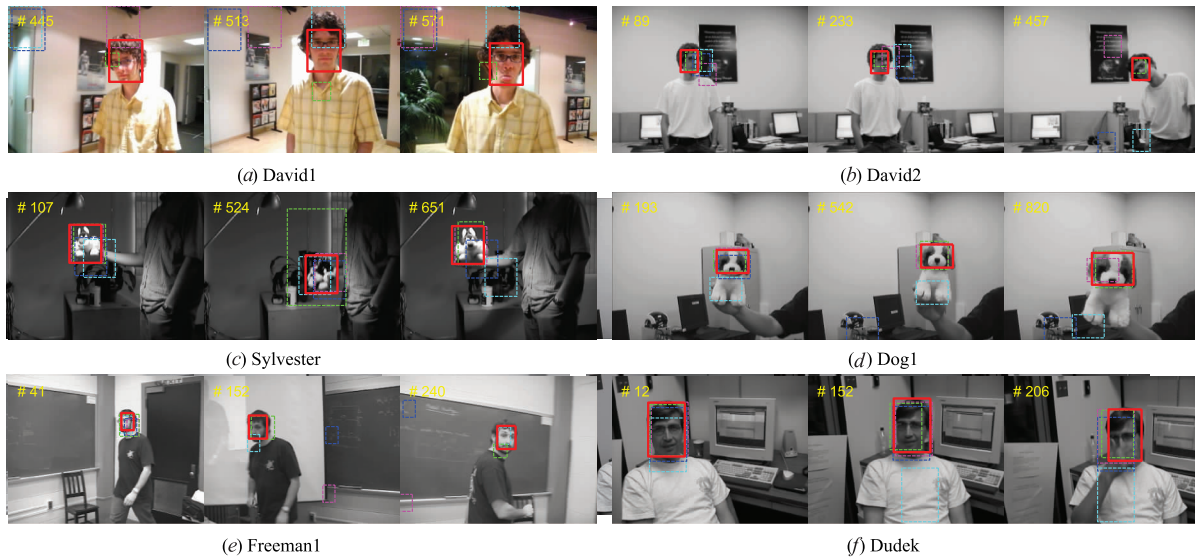


图5 视觉跟踪的部分实验结果截图(不同颜色代表的特征见图4)

## (2) 结果分析

图5所示为几组具有代表性的视频序列跟踪实验结果. 图5(a)为“David1”序列, 跟踪目标为人脸. 由于目标在相邻帧之间发生较小的光照改变, 导致参考特征跟踪逐渐发生偏移, 最终无法实现对目标的有效跟踪. 图5(b)为“David2”序列, 被跟踪目标的难点在于存在颜色分布与目标相似的干扰区域. 由图457帧可以看出, HSH 较参考特征能够更加准确地锁定跟踪区域. 图5(c)和图(d)跟踪的分别为“Sylvester”序列和“Dog1”序列. 前者的难点是存在较大的姿态和光照变化, 后者的难点是目标尺度变化. 由实验结果可以看出, Spatiogram 和 HSH 跟踪所得的结果更加理想. 图5(e)和图(f)中, HSH 亦取得了相对更优的跟踪结果.

为进一步对跟踪结果进行评价, 选用了中心位置误差指标对跟踪结果进行定量分析. 中心位置误差定义为跟踪结果的中心位置与目标真实位置之间的误差, 单位为像素. 通常而言, 中心位置误差小, 表明跟

踪的精度越高, 性能越好. 图6给出了5种特征跟踪所得的中心位置误差曲线的比较. 由图可以看出, HSH 跟踪所得的中心位置误差曲线在绝大部分情况下均位于最下方, 表明该特征跟踪所得的结果更加准确.

表2给出了中心位置误差的统计数据. 均值越小, 表明跟踪的精度越高; 方差越小, 表明跟踪的结果越稳定. 对比可知, 传统的 Histogram 特征的性能最差, Spatiogram 特征能够获得相对较好的跟踪性能, 但较 HSH 仍然存在一定的差距.

综合上述分析可知, 本文所提特征在视觉跟踪中较参考特征有较为明显的优势.

### 3.3 局部特征描述

与 SIFT 特征以及文献[6]中所提出的 MROGH (Multi-support Region Order-Based Gradient Histogram) 特征不同的是, 上述应用主要是利用了分层结构直方图的全局特征描述能力. 事实上, 分层结构直方图也可以用于局部特征描述.

表2 五种特征跟踪所得中心位置误差的均值和方差比较

Sequences	Histogram		LBP		HoG		Spatiogram		HSH	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
David1	98.6	71.1	96.2	44.6	173.0	59.3	28.5	7.5	7.7	5.4
David2	53.6	45.1	88.7	75.2	77.7	63.3	5.3	2.3	3.0	1.5
Sylvester	17.6	27.4	57.8	26.2	35.3	36.5	12.9	16.4	6.3	9.1
Dog1	44.8	68.8	83.1	47.8	121.2	68.0	8.1	8.4	6.7	8.4
Freeman1	216.6	92.0	27.1	37.1	144.6	112.4	12.2	8.2	4.5	3.7
Dudek	76.2	76.2	313.4	313.4	163.9	163.9	11.0	11.0	2.8	2.8

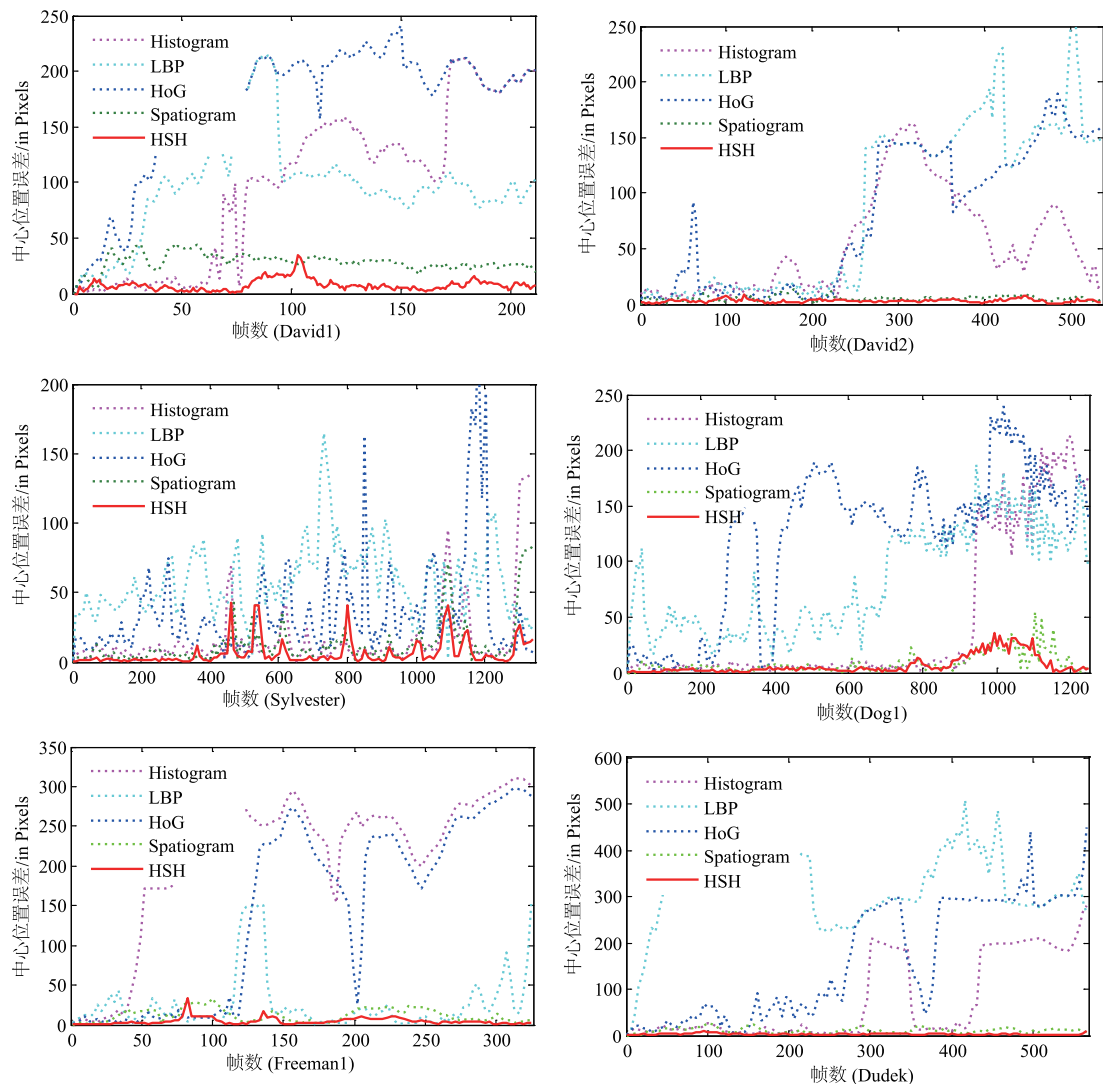


图6 五种特征跟踪所得的中心位置误差比较

### (1) 实验设计

为验证分层结构直方图对局部图像特征的描述能力,设计实验方案如下:

第一步:对输入的图像  $A$  和图像  $B$  分别提取 Hessian-Affine 区域.

第二步:利用特征描述子对 Hessian-Affine 区域进行特征描述.

第三步:对图像  $A$  和图像  $B$  进行特征匹配,分别记录返回匹配结果和正确匹配结果.

第四步:计算 recall 和 1-precision,并绘制曲线进行比较.

Hessian-Affine 区域的具体提取方法以及特征匹配策略请参考文献[13]. 参考对照的局部特征描述子为 SIFT 和 MROGH. 本文的 HSH 与 MROGH 特征类似, Hessian-Affine 区域均归一化为边长为 21 像素的矩形区域.

### (2) 结果分析

实验采用的是测试图片如图 7 所示,图片来源于 Oxford Dataset<sup>[14]</sup>. 三种局部特征描述子的实验结果的如图 8 所示. 在“recall-precision”图中,位于左上方的曲线代表性能更优. 由图可知, HSH 在图像模糊以及光照变化的情况下取得了与 MROGH 相当或者更优的性能. 在尺度和旋转情况下, MROGH 特征的性能最优,但 HSH 的性能仍然明显优于 SIFT 特征.

需要指出的是,本文算法并未有进行专门针对抗旋转变化的设计. 如果专门针对旋转情况对 HSH 进行优化,性能应当会得到进一步提升.

### 3.4 分析与讨论

在本文所讨论的应用中,分层结构直方图在提高匹配精度的同时能够减小误匹配发生的概率. 在视觉跟踪中,分层结构直方图显著降低了跟踪的中心位置误差. 需要指出的是,本文在跟踪实验中未进行模型更



图7 局部特征描述实验所采用的测试图像

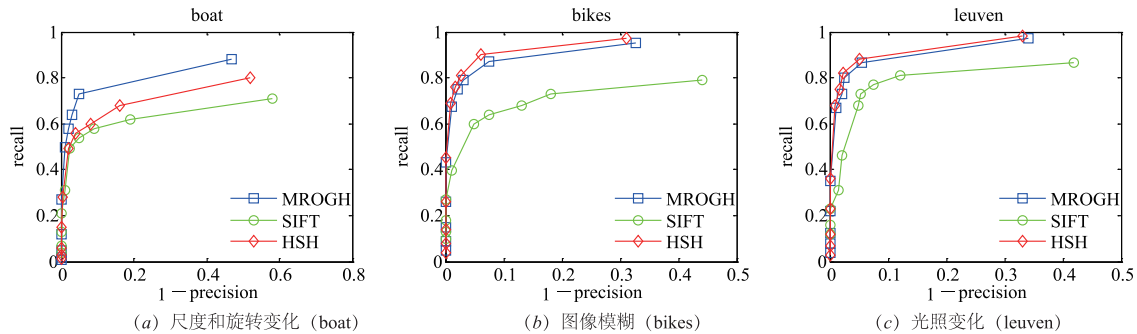


图8 基于Hessian-Affine区域的局部特征描述实验结果比较

新处理. 除上述应用外, 本文特征还可以应用于图像检索、特征匹配、图像配准等领域. 与关键点或关键区域检测算子结合, 本文特征还可以用于局部特征描述.

实验中发现, 分层结构直方图的计算时间开销与所提取图像区域的大小相关, 当图像区域较大时, 提取的效率会降低. 虽然较 LBP、HoG 以及 Spatiogram 等特征的计算速度快, 但尽管如此, 分层结构直方图在某些应用场合实现实时处理还存在一定的困难, 因此研究一种快速提取分层结构直方图的方法非常关键.

#### 4 结论

本文提出了一种基于结构图元统计分层图像结构信息的分层结构直方图, 并研究了其在图像匹配、视觉跟踪和局部特征描述中的应用. 实验结果表明, 该特征较目前一些常用特征的分辨性能更优, 匹配所得的相似性分布图的峰值更加明显, 显著减小了视觉跟踪中的跟踪误差. 后续工作中, 我们将进一步研究该特征在其他领域的应用.

#### 参考文献

- [1] 余旺盛, 等. 基于区域边缘统计的图像特征描述新方法 [J]. 计算机学报. 2014, 37(6): 1398 - 1410.  
W S Yu, et al. A new image feature descriptor based on region edge statistical [J]. Chinese Journal of Computers. 2014, 37(6): 1398 - 1410. (in Chinese)
- [2] T Ojala, et al. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2002, 24(7): 972 - 987.
- [3] N Dalal, et al. Histograms of oriented gradients for human

detection [A]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition [C]. San Diego, USA; IEEE Press, 2005. 20 - 25.

- [4] S T Birchfield, et al. Spatiograms versus histograms for region-based tracking [A]. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition [C]. San Diego, USA; IEEE Press, 2005. 1158 - 1163.
- [5] D Lowe. Distinctive image features from scale-invariant keypoints [J]. International Journal of Computer Vision. 2004, 60(2): 91 - 110.
- [6] B Fan, et al. Rotationally invariant descriptors using intensity order pooling [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(10): 2031 - 2045.
- [7] C Ma, et al. Hierarchical convolutional features for visual tracking [A]. Proceedings of the IEEE International Conference on Computer Vision [C]. Santiago, Chile; Springer-Verlag, 2015. 3074 - 3082.
- [8] F Zhao, et al. Deep semantic ranking based hashing for multi-label image retrieval [A]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition [C]. Boston, USA; IEEE Press, 2015. 1556 - 1564.
- [9] K Zhang, et al. Robust visual tracking via convolutional networks without training [J]. IEEE Transactions on Image Processing, 2016, 25(4): 1779 - 1792.
- [10] T Liu, et al. Real-time part-based visual tracking via adaptive correlation filters [A]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015 [C]. Boston, USA; IEEE Press, 2015. 4902 - 4912.
- [11] P Perez, et al. Color-based probabilistic tracking [A]. In: Proceedings of the European Conference on Computer Vision [C]. Berlin, Germany; Springer-Verlag, 2002. 661

-75.

[12] Y Wu, et al. Object tracking benchmark[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2015, 37(9):1834-1848.

[13] K Mikolajczyk, et al. A performance evaluation of local

descriptors[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2005, 27(10):1615-1630.

[14] B Fan, et al. Receptive fields selection for binary feature description[J]. IEEE Transactions on Image Processing. 2014, 23(6):2583-2595.

#### 作者简介



**余旺盛(通信作者)** 男,1985年出生于湖南平江,2014年获空军工程大学信息与导航学院信息与通信工程专业博士学位,现为空军工程大学信息与导航学院讲师,主要研究方向为图像处理、计算机视觉与模式识别等。

E-mail: xing\_fu\_yu@sina.com



**李卫华** 男,1964年出生于河南郑州,2006年获西北工业大学导航、制导与控制专业博士学位,现为空军工程大学信息与导航学院教授,主要研究领域为指挥控制与信息融合技术、图像处理与人工智能技术等。

E-mail: li\_wei\_hua@yahoo.com.cn